Probability in Action

edited by

Zbigniew A. Łagodowski

Politechnika Lubelska

Lublin 2015

Probability in Action

Volume 2

Monografie – Politechnika Lubelska

Politechnika Lubelska
Wydział Elektrotechniki i Informatyki
ul. Nadbystrzycka 38A
20-618 Lublin

Probability in Action
Volume 2

edited by
Zbigniew A. Łagodowski

Politechnika Lubelska
Lublin 2015

Reviewer:
prof. dr hab. Jurij Kozicki, Maria Curie Skłodowska University

LaTeX Editor: Adam Gregosiewicz

Publication approved by the Rector of Lublin University of Technology

ISBN: 978-83-7947-173-7

Publisher: Lublin University of Technology
ul. Nadbystrzycka 38D, 20-618 Lublin, Poland
Realization: Lublin University of Technology Library
ul. Nadbystrzycka 36A, 20-618 Lublin, Poland
tel. (81) 538-46-59, email: wydawca@pollub.pl
www.biblioteka.pollub.pl
Printed by : TOP Agencja Reklamowa Agnieszka Łuczak
www.agencjatop.pl

The digital version is available at the Digital Library of Lublin University of Technology: www.bc.pollub.pl
Impression: 100 copies

Contents

Preface

We present to the readers the second volume of the series "**Probability in Action**". The motivation to launch this series was presented in the first volume and has not since changed. In view of the good reception of the first publication, the second is written in the same style. The main aim is the presentation in book form of the current research of scientists at the Lublin University of Technology (some chapters have co-authors from other cooperating academic institutions) in the field of probability theory and its applications. Another objective is an offer and an invitation to collaboration to a very broad group of specialists representing pure and applied mathematics, mathematical biology, statistics, engineering, economy and social sciences.

The book is organized as a series of nine research articles.

Zbigniew A. Łagodowski

Averaging discrete-time signals having finite energy

Tadeusz Banek[1], **August Zapała**[2]

Abstract

The general method for averaging functionals of any discrete−time signals having finite energy is presented. The method uses Banach's general concept of the Lebesgue integration in abstract spaces, which is restricted here to the separable Hilbert space. The described method allows us to evaluate any characteristic which can be expressed as a function of data. In addition to the integral representation we offer a numerical Monte Carlo type integration procedure which is of independent interest.

1. Introduction

Assuming that signals emitted by a finite number of sources have finite total energy in any bounded time interval, we propose an averaging procedure based on Banach's concept of the Lebesgue integral in abstract spaces [1]. More precisely, we are going to use a particular variant of Banach's theory, namely those which is connected with integration in the infinite dimensional separable Hilbert space, i.e. the space

$$\ell^2 = \left\{ x = (x_1, x_2, ...) \in \mathbb{R}^\infty : \sum_{n=1}^{\infty} x_n^2 < \infty \right\},$$

of all sequences of real numbers having sum of squares finite. By $S_n(r)$ we denote the set

$$S_n(r) = \left\{ x = (x_1, x_2, ..., x_n, 0, ...) \in \ell^2 : \sum_{k=1}^{n} x_k^2 \leq r^2 \right\} \subset \ell^2,$$

[1]Lublin Technical University, Faculty of Management, Nadbystrzycka 38A, 20-618 Lublin, Poland, Polish Air Force Academy, Aeronautics Faculty, Department of Avionics and Control Systems, *e-mail: kenabt@gmail.com*

[2]The John Paul II Catholic University of Lublin, Faculty of Mathematics, IT and Landscape Architecture, *e-mail: azapala@kul.pl*

and put

$$S(r) = \overline{\bigcup_{n \geq 1} S_n(r)} = \left\{ x \in \ell^2 : \sum_{n=1}^{\infty} x_n^2 \leq r^2 \right\} \subset \ell^2.$$

According to Banach's [1] Theorem 1, the most general non-negative linear functional F defined on the linear set $\mathfrak{C}$ of real-valued, bounded, Borel measurable functions Φ on the space $S(r)$, satisfying additional conditions (i)–(ii) of Banach's paper, is of the form

$$\begin{aligned} F(\Phi) &= \lim_{n \to \infty} F_n(\Phi), \\ F_n(\Phi) &= \int\limits_{S_n(r)} \frac{\Phi(x_1, ..., x_n, 0, ...)\, dx_1 ... dx_n}{2^n r \sqrt{\left(r^2 - x_1^2\right) \cdot ... \cdot \left[r^2 - \left(x_1^2 + ... + x_{n-1}^2\right)\right]}}, \end{aligned}$$

where $\Phi : S(r) \subset \ell^2 \to \mathbb{R}$ is a bounded, Borel measurable function of infinite number of variables. The mentioned functional F possesses analogous properties as the Lebesgue integral, thus it is called functional integral, or in short $\mathfrak{L}$-integral. Having in mind applications in physics, $S(r)$ may be interpreted as the set of flows or impulses with a finite energy, which stimulate the behaviour of the observed object and exert influence on its numerical characteristic Φ, and *a fortiori* – on $F(\Phi)$. To illustrate the wide spectrum of all the possible forms of Φ we consider a few examples.

Example 1.1. Let $\Phi(x) = P(|x|_2)$, $x \in S(r)$, where P is a real-valued polynomial, and $|\cdot|_2$ is the norm of ℓ^2. More generally, if $\Phi(x) : S(r) \to \mathbb{R}$ is Borel measurable and $|\Phi(x)|$ is bounded from above by a polynomial $P(|x|_2)$, then the Banach $\mathfrak{L}$-integral $F(\Phi)$ is well-defined. In particular, for functions of the form $\Phi(x) = \sin(P(|x|_2))$, $\cos(P(|x|_2))$, $\arctan(P(|x|_2))$, $\operatorname{arccot}(P(|x|_2))$ etc., the Banach $\mathfrak{L}$-integral $F(\Phi)$ can be considered.

Example 1.2. Since for each real-valued polynomial P we have $|P(y)| \leq m < \infty$ for all $0 \leq y \leq r$, where $0 < m < \infty$ denotes a constant, the Banach integrals for $\Phi(x) = \exp\{\pm P(|x|_2)\}$, $\Phi(x) = \log[1 + |P(|x|_2)|]$ etc., are defined as well.

Example 1.3. Let $L(x) = \sum_{k=1}^{\infty} x_{i(k)}^2$ be a lacunary series of $x = (x_1, x_2, ...)$ in ℓ^2, where $1 \leq i(1) < i(2) < ...$ is an arbitrary but fixed sequence of increasing indices. Then clearly $L(x) \leq |x|_2^2$, therefore all the above examples with polynomial $P(|x|_2)$ replaced by $P(L(x))$ provide bounded, Borel mappings on $S(r)$ for which the Banach functional integrals are well-defined.

Example 1.4. Notice that if $x \in S(r)$, then $|x_k| \leq r$ for all $k \geq 1$. Moreover, the projection $\ell^2 \ni x \to x_k \in \mathbb{R}$ is a continuous function of

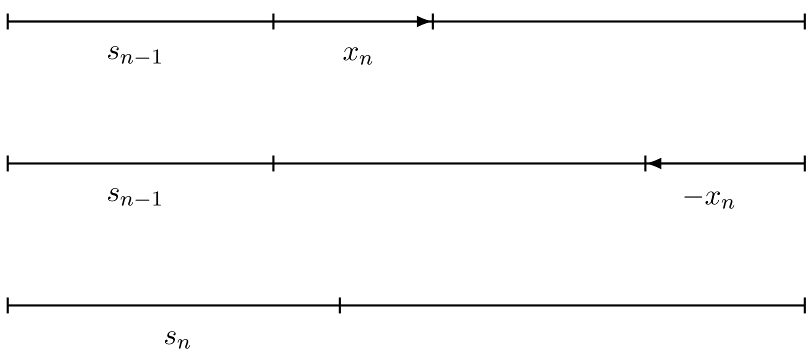
$x \in \ell^2$, and so Borel-measurable. Hence it follows that mappings of the form $f_{k_1}(x) = \sup\left\{x_{i_1(j)},\ 1 \le j \le k_1\right\}$, $g_{k_2}(x) = \sup\left\{\left|x_{i_2(j)}\right|,\ 1 \le j \le k_1\right\}$, $h_{k_3}(x) = \inf\left\{x_{i_3(j)},\ 1 \le j \le k_3\right\}$, $l_{k_4}(x) = \inf\left\{\left|x_{i_4(j)}\right|,\ 1 \le j \le k_4\right\}$, $f_\infty(x) = \sup\left\{x_{i_5(k)}, k \ge 1\right\}$, $g_\infty(x) = \sup\left\{\left|x_{i_6(k)}\right|, k \ge 1\right\}$, $h_\infty(x) = \inf\{x_{i_7(k)}, k \ge 1\}$, or $l_\infty(x) = \inf\left\{\left|x_{i_8(k)}\right|, k \ge 1\right\}$, are bounded and Borel-measurable on $S(r)$. Therefore compositions of any bounded Borel function $\Psi : [-r,r]^s \to \mathbb{R}$, $1 \le s < \infty$, with the last mappings, as well as with $x_i(k)$, $|x|_2$ and $L(x)$, i.e. maps of the form

$$\Psi\left(f_{k_1}, g_{k_2}, h_{k_3}, l_{k_4}, f_\infty, g_\infty, h_\infty, l_\infty, |\cdot|_2, L(\cdot), \pi_{i(1)}, \pi_{i(2)}, ..., \pi_{i(r)}\right)(x),$$

where $\pi_{i(k)}(x) = x_{i(k)}$, $s = r + 10$, are also admissible integrands for the Banach $\mathfrak{L}$-integral.

To describe more precisely everyday life practical applications of the developed here general theory, we sketch some situations when it may be useful.

1. Consider some electrical equipment battery, e.g. laptop, tablet, video camera, smartphone, mobile telephone, GPS etc. Usually the battery needs to be fully recharged when it is quite flat, but sometimes (in random cases) the process of recharging starts before the battery is entirely discharged, and sometimes (in other cases) the process of recharging is not finished as it should be. In such cases the total capacity of the battery is reduced, but the downfall does not change linearly, i.e. we may expect that the remaining part of unattainable capacity after recharging ($= s_n$) is smaller than the sum of unattainable capacity before this process ($= s_{n-1}$) and unused capacity ($= |\pm x_n|$); see the pictures below.



In other words, one may expect that

$$s_n \approx \sqrt{s_{n-1}^2 + x_n^2} \le s_{n-1} + |\pm x_n|.$$

In this model the quantity $\left(\sum_{i=1}^{n} x_i^2\right)^{1/2}$ expresses the unattainable capacity of the battery and when it assumes the maximal possible value r, the battery becames unusable. Notice that the limit point $x = \lim_n \sum_{i=1}^{n} x_i e_i \in \ell^2$ of this process represents the whole history of life of the battery.

2. Suppose that a PC is used by a couple of people and the private files written on the hard disk by one or the other person are denoted by signs $+$,$-$ resp. The files cannot be written linearly one after another, so that the total capacity occupied by a finite number n of files is greater than the sum $|\pm x_1| + |\pm x_2| + \ldots + |\pm x_n|$ of their sizes $|\pm x_i|$. When one performs the backup of the hard disk, say on a DVD-ROM, then the files are compressed in such a way that they occupy less capacity then the sum of their sizes. Thus we may expect that the total capacity of n files written on a DVD-ROM is equal to

$$s_n \approx \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} \leq |\pm x_1| + |\pm x_2| + \ldots + |\pm x_n|.$$

Now when the quantity s_n attains the capacity r of the DVD-ROM, the DVD-ROM is full, the backup is stopped and the DVD-ROM should be changed.

Although our study is well-founded to macroscopic situations, it may be also used for the description and analysis of some phenomena in nuclear physics.

3. Recall that the main ideas of quantum mechanics are the following: the possible (pure) states of a quantum mechanical system (e.g. a particle) are represented by some unit vectors of a complex separable Hilbert space $(H, \|\cdot\|)$, called state space, well defined up to a complex constant c, $|c| = 1$ – the phase factor; physical properties of the system are described by means of wave functions $x(u,t)$ of the position $u \in \mathbb{R}^3$ and time $t \in \mathbb{R}$, taking values in H, such that for a fixed t, $\|x(u,t)\|^2$ is a probability density satisfying condition: the probability that the particle in the state x can be found in a region $\Delta \subset \mathbb{R}^3$ at the moment of time t is equal to

$$P[x, u \in \Delta] = \int_{\Delta} \|x(u,t)\|^2 \, du;$$

each physical quantity, say b, is associated with a (Hermitian) linear operator B acting on H, and all the possible values of b are certain eigenvalues of B. Moreover, if x is expanded in the CONS $\{e_i\}$ consisted of eigenvectors of B, then the process of measurement of it can give merely the value of some coefficient x_i with probability $|x_i|^2$ in the series expansion

$$x = \sum_{i=1}^{\infty} x_i e_i,$$

and then the wave function reduces to $x_i e_i$. In the presented below approach we consider only a real Hilbert space $H = \ell^2$ and we are not interested at all in quantization, thus it is a sort of simplification of the quantum mechanics theory, but on the other hand, contrary to postulates of quantum mechanics where coefficients x_i of vectors $x \in H$ are assumed to be constant (for fixed t), we consider x_i, $i \geq 1$, as random variables. Furthermore, taking $r = 1$ we introduce normalization which entails that

$$\sum_{ni=1}^{\infty} x_i^2 = 1,$$

i.e. $\left\{x_i^2, i \geq 1\right\}$ is a discrete probability distribution with probability 1. The last condition is consistent with requirements of quantum mechanics. Hence our considerations can also shed light on the properties of quantum mechanical systems. More precisely, if the procedure of measurements enables us to obtain consecutive values of coefficients x_i, treated as experimental data of random variables for the flow of homogeneous particles, then one is able to approximate x, and in consequence – the expected value of any bounded numerical characteristic $\Phi(x) = \Phi(x_1, x_2, ...)$ of the wave function x. Therefore one may expect that in such a way the physical quantities like energy, momentum or position of the particle would be estimated (with accuracy which is admissible by Heisenberg's uncertainty principle).

As is clearly visible from the above examples, Φ may have a complicated form, hence analytic integration is rarely possible. Therefore in most situations we have to approximate the proper value of $F(\Phi)$. The first step in this direction requires an approximation of $F(\Phi)$ by means of $F_n(\Phi)$.

2. Approximation of the Banach functional integral with increasing dimension of the space

From the general theory developed in Banach's paper [1] we know that $F(\Phi) = \lim_{n\to\infty} F_n(\Phi)$, but for many practical reasons the most important is the rate in this convergence. Under some realistic restrictions we are able to estimate the rate of convergence of $F_n(\Phi)$ to $F(\Phi)$. Since Φ is a bounded function of $x = (x_1, x_2, ...) \in S(r) \subset \ell^2$, where $\sum_{n=1}^{\infty} x_n^2 \leq r^2 < \infty$, it is natural to demand that there can be found a constant $0 < C < \infty$, such that

$$|\Phi(x_1, ..., x_{n-1}, x_n, 0, ...) - \Phi(x_1, ..., x_{n-1}, 0, ...)| \leq C \cdot x_n^2 \tag{2.1}$$

for all sufficiently large $n \geq n_C$. Obviously, the constant C in (2.1) may be dependent on r, thus for an arbitrary upper bound that can be expressed in terms of $|x_n|^\beta$, $\beta > 0$, we can always write $|x_n|^\beta \leq r^{\beta-2} \cdot x_n^2$ when $2 < \beta < \infty$. On the other hand, if $0 < \beta < 2$, then one can estimate from above the integral of $|x_n|^\beta$ by the integral of x_n^2 using the Hölder inequality. Thus the assumption (2.1) is in fact quite natural.

Lemma 2.1. *If* $\Phi(x_1, ..., x_n, 0, ...) = x_n^2$, *then we have*

$$M_n^2(r) := \int\limits_{S_n(r)} \frac{x_n^2 \quad dx_1 ... dx_n}{2^n r \sqrt{\left(r^2 - x_1^2\right) \cdot ... \cdot \left[r^2 - \left(x_1^2 + ... + x_{n-1}^2\right)\right]}} = \frac{r^2 \cdot 2^{n-1}}{3^n}.$$

Proof. Substitute $x_i = r y_i$ for $i = 1, 2, ..., n$, and note that for $n > 1$,

$$M_n^2(r) = \int\limits_{S_n(1)} \frac{r^2 y_n^2}{2^n \sqrt{\left(1 - y_1^2\right) \cdot ... \cdot \left[1 - \left(y_1^2 + ... + y_{n-1}^2\right)\right]}} dy_1 ... dy_n =$$

$$= r^2 \int\limits_{S_{n-1}(1)} \frac{1}{2^n \sqrt{\left(1 - y_1^2\right) \cdot ... \cdot \left[1 - \left(y_1^2 + ... + y_{n-1}^2\right)\right]}} \cdot$$

$$\cdot \left(2 \int\limits_0^{\sqrt{1 - \left(y_1^2 + ... + y_{n-1}^2\right)}} y_n^2 dy_n \right) dy_1 dy_2 ... dy_{n-1} =$$

$$= \quad r^2 \int\limits_{S_{n-1}(1)} \frac{\left. \frac{y_n^3}{3} \right|_0^{\sqrt{1 - \left(y_1^2 + ... + y_{n-1}^2\right)}} dy_1 dy_2 ... dy_{n-1}}{2^{n-1} \sqrt{\left(1 - y_1^2\right) \cdot ... \cdot \left[1 - \left(y_1^2 + ... + y_{n-1}^2\right)\right]}} =$$

$$= \quad r^2 \int\limits_{S_{n-1}(1)} \frac{\left[1 - \left(y_1^2 + ... + y_{n-1}^2\right)\right] \quad dy_1 dy_2 ... dy_{n-1}}{3 \cdot 2^{n-1} \sqrt{\left(1 - y_1^2\right) \cdot ... \cdot \left[1 - \left(y_1^2 + ... + y_{n-2}^2\right)\right]}} =$$

$$= \quad \frac{r^2}{3} \left\{1 - \left(M_1^2(1) + M_2^2(1) + ... + M_{n-1}^2(1)\right)\right\}.$$

Furthermore,

$$M_1^2(r) = \int\limits_{-r}^{r} \frac{x_1^2}{2r} dx_1 = \left. \frac{x_1^3}{3 \cdot 2r} \right|_{-r}^{r} = \frac{r^2}{3},$$

thus

$$M_2^2(1) = \frac{1}{3} \left\{1 - M_1^2(1)\right\} = \frac{1}{3} \cdot \frac{2}{3},$$

and, by induction,

$$M_k^2(1) = \frac{1}{3}\left\{1 - \left(\frac{1}{3} + \frac{2}{3^2} + \dots + \frac{2^{k-2}}{3^{k-1}}\right)\right\} = \frac{1}{3}\left\{1 - \frac{1}{3} \cdot \frac{1-(2/3)^{k-1}}{1-2/3}\right\}$$
$$= \frac{2^{k-1}}{3^k}.$$

Hence

$$M_n^2(r) = \frac{r^2}{3}\left\{1 - \left(M_1^2(1) + M_2^2(1) + \dots + M_{n-1}^2(1)\right)\right\} =$$

$$= \frac{r^2}{3}\left\{1 - \left(\frac{1}{3} + \frac{2}{3^2} + \dots + \frac{2^{n-2}}{3^{n-1}}\right)\right\} = \frac{r^2 \cdot 2^{n-1}}{3^n},$$

which concludes the proof. □

Theorem 2.2. *Under the assumption (2.1) we have*

$$|F(\Phi) - F_n(\Phi)| \le C \cdot \frac{r^2 \cdot 2^n}{3^n}.$$

Proof. Notice first that

$$|\Phi(x_1, \dots, x_n, \dots, x_{n+N}, 0, \dots) - \Phi(x_1, \dots, x_n, 0, \dots)| \le C \cdot \left(x_{n+1}^2 + \dots + x_{n+N}^2\right).$$

Therefore

$$|F_{n+N}(\Phi) - F_n(\Phi)| \le C \cdot \left(M_{n+1}^2(r) + \dots + M_{n+N}^2(r)\right) \le$$

$$\le C \cdot \frac{r^2}{3}\left(\frac{2^n}{3^n} + \dots + \frac{2^{n+N-1}}{3^{n+N-1}}\right) = C \cdot \frac{r^2}{3} \cdot \frac{2^n}{3^n}\left(\frac{1-(2/3)^N}{1-(2/3)}\right).$$

Passing to the limit as $N \to \infty$ we obtain

$$|F(\Phi) - F_n(\Phi)| = \lim_{N\to\infty} |F_{n+N}(\Phi) - F_n(\Phi)| \le$$

$$\le \lim_{N\to\infty} C \cdot \frac{r^2}{3} \cdot \frac{2^n}{3^n}\left(\frac{1-(2/3)^N}{1-(2/3)}\right) = C \cdot r^2 \cdot \frac{2^n}{3^n},$$

which entails the desired estimate. □

3. Random walk in the ball

As the first step, choose a random point $x_1 = X_1(\omega)$ from $S_1(r) = [-r, r]$, where the interval $[-r, r]$ is equipped with the uniform distribution. Denote by $l_2(x_1)$ a segment of the line orthogonal to the 1-dimensional sphere $S_1(r) = [-r, r]$, which is a chord of $S_2(r)$ that passes by the point x_1. At the second step select randomly a point $x_2 = X_2(\omega)$ from $l_2(x_1)$ according to the uniform distribution on it. Being at x_2, find a segment $l_3(x_2)$ orthogonal to the plane spanned by $l_2(x_1)$ and $S_1(r)$ and crossing $l_2(x_1)$ at x_2, such that $l_3(x_2)$ is a chord of $S_3(r)$. Choose randomly a point $x_3 = X_3(\omega)$ according to the uniform distribution on $l_3(x_2)$, etc. The sequence $(x_1, ..., x_n) = (X_1(\omega), ..., X_n(\omega))$ is then a random vector and the probability density function corresponding to this sample is given by the formula

$$g_n(x_1, ..., x_n) = \frac{\mathbb{1}_{S_n(r)}(x_1, ..., x_n)}{2^n r\sqrt{\left(r^2 - x_1^2\right) \cdot ... \cdot \left[r^2 - \left(x_1^2 + ... + x_{n-1}^2\right)\right]}},$$

i.e. it is the same mapping as the one defining the Banach $\mathfrak{L}$-integral. To check that $g_n(x_1, ..., x_n)$ is in fact a probability density it is enough to show that

$$I_n := \int\limits_{S_n(r)} \frac{dx_1 ... dx_n}{2^n r\sqrt{\left(r^2 - x_1^2\right) \cdot ... \cdot \left[r^2 - \left(x_1^2 + ... + x_{n-1}^2\right)\right]}} = 1$$

for an arbitrary $n \geq 1$. Indeed,

$$\begin{aligned} I_{n+1} &= \int\limits_{S_n(r)} \frac{\left[\frac{1}{2}\int_{-\sqrt{\left(r^2 - x_1^2 - ... - x_n^2\right)}}^{\sqrt{\left(r^2 - x_1^2 - ... - x_n^2\right)}} dx_{n+1}\right] dx_1 ... dx_n}{2^n r\sqrt{\left(r^2 - x_1^2\right) \cdot ... \cdot \left[r^2 - \left(x_1^2 + ... + x_n^2\right)\right]}} = \\ &= I_n = ... = I_1 = \int_{-r}^{r} \frac{1}{2r} dx_1 = 1. \end{aligned}$$

Denote by $X_1, ..., X_n$ a random sequence having realization $x_1, ..., x_n$, i.e. having probability density $g_n(x_1, ..., x_n)$ as in the Banach functional integral. Observe that in this context the considered above quantities $M_n^2(r)$ are equal to the second moments of r.v's X_n. On the basis of the presented argumentation we obtain immediately the following result.

Proposition 3.1. *The Banach functional $F_n(\Phi)$ that approximates $\mathfrak{L}$-integral is equal to*

$$\int\limits_{S_n(r)} \Phi_n(x_1, ..., x_n) \frac{dx_1 ... dx_n}{2^n r \sqrt{\left(r^2 - x_1^2\right) \cdot ... \cdot \left[r^2 - \left(x_1^2 + ... + x_{n-1}^2\right)\right]}}$$
$$= \mathbb{E}_{RW}\left[\Phi_n(X_1, ..., X_n)\right],$$

where Φ_n is the restriction of Φ to $S_n(r)$, and the expectation $\mathbb{E}_{RW}$ is taken with respect to the measure $\mathbb{P}_{RW}$ induced by the random walk $X_1, ..., X_n$ on $S_n(r)$.

Proof. The proof is obvious, so the details are omitted. $\square$

In the next sections we show how to compute the Banach functional integral using random walk on the space $S_n(r)$, i.e. we demonstrate the power of the above Proposition 3.1.

4. Monte Carlo integration

Since $x_1, ..., x_n$ is a sample of $X_1, ..., X_n$, it is natural to expect that

$$\begin{aligned} F_n(\Phi_n) &= \int_{S_n(r)} \Phi_n(x_1, ..., x_n) \frac{dx_1 ... dx_n}{2^n r \sqrt{\left(r^2 - x_1^2\right) ... \left(r^2 - x_1^2 - ... - x_{n-1}^2\right)}} \\ &\approx \frac{1}{N} \sum_{i=1}^{N} \Phi_n(x_1(i), ..., x_n(i)), \end{aligned}$$

where $x_1(i), ..., x_n(i)$ is the i-th sample of $X_1, ..., X_n$. Such a method follows from the general theory of Monte Carlo integration, cf. e.g. [2], and on the other hand – it is a straightforward consequence of the law of large numbers. This is also justified by the following theorem.

Proposition 4.1. *Let $X^n = \mathrm{col}(X_1, ..., X_n)$, and let $X^n(1), ..., X^n(N)$, be a collection of $N-$ identical copies of X^n. Then*

$$\mathbb{P}_{RW}\left[\left|\frac{1}{N}\sum_{i=1}^{N} \Phi_n(X^n(i)) - F_n(\Phi_n)\right| > \varepsilon\right] \leq \frac{\mathrm{Var}\left[\Phi_n(X^n)\right]}{N\varepsilon^2},$$

and

$$\mathbb{P}_{RW}\left[\left|\frac{1}{N}\sum_{i=1}^{N} \Phi_n(X^n(i)) - F_n(\Phi_n)\right| > \varepsilon\right] \leq$$
$$\leq \frac{1}{N^3\varepsilon^4}\mathbb{E}_{RW}\left[\Phi_n(X^n) - F_n(\Phi_n)\right]^4 + \frac{6}{N^2\varepsilon^4}\left\{\mathrm{Var}\left[\Phi_n(X^n)\right]\right\}^2.$$

Proof. The first assertion follows by an application of Chebyshev's inequality, since

$$\mathbb{E}_{RW}\left[\frac{1}{N}\sum_{i=1}^{N}\Phi_n\left(X^n\left(i\right)\right)\right]=\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}_{RW}\left[\Phi_n\left(X^n\left(i\right)\right)\right]=\mathbb{E}_{RW}\left[\Phi_n\left(X^n\right)\right]$$
$$=F_n\left(\Phi_n\right),$$

and in addition

$$\mathrm{Var}\left[\frac{1}{N}\sum_{i=1}^{N}\Phi_n\left(X^n\left(i\right)\right)\right]=\mathbb{E}_{RW}\left[\frac{1}{N}\sum_{i=1}^{N}\left[\Phi_n\left(X^n\left(i\right)\right)-\mathbb{E}_{RW}\Phi_n\left(X^n\right)\right]\right]^2$$
$$=\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbb{E}_{RW}\left[\Phi_n\left(X^n\left(i\right)\right)-\mathbb{E}_{RW}\Phi_n\left(X^n\right)\right]$$
$$\cdot\left[\Phi_n\left(X^n\left(j\right)\right)-\mathbb{E}_{RW}\Phi_n\left(X^n\right)\right]=$$
$$=\frac{1}{N^2}\sum_{i=1}^{N}\mathbb{E}_{RW}\left[\Phi_n\left(X^n\left(i\right)\right)-\mathbb{E}_{RW}\Phi_n\left(X^n\right)\right]^2=$$
$$=\frac{1}{N^2}\cdot N\cdot\mathbb{E}_{RW}\left[\Phi_n\left(X^n\right)-F_n\left(\Phi_n\right)\right]^2=\frac{1}{N}\mathrm{Var}\left[\Phi_n\left(X^n\right)\right].$$

To obtain the second statement we apply Markov's inequality with parameter $p=4$, which gives

$$\mathbb{P}_{RW}\left[\left|\frac{1}{N}\sum_{i=1}^{N}\Phi_n\left(X^n\left(i\right)\right)-F_n\left(\Phi_n\right)\right|>\varepsilon\right]\leq$$
$$\leq\quad\frac{1}{\varepsilon^4}\mathbb{E}_{RW}\left[\frac{1}{N}\sum_{i=1}^{N}\left[\Phi_n\left(X^n\left(i\right)\right)-\mathbb{E}_{RW}\Phi_n\left(X^n\right)\right]\right]^4=$$
$$=\quad\frac{1}{N^4\varepsilon^4}\sum_{i,j,k,l=1}^{N}\mathbb{E}_{RW}\left[\Phi_n\left(X^n\left(i\right)\right)-\mathbb{E}_{RW}\Phi_n\left(X^n\right)\right]$$
$$\cdot\left[\Phi_n\left(X^n\left(j\right)\right)-\mathbb{E}_{RW}\Phi_n\left(X^n\right)\right]\cdot$$
$$\cdot\mathbb{E}_{RW}\left[\Phi_n\left(X^n\left(k\right)\right)-\mathbb{E}_{RW}\Phi_n\left(X^n\right)\right]$$
$$\cdot\left[\Phi_n\left(X^n\left(l\right)\right)-\mathbb{E}_{RW}\Phi_n\left(X^n\right)\right].$$

This equals

$$\frac{1}{N^4\varepsilon^4}\sum_{i=1}^{N}\mathbb{E}_{RW}\left[\Phi_n\left(X^n\left(i\right)\right)-\mathbb{E}_{RW}\Phi_n\left(X^n\right)\right]^4+$$
$$+\frac{6}{N^4\varepsilon^4}\sum_{\substack{i,j=1\\ i\neq j}}^{N}\mathbb{E}_{RW}\left[\Phi_n\left(X^n\left(i\right)\right)-\mathbb{E}_{RW}\Phi_n\left(X^n\right)\right]^2$$
$$\cdot\left[\Phi_n\left(X^n\left(j\right)\right)-\mathbb{E}_{RW}\Phi_n\left(X^n\right)\right]^2.$$

Moreover,

$$\sum_{i=1}^{N}\mathbb{E}_{RW}\left[\Phi_n\left(X^n\left(i\right)\right)-\mathbb{E}_{RW}\Phi_n\left(X^n\right)\right]^4=N\cdot\mathbb{E}_{RW}\left[\Phi_n\left(X^n\right)-F_n\left(\Phi_n\right)\right]^4,$$

and for $i\neq j$, the r.v's $\left[\Phi_n\left(X^n\left(i\right)\right)-\mathbb{E}_{RW}\Phi_n\left(X^n\right)\right]^2$ and $\left[\Phi_n\left(X^n\left(j\right)\right)-\mathbb{E}_{RW}\Phi_n\left(X^n\right)\right]^2$ are independent, thus

$$\begin{aligned}
&\sum_{\substack{i,j=1\\ i\neq j}}^{N}\mathbb{E}_{RW}\left[\Phi_n\left(X^n\left(i\right)\right)-\mathbb{E}_{RW}\Phi_n\left(X^n\right)\right]^2\left[\Phi_n\left(X^n\left(j\right)\right)-\mathbb{E}_{RW}\Phi_n\left(X^n\right)\right]^2\\
=\;&\sum_{\substack{i,j=1\\ i\neq j}}^{N}\mathbb{E}_{RW}\left[\Phi_n\left(X^n\left(i\right)\right)-\mathbb{E}_{RW}\Phi_n\left(X^n\right)\right]^2\\
&\qquad\mathbb{E}_{RW}\left[\Phi_n\left(X^n\left(j\right)\right)-\mathbb{E}_{RW}\Phi_n\left(X^n\right)\right]^2=\\
=\;&N\left(N-1\right)\left\{\mathbb{E}_{RW}\left[\Phi_n\left(X^n\right)-\mathbb{E}_{RW}\Phi_n\left(X^n\right)\right]^2\right\}^2\\
=\;&\left(N^2-N\right)\left\{\mathrm{Var}\left[\Phi_n\left(X^n\right)\right]\right\}^2.
\end{aligned}$$

The last conclusion of Proposition 4.1 is an immediate consequence of the above arguments. □

Corollary 4.2. *If Φ_n is bounded, say $|\Phi_n|\leq c<\infty$, then we have*

$$\frac{1}{N}\sum_{i=1}^{N}\Phi_n\left(X^n\left(i\right)\right)\to F_n\left(\Phi_n\right)\ \text{ a.s. and in } L^2 \text{ as } N\to\infty;$$

in particular, for each $\varepsilon>0$,

$$\mathbb{P}_{RW}\left[\left|\frac{1}{N}\sum_{i=1}^{N}\Phi_n\left(X^n\left(i\right)\right)-F_n\left(\Phi_n\right)\right|>\varepsilon\right]\leq\frac{16c^4}{N^3\varepsilon^4}+\frac{6\cdot 16c^4}{N^2\varepsilon^4}\to 0$$

as $N\to\infty$.

Proof. Clearly, $F_n(\Phi_n) = \mathbb{E}_{RW}\Phi_n(X^n)$, thus $|\Phi_n(X^n) - \mathbb{E}_{RW}\Phi_n(X^n)| \leq 2c$, and consequently $\operatorname{Var}[\Phi_n(X^n)] \leq 4c^2$ and $\mathbb{E}_{RW}[\Phi_n(X^n) - F_n(\Phi_n)]^4 \leq 16c^4$. Hence and from Proposition 4.1 the assertion of Corollary 4.2 easily follows; namely, a.s. convergence is a consequence of the estimate: for every fixed $\varepsilon > 0$,

$$\mathbb{P}_{RW}\left[\bigcup_{N\geq M}\left\{\left|\frac{1}{N}\sum_{i=1}^{N}\Phi_n(X^n(i)) - F_n(\Phi_n)\right| > \varepsilon\right\}\right] \leq \sum_{N\geq M}\left(\frac{16c^4}{N^3\varepsilon^4} + \frac{6\cdot 16c^4}{N^2\varepsilon^4}\right) \to 0$$

as $M \to \infty$, cf. [3], p. 151, and the inequality

$$\mathbb{E}_{RW}\left[\frac{1}{N}\sum_{i=1}^{N}\Phi_n(X^n(i)) - F_n(\Phi_n)\right]^2 = \operatorname{Var}\left[\frac{1}{N}\sum_{i=1}^{N}\Phi_n(X^n(i))\right] = \frac{1}{N}\operatorname{Var}[\Phi_n(X^n)] \leq \frac{4c^2}{N}$$

derived in the proof of Proposition 4.1 implies evidently convergence in L^2. □

Although the last Corollary justifies the approach based on the random walk on $S_n(r)$-space in Monte Carlo simulation of the proper value for $F_n(\Phi_n)$, it remains an open question how this random walk can be simulated by computer equipped with a generator of (pseudo)random numbers equally distributed on $[0,1]$.

5. Distributional mapping

In this section we are going to demonstrate how the random walk X^n on the $S_n(r)$-space can be simulated by computer.

Suppose $W^n = \operatorname{col}(W_1, ..., W_n)$ is a sequence of i.i.d. random variables with uniform density on $[0,1]$. We are looking for a smooth mapping $f : [0,1]^n \to \mathbb{R}^n$, such that

$$f(W^n) \overset{D}{=} X^n,$$

where $\overset{D}{=}$ means distributional equality. In oter words, we seek for f satisfying condition

$$F \circ f = G,$$

where F, G are cumulative distribution functions of X^n and W^n resp.

Theorem 5.1. *The smooth mapping $f : [0,1]^n \to \mathbb{R}^n$, satisfying condition $f(W^n) \stackrel{D}{=} X^n$, is given by the formula*

$$f_1(w_1, ..., w_n) = r(2w_1 - 1),$$
$$f_2(w_1, ..., w_n) = (2w_2 - 1)\sqrt{r^2 - r^2(2w_1 - 1)^2},$$
$$\vdots$$
$$f_n(w_1, ..., w_n) = (2w_n - 1)\sqrt{r^2 - \left[r^2(2w_1 - 1)^2 + ... + r^2(2w_{n-1} - 1)^2\right]}.$$

Proof. We have to check that the random vector $f(W^n)$ has the desired distribution. To this end, note first that the map f restricted to $[0,1]^n$ is 1 to 1, and the inverse mapping $f^{-1} = \left(f_1^{-1}, ..., f_n^{-1}\right)$ to f can be easily derived step by step, namely

$$w_1 = f_1^{-1}(x_1, ..., x_n) = \frac{(x_1/r + 1)}{2},$$
$$w_2 = f_2^{-1}(x_1, ..., x_n) = \frac{x_2}{2\sqrt{r^2 - x_1^2}} + \frac{1}{2},$$
$$\vdots$$
$$w_n = f_n^{-1}(x_1, ..., x_n) = \frac{x_n}{2\sqrt{r^2 - \left(x_1^2 + ... + x_{n-1}^2\right)}} + \frac{1}{2}.$$

Therefore the Jacobian of f^{-1} is equal to

$$Jf^{-1}(x) = \frac{1}{2r} \cdot \frac{1}{2\sqrt{r^2 - x_1^2}} \cdot ... \cdot \frac{1}{2\sqrt{r^2 - \left(x_1^2 + ... + x_{n-1}^2\right)}}.$$

Since W^n is distributed uniformly on $[0,1]^n$, by the Jacobi change of variables formula we obtain

$$\int_{[0,1]^n} \Phi_n(f_1(w_1, ..., w_n), ..., f_n(w_1, ..., w_n))\, dw_1 ... dw_n =$$
$$= \int_{f^{-1}([0,1]^n)} \Phi_n(x_1, ..., x_n) \left|Jf^{-1}(x)\right| dx_1 ... dx_n =$$
$$= \int_{S_n(r)} \Phi_n(x_1, ..., x_n) \frac{dx_1 ... dx_n}{2^n r\sqrt{\left(r^2 - x_1^2\right) \cdot ... \cdot \left[r^2 - \left(x_1^2 + ... + x_{n-1}^2\right)\right]}}.$$

In this way Theorem 5.1 is proved. $\square$

Corollary 5.2. *Let* $W^n = \mathrm{col}(W_1, ..., W_n)$ *be a random vector with the uniform distribution on* $[0,1]^n$, *and let* $X^n = \mathrm{col}(X_1, ..., X_n)$ *be the random walk on* $S_n(r)$, *i.e. a r. v. which has distribution concentrated on* $S_n(r)$ *according to the same density* $g_n(x_1, ..., x_n)$ *as in the Banach* $\mathfrak{L}$*-integral. Then the mapping* f *defined in the last Theorem transforms computer's drawing of (pseudo)random numbers* $(w_1, ..., w_n) := w^n$ *taken from* $[0,1]$ *into the simulated values* $(x_1, ..., x_n) = X^n(\omega)$ *of the random walk on* $S_n(r)$. *Consequently, given any bounded Borel-measurable mapping* $\Phi : S(r) \to \mathbb{R}$, *the approximated value for the Banach integral* $F_n(\Phi_n)$ *is equal to*

$$\frac{1}{N}\sum_{i=1}^{N}\Phi_n\left(X^n(i)\right) = \frac{1}{N}\sum_{i=1}^{N}\Phi_n\left(f_1\left(w^n(i)\right), ..., f_n\left(w^n(i)\right)\right),$$

where Φ_n *denotes the restriction of* Φ *to* $S_n(r)$, *and* $w^n(i)$ *is the i-th computer sample of (pseudo)random numbers* $(w_1, ..., w_n)$. *Furthermore, the error of approximation of* $F_n(\Phi_n)$ *by the above value exceeds* $\varepsilon > 0$ *with probability less than* $\dfrac{16c^4}{N^3\varepsilon^4} + \dfrac{6\cdot 16c^4}{N^2\varepsilon^4}$.

Combining the above Corollary with the previously given Theorem 2.2 and Corollary 4.2 we obtain finally the following result.

Corollary 5.3. *Let* $\Phi : S(r) \to \mathbb{R}$ *be a bounded, i.e.* $|\Phi| \le c < \infty$, *Borel-measurable mapping satisfying condition (2.1), let* $\{\varepsilon_n, n \ge 1\}$ *be a sequence of positive numbers such that* $0 < \dfrac{2c}{N}\sqrt[4]{1+6N^2} < \varepsilon_n$, *and let* $F(\Phi)$ *denote the Banach* $\mathfrak{L}$*-integral. Then with probability greater than* $1 - \dfrac{16c^4}{N^3\varepsilon_n^4} + \dfrac{6\cdot 16c^4}{N^2\varepsilon_n^4}$ *we have for sufficiently large* $n \ge n_C$,

$$\left|F(\Phi) - \frac{1}{N}\sum_{i=1}^{N}\Phi_n\left(f_1\left(w^n(i)\right), ..., f_n\left(w^n(i)\right)\right)\right| \le C\cdot\frac{r^2\cdot 2^n}{3^n} + \varepsilon_n\,, \quad (5.1)$$

where $f = (f_1, ..., f_n) : [0,1]^n \to \mathbb{R}^n$ *is the smooth mapping specified in Theorem 5.1,* Φ_n *is the restriction of* Φ *to* $S_n(r)$, *and* $w^n(i)$ *is the i-th computer sample of (pseudo)random numbers* $(w_1, ..., w_n)$ *drawn according to the uniform distribution from the interval* $[0,1]$.

Proof. As we know from Theorem 2.2, $|F(\Phi) - F_n(\Phi)| \le C\cdot r^2\cdot(2/3)^n$ for all sufficiently large $n \ge n_C$. Therefore, on account of Corollary 4.2, we conclude that

$$\begin{aligned}
&\mathbb{P}_{RW}\left[\left|F(\Phi) - \frac{1}{N}\sum_{i=1}^{N}\Phi_n\left(f_1\left(w^n(i)\right),...,f_n\left(w^n(i)\right)\right)\right| > C\cdot\frac{r^2\cdot 2^n}{3^n} + \varepsilon_n\right] \leq \\
&\quad \leq \mathbb{P}_{RW}\left[\left|F(\Phi) - F_n(\Phi)\right| > C\cdot\frac{r^2\cdot 2^n}{3^n}\right] \\
&\quad + \mathbb{P}_{RW}\left[\left|F_n(\Phi) - \frac{1}{N}\sum_{i=1}^{N}\Phi_n\left(f_1\left(w^n(i)\right),...,f_n\left(w^n(i)\right)\right)\right| > \varepsilon_n\right] \\
&\quad \leq 0 + \frac{16c^4}{N^3\varepsilon_n^4} + \frac{6\cdot 16c^4}{N^2\varepsilon_n^4},
\end{aligned}$$

whenever $n \geq n_C$ is sufficiently large. □

Evidently, the sequence of numbers $\varepsilon_n = \varepsilon_n(N)$ should be chosen here in such a way that

$$1 > \frac{16c^4}{N^3\varepsilon_n^4} + \frac{6\cdot 16c^4}{N^2\varepsilon_n^4} \to 0 \quad \text{as} \quad N, n \to \infty,$$

thus we propose the following procedure. Choose first ε_n, for instance $\varepsilon_n = C\cdot r^2\cdot 2^n/3^n$, so that the right-hand side of (5.1) is less than $2C\cdot r^2\cdot 2^n/3^n$, next fix $n \geq n_C$ to obtain the desired accuracy of approximation in (5.1), and then select N so large that the probability

$$1 - \frac{16c^4}{N^3\varepsilon_n^4} + \frac{6\cdot 16c^4}{N^2\varepsilon_n^4}$$

is sufficiently close to 1.

References

1. S. Banach, *The Lebesgue integral in abstract spaces*, in S. Saks, Theory of the Integral, Monografie Matematyczne, Warszawa - Lwów 1937.
2. G. S. Fishman, *Monte Carlo*, Springer, Berlin, Heidelberg, New York 1996.
3. M. Loève, *Probability Theory*, Van Nostrand, Princeton 1960.

Lord Kelvin's method of images approach to the Rotenberg model

Adam Gregosiewicz[1]

Abstract

We study a mathematical model of cell populations dynamics proposed by M. Rotenberg [14] and investigated by M. Boulanouar [7]. Here, a cell is characterized by her maturity and speed of maturation. The growth of cell populations is described by a partial differential equation with a boundary condition. We use semigroup theory approach and apply Lord Kelvin's method of images to give a new proof that the model is well posed.

1. Introduction

In the Rotenberg model of cell populations dynamics [14] a cell is characterized by two variables, its maturity and speed of maturation. We assume that the maturity is a real number x that belongs to the interval $I = (0, 1)$ and speed of maturation v belongs to a set $V = (a, b)$, where a and b are non-negative real numbers such that $a < b < +\infty$. The growth of the cell's population density is governed by the partial differential equation

$$\frac{\partial f}{\partial t} = -v\frac{\partial f}{\partial x}, \tag{1.1}$$

where $f = f(x, v, t)$, $t \geqslant 0$, is the cell's density at (x, v) at time t. In this model a cell is born at $x = 0$ and dies at $x = 1$, and the boundary condition

$$vf(0, v, t) = p\int_V wk(1, w)f(1, w, t)\,\mathrm{d}w$$

[1]Lublin University of Technology, Nadbystrzycka 38A, 20-618 Lublin, Poland, *e-mail: a.gregosiewicz@pollub.pl*

Keywords: Method of images, Strongly continuous semigroup, Cell population dynamics, Partial differential equations.

The author's research was partly supported by Polish National Science Centre grant 2014/15/N/ST1/03110.

describes the reproduction rule. Here k satisfies

$$\int_V k(v,w)\,\mathrm{d}w = 1 \tag{1.2}$$

for any $v \in V$, and $k(v,w)$ is the probability density of daughter velocity, conditional on v being the velocity of the mother. Furthermore, it is assumed that $p \geqslant 0$ is the average number of viable daughters per mitosis. However (see [7]), it may be also important to consider the case when a cell degenerate in the sense that its daughters inherit mother's velocity. Such behaviour is described by

$$f(0,v,t) = qf(1,v,t),$$

where $q \geqslant 0$ is the average number of viable daughters per mitosis. Therefore, we combine this two cases and assume that the reproduction rule is characterized by the boundary condition

$$vf(0,v,t) = p\int_V wk(1,w)f(1,w,t)\,\mathrm{d}w + qvf(1,v,t), \qquad v \in V. \tag{1.3}$$

It is also biologically interesting when $V \subset (a,b)$ is a discrete set, that is cells mature only at certain (at most countably many) velocities. In this case (1.2) becomes

$$\sum_{w\in V} k(v,w) = 1$$

with the boundary condition

$$vf(0,v,t) = p\sum_{w\in V} wk(1,w)f(1,w,t) + qvf(1,v,t), \qquad v \in V.$$

Well-posedness of the (generalized) Rotenberg model may by equivalently rephrase in the semigroup theory. Roughly speaking, see [11, II.1.2], the model (1.1)-(1.3) is well-posed if and only if the operator

$$f \mapsto -v\frac{\partial f}{\partial x}$$

with domain related to (1.3) is the generator of a strongly continuous semigroup.

In this paper we give a new proof of the generation theorem of Boulanouar [7, Theorem 2.2, Theorem 3.1]. To this end we use Lord Kelvin's method of images. For detailed introduction to the method of images see [3], and [2, 4, 5, 6, 8, 10, 12, 13, 15] for some examples. As a by-product we obtain an explicit formula for the semigroup related to Rotenberg model.

2. Notations

In this section we recall some basic preliminaries and introduce notations. Given $\omega \in \mathbb{R}$ we define a function $\mathrm{e}_\omega \colon \mathbb{R} \to \mathbb{R}$ by the formula $\mathrm{e}_\omega(x) = \mathrm{e}^{\omega x}$, $x \in \mathbb{R}$. Let I be an open set of real numbers. Moreover, let $V \subset \mathbb{R}$ and ν be a measure on V. Then for $\Omega = I \times V$ we define $L^1(\Omega, \mu_\omega)$ to be the space of absolutely integrable real functions with respect to the measure

$$\mu_\omega(\mathrm{d}x, \mathrm{d}v) = \mathrm{e}_\omega\big(v^{-1}(x-1)\big)\,\mathrm{d}x\,\nu(\mathrm{d}v),$$

where $\mathrm{d}x$ is the Lebesgue measures on I. That is, $f \colon \Omega \to \mathbb{R}$ belongs to $L^1(\Omega, \mu_\omega)$ if and only if

$$\|f\|_{L^1(\Omega,\mu_\omega)} := \int_V \int_I \mathrm{e}_\omega\big(v^{-1}(x-1)\big)|f(x,v)|\,\mathrm{d}x\,\nu(\mathrm{d}v) < +\infty.$$

Then $L^1(\Omega, \mu_\omega)$ is a Banach space with the norm $\|\cdot\|_{L^1(\Omega,\mu_\omega)}$. Let also $W^1(\Omega)$ be the space of functions $f \in L^1(\Omega, \mu_\omega)$ such that given $v \in V$ we have $f(\cdot, v) \in W^{1,1}(I, \mathrm{d}x)$, where $W^{1,1}(I, \mathrm{d}x)$ is the Sobolev space of absolutely continuous functions on I such that $D_1 f$ is a member of $L^1(\Omega, \mu_\omega)$. Here $D_1 f$ is the partial derivative of f with respect to the first variable, that is $D_1 f(x,v) = \partial/\partial x f(x,v)$. Furthermore, if $\omega = 0$, then we denote the spaces $L^1(\Omega, \mu_0)$ by $L^1(\Omega)$ and $W^1(\Omega, \mu_0)$ by $W^1(\Omega)$.

3. Generation theorem

As in the introduction, let $I = (0,1)$ and given $a, b \in \mathbb{R}$ such that $0 \leqslant a < b$ let $(V, \mathcal{V}, \nu)$ be a measure space where $V \subset (a,b)$. As we said before, most interesting cases from a biological point of view would be when V equals (a,b) or is its finite subset, and ν is the Lebesgue measure or the counting measure on V, respectively. However, we do not need to assume that and can work in an abstract setup.

Let $k \colon V \times V \to [0, +\infty)$ be a measurable, non-negative real function such that

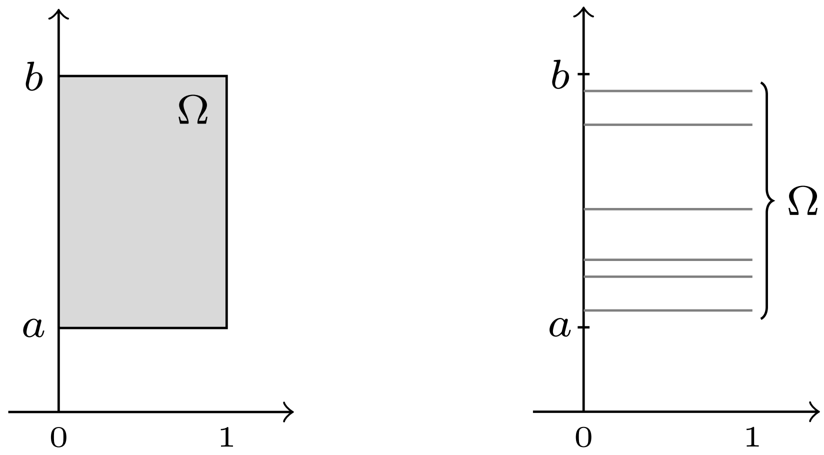
$$\int_V k(v,w)\,\nu(\mathrm{d}w) = 1, \qquad v \in V. \tag{3.1}$$

Then for $\Omega = I \times V$, see Figure 1, we consider the operator A in $L^1(\Omega)$ given by

$$Af(x,v) = -vD_1 f(x,v), \qquad (x,v) \in \Omega, \tag{3.2}$$

with the domain $D(A)$ composed of functions $f \in W^1(\Omega)$ satisfying a the boundary condition

$$vf(0,v) = p\int_V wk(w,v)f(1,w)\,\nu(\mathrm{d}w) + qvf(1,v), \qquad v \in V, \tag{3.3}$$



(a) Continuous case.

(b) Discrete case.

Figure 1: The set Ω.

where p, q are non-negative real numbers such that $p + q > 0$.

Theorem 3.1. *The operator A generates a strongly continuous semigroup in $L^1(\Omega)$.*

Formula (3.2) indicates that for fixed $v \in V$ a desired semigroup should resemble a translation semigroup. Hence, we would like to define $\{T(t),\, t \geqslant 0\}$ in $L^1(\Omega)$ by

$$T(t)f(x,v) = \tilde{f}(x - tv, v), \qquad (x,v) \in \Omega,\ t \geqslant 0,\ f \in L^1(\Omega), \tag{3.4}$$

where $\tilde{f}$ is a measurable function defined on $\tilde{\Omega} = (-\infty, 1) \times V$. Since $T(0)f$ should equals f, it follows that $\tilde{f}$ must be an extension of f. Moreover, because a semigroup leave its domain invariant, given $f \in L^1(\Omega)$ we are looking for $\tilde{f} \colon \tilde{\Omega} \to \mathbb{R}$ such that

(E1) the restriction of $\tilde{f}$ to Ω equals f, that is $\tilde{f}_{|\Omega} = f$,

(E2) if $f \in D(A)$, then $T(t)\tilde{f}$ given by (3.4) belongs to $D(A)$ for $t \geqslant 0$.

Let $f \in D(A)$. In particular, (E2) implies that $\tilde{f}$ must be chosen in such a way that $T(t)f$ given by (3.4) satisfies the boundary condition (3.3). Hence, we must have

$$v\tilde{f}(-tv, v) = p \int_V wk(w,v)\tilde{f}(1 - tw, w)\, \nu(\mathrm{d}w) + qv\tilde{f}(1 - tv, v),$$

where $v \in V$, $t \geqslant 0$. If we denote $x = -tv$, $x \leqslant 0$, $v \in V$, this may be rewritten as

$$\tilde{f}(x,v) = pv^{-1}\int_V wk(w,v)\tilde{f}(1+xwv^{-1},w)\,\nu(\mathrm{d}w) + q\tilde{f}(1+x,v). \qquad (3.5)$$

Let $i \geqslant 1$ be a non-negative integer and set

$$\Omega_i = \{(x,v) \in \mathbb{R}^2 \colon v \in V,\ -ivb^{-1} < x \leqslant -(i-1)vb^{-1}\},$$

see Figure 2.

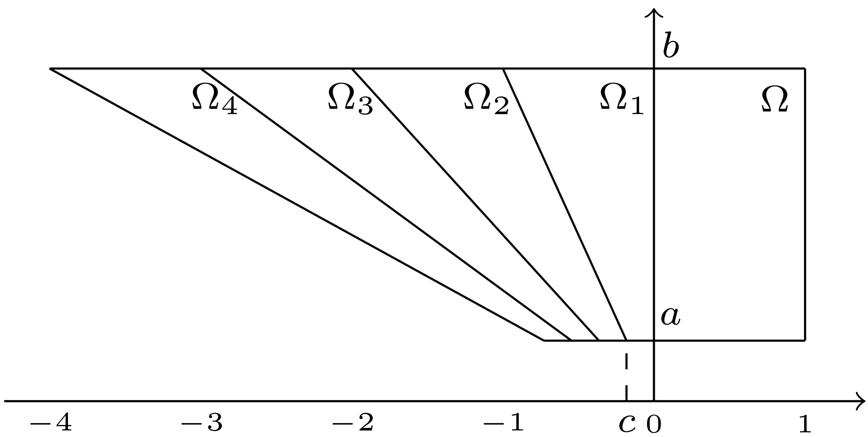


Figure 2: The set Ω_i, where $c = ab^{-1}$.

For $w \in V$ it follows that $(x,v) \in \Omega_j$, $j \geqslant 1$, implies $(1+xwv^{-1},w) \in \bigcup_{i=0}^{j-1}\Omega_i$, where by convention $\Omega_0 = \Omega$. Therefore, we may define $\tilde{f}$ by induction. Having defined it on $\bigcup_{i=0}^{j}\Omega_i$, $j \geqslant 0$, for $(x,v) \in \Omega_{j+1}$ we let $\tilde{f}(x,v)$ be given by (3.5). This shows that if $\tilde{f} \colon \tilde{\Omega} \to \mathbb{R}$ satisfying (E1) and (E2) exists, then it is uniquely determined.

Definition 3.2. Given $f \in L^1(\Omega)$ we denote by $\tilde{f} \colon \tilde{\Omega} \to \mathbb{R}$ its unique extension satisfying

$$\tilde{f}(x,v) = \begin{cases} f(x,v), & x > 0,\\ \frac{p}{v}\int_V wk(w,v)\tilde{f}(1+x\frac{w}{v},w)\,\nu(\mathrm{d}w) + q\tilde{f}(1+x,v), & x \leqslant 0, \end{cases}$$

almost everywhere and call it the *boundary extension* of f.

It is worth noting that we do not assume that f belongs to $D(A)$ in order to define $\tilde{f}$. However, what is crucial, boundary extensions of functions from the domain of A posses important property which we describe in Lemma 3.4.

Lemma 3.3. *Let*

$$\omega > \max(b\log(p+q), 0). \tag{3.6}$$

Then for $f \in L^1(\Omega)$ the boundary extension $\tilde{f}$ belongs to $L^1(\tilde{\Omega}, \mu_\omega)$ and

$$\|\tilde{f}\|_{L^1(\tilde{\Omega},\mu_\omega)} \leqslant M_\omega \|f\|_{L^1(\Omega)}, \tag{3.7}$$

where $M_\omega = (1 - C_\omega)^{-1}$ and $C_\omega = (p+q)\mathrm{e}^{-\omega b^{-1}}$.

Proof. Let $\omega > 0$, $f \in L^1(\Omega, \mu)$ and $\tilde{f}$ be its extension. For $i \geqslant 1$, $v \in V$, we denote

$$\Omega_{i,v} = \{x \in \mathbb{R}\colon\ -ivb^{-1} < x \leqslant -(i-1)vb^{-1}\}.$$

It follows that

$$\begin{aligned}\int_{\Omega_i} |\tilde{f}|\mu_\omega &= \int_V \int_{\Omega_{i,v}} \mathrm{e}^{\omega v^{-1}(x-1)} |\tilde{f}(x,v)| \,\mathrm{d}x\, \nu(\mathrm{d}v) \\ &= p \int_V \int_V \frac{w}{v} k(w,v) \int_{\Omega_{i,v}} \mathrm{e}^{\frac{\omega}{v}(x-1)} |\tilde{f}(1 + x\frac{w}{v}, w)| \,\mathrm{d}x\, \nu(\mathrm{d}w)\, \nu(\mathrm{d}v) \\ &\quad + q \int_V \int_{\Omega_{i,v}} \mathrm{e}^{\omega v^{-1}(x-1)} |\tilde{f}(1+x, v)| \,\mathrm{d}x\, \nu(\mathrm{d}v).\end{aligned}$$

Changing variables leads to

$$\begin{aligned}\int_{\Omega_i} |\tilde{f}|\mu_\omega = p \int_V \int_V k(w,v) \int_{1+\Omega_{i,w}} \mathrm{e}^{\omega w^{-1}(x-1)} \mathrm{e}^{-\omega v^{-1}} |\tilde{f}(x,w)| \,\mathrm{d}x\, \nu(\mathrm{d}w)\, \nu(\mathrm{d}v) \\ + q \int_V \int_{1+\Omega_{i,v}} \mathrm{e}^{\omega v^{-1}(x-1)} \mathrm{e}^{-\omega v^{-1}} |\tilde{f}(x,v)| \,\mathrm{d}x\, \nu(\mathrm{d}v),\end{aligned} \tag{3.8}$$

where $1 + \Omega_{i,w}$ is the algebraic sum $\{1\} + \Omega_{i,w}$. Since $\mathrm{e}^{-\omega v^{-1}} < \mathrm{e}^{-\omega b^{-1}}$ for $(x,v) \in \Omega_i$, $i \geqslant 1$, using (3.1) it follows that

$$\int_{\Omega_i} |\tilde{f}|\mu_\omega \leqslant C_\omega \int_{1+\Omega_i} |\tilde{f}|\mu_\omega, \tag{3.9}$$

where $1 + \Omega_i = \{(1+x, v) \in \mathbb{R}^2\colon (x,v) \in \Omega_i\}$. Furthermore, we have

$$\bigcup_{i=1}^{j} (1 + \Omega_i) \subset \bigcup_{i=0}^{j-1} \Omega_i, \qquad j \geqslant 1,$$

see Figure 3. Combining this with (3.9), we obtain

$$\int_{\bigcup_{i=0}^{j} \Omega_i} |\tilde{f}|\mu_\omega = D_\omega + \int_{\bigcup_{i=1}^{j} \Omega_i} |\tilde{f}|\mu_\omega \leqslant \|f\|_{L^1(\Omega,\mu_\omega)} + C_\omega \int_{\bigcup_{i=0}^{j-1} \Omega_i} |\tilde{f}|\mu_\omega.$$

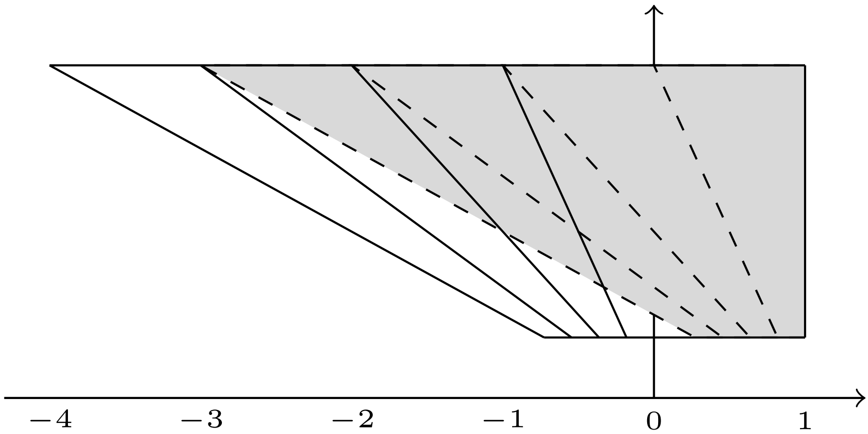


Figure 3: The set $\bigcup_{i=1}^{4}(1+\Omega_i)$ is colored blue.

Hence the inductive argument shows that

$$\int_{\bigcup_{i=0}^{j}\Omega_i} |\tilde{f}|\mu_\omega \leqslant \|f\|_{L^1(\Omega,\mu_\omega)} \sum_{i=0}^{j} C_\omega^i. \tag{3.10}$$

By (3.6) we have $C_\omega < 1$, hence the last sum converges as $j \to +\infty$, and

$$\int_{\tilde{\Omega}} |\tilde{f}|\mu_{\omega_0} < M_\omega \|f\|_{L^1(\Omega,\mu_\omega)}.$$

Finally, since $\mathrm{e}^{\omega v^{-1}(x-1)} < 1$ for $(x,v) \in \tilde{\Omega}$, it follows that

$$\|f\|_{L^1(\Omega,\mu_\omega)} \leqslant \|f\|_{L^1(\Omega)},$$

which proves (3.7). $\square$

In view of Lemma 3.3 let here and subsequently fix $\omega > \max(b\log(p+q),0)$. The set $\mathbb{Y}$ of all boundary extensions $\tilde{f}$ of $f \in L^1(\Omega)$ is a subspace of $L^1(\tilde{\Omega},\mu_\omega)$. Hence we may define the extension operator

$$E\colon L^1(\Omega) \to \mathbb{Y} \hookrightarrow L^1(\tilde{\Omega},\mu_\omega), \qquad Ef = \tilde{f}.$$

which by (3.7) is bounded with

$$\|E\|_{L^1(\Omega)\to L^1(\tilde{\Omega},\mu_\omega)} \leqslant M_\omega.$$

Moreover, E is one-to-one since $Ef = 0$ implies $0 = \tilde{f}_{|\Omega} = f$. Therefore E is an isomorphism of $L^1(\Omega)$ and $\mathbb{Y}$, and of course E^{-1} is the restriction operator, that is $E^{-1} = R$, where $Rf = f_{|\Omega}$, and hence

$$\|E^{-1}\|_{L^1(\tilde{\Omega},\mu_\omega)\to L^1(\Omega)} = 1.$$

Let now $\{\tilde{T}(t),\, t \geqslant 0\}$ be the family of operators in $L^1(\tilde{\Omega}, \mu_\omega)$ given by

$$\tilde{T}(t)f(x,v) = f(x - tv, v), \qquad (x,v) \in \tilde{\Omega},\ t \geqslant 0,\ f \in L^1(\tilde{\Omega}, \mu_\omega).$$

It is easy to show that $\{\tilde{T}(t),\, t \geqslant 0\}$ is a strongly continuous semigroup and its generator $\tilde{A}$ is given by

$$\tilde{A}f(x,v) = -vD_1 f(x,v), \qquad (x,v) \in \tilde{\Omega},\ f \in L^1(\tilde{\Omega}, \mu_\omega),$$

with domain $D(\tilde{A}) = W^1(\tilde{\Omega}, \mu_\omega)$.

Lemma 3.4. *Given $f \in L^1(\Omega)$ we have $f \in D(A)$ if and only if $\tilde{f} \in D(\tilde{A})$.*

Proof. Assume that $\tilde{f} \in D(\tilde{A})$. Then of course $f = \tilde{f}_{|\Omega} \in W^1(\Omega)$. Given $v \in V$, by the continuity of $\tilde{f}(\cdot, v)$ and (3.5) we have

$$\begin{aligned} f(0,v) = \tilde{f}(0,v) &= \lim_{x\to 0^-} \tilde{f}(x,v) = pv^{-1}\int_V wk(w,v)\tilde{f}(1,w)\,\nu(\mathrm{d}w) + q\tilde{f}(1,v) \\ &= pv^{-1}\int_V wk(w,v)f(1,w)\,\nu(\mathrm{d}w) + qf(1,v) \end{aligned}$$

by the Lebesgue dominated convergence theorem. This shows that f satisfies (3.3), and hence $f \in D(A)$.

On the other hand, let $f \in D(A)$ and fix $v \in V$. For $j \geqslant 0$ denote $\Gamma_{j,v} = \bigcup_{i=0}^{j} \Omega_{i,v}$. We have $\tilde{f} \in W^1(\Gamma_{0,v})$. Assume now that $\tilde{f} \in W^1(\Gamma_{j,v})$ for some $j \geqslant 0$. Then for $x \in \Omega_{j+1,v}$ by the Fubini theorem and (3.3) we have

$$\begin{aligned} \int_0^x D_1\tilde{f}(y,v)\,\mathrm{d}y = pv^{-1}&\int_V wk(w,v)wv^{-1}\int_0^x D_1\tilde{f}(1+ywv^{-1}, w)\,\mathrm{d}y\,\nu(\mathrm{d}w) \\ &+ q\int_0^x D_1\tilde{f}(1+y,v) \\ &= pv^{-1}\int_V wk(w,v)\tilde{f}(1+x\frac{w}{v}, w)\,\mathrm{d}y\,\nu(\mathrm{d}w) + q\tilde{f}(1+x,v) \\ &\quad - pv^{-1}\int_V wk(w,v)\tilde{f}(1,w)\,\mathrm{d}y\,\nu(\mathrm{d}w) - q\tilde{f}(1,v) \\ &= \tilde{f}(x,v) - \tilde{f}(0,v). \end{aligned}$$

This proves that $\tilde{f} \in W^1(\Gamma_{j+1,v})$. Using the induction argument it follows that $\tilde{f} \in W^1(\tilde{\Omega}) = D(\tilde{A})$. $\square$

Lemma 3.5. *The space $\mathbb{Y}$ is invariant for the semigroup $\{\tilde{T}(t),\, t \geqslant 0\}$.*

Proof. Let $f \in L^1(\Omega)$ and $\tilde{f} \in \mathbb{Y}$ be its boundary extension. By (3.5) we have

$$\tilde{T}(t)\tilde{f}(x,v) = pv^{-1}\int_V wk(w,v)\tilde{f}(1+xwv^{-1}-tw,w)\,\nu(\mathrm{d}w)+q\tilde{f}(1+x-tv,v),$$

for $x < tv$ and $(x,v) \in \tilde{\Omega}$. Hence $\tilde{T}(t)\tilde{f}$ is the extension of $g \in L^1(\Omega)$, where $g(x,v) = \tilde{f}(x-tv,v)$, $(x,v) \in \Omega$. □

Proof of Theorem 3.1. By Lemma 3.5 the part $\tilde{A}_{\mathbb{Y}}$ of $\tilde{A}$ in $\mathbb{Y}$ generates the strongly continuous semigroup $\{\tilde{T}_{\mathbb{Y}}(t),\, t \geqslant 0\}$ in $\mathbb{Y}$ given by

$$\tilde{T}_{\mathbb{Y}}(t)\tilde{f}(x,v) = \tilde{f}(x-tv,v), \qquad t \geqslant 0,\ (x,v) \in \Omega,\ f \in \mathbb{Y}; \tag{3.11}$$

see e.g. [9, Corollary II.2.3]. This proves that $\{T(t),\, t \geqslant 0\}$, where

$$T(t) = R\tilde{T}_{\mathbb{Y}}(t)E, \tag{3.12}$$

is a strongly continuous semigroup in $L^1(\Omega)$ similar to $\{\tilde{T}_{\mathbb{Y}}(t),\, t \geqslant 0\}$, see e.g. [1, 7.4.22]. Moreover, its generator is the operator $R\tilde{A}_{\mathbb{Y}}E$ and with domain $RD(\tilde{A})$. However,

$$R\tilde{A}_{\mathbb{Y}}Ef(x,v) = -vD_1f(x,v) = Af(x,v), \qquad (x,v) \in \Omega,$$

and by Lemma 3.4 it follows that $RD(\tilde{A}) = D(A)$, which completes the proof. □

As in the proof of Theorem 3.1 denote by $\{T(t),\, t \geqslant 0\}$ the semigroup generated by A. Then by (3.12) we have

$$T(t)f(x,v) = \tilde{f}(x-tv), \qquad t \geqslant 0,\ (x,v) \in \Omega,\ f \in L^1(\Omega),$$

as conjectured in (3.4).

Lemma 3.6. *We have*

$$\|T(t)\|_{L^1(\Omega)\to L^1(\Omega)} = \max(1, p+q)$$

for $0 < t < b^{-1}$.

Proof. Fix $0 < t < b^{-1}$. Then

$$\|T(t)f\|_{L^1(\Omega)} = \int_V\int_{tv}^1 f(x-tv,v)\,\mathrm{d}x\,\nu(\mathrm{d}v) + \int_V\int_0^{tv}\tilde{f}(x-tv,v)\,\mathrm{d}x\,\nu(\mathrm{d}v).$$

The first term on the right-hand side equals $\int_V \int_0^{1-tv} f(x,v)\,\mathrm{d}x\,\nu(\mathrm{d}v)$. Moreover, since for $x \in (0, tv)$ it follows that $(x - tv, v) \in \Omega_1$, we obtain

$$\begin{aligned}\int_0^{tv} \tilde{f}(x-tv,v)\,\mathrm{d}x &= pv^{-1}\int_V wk(w,v)\int_0^{tv} f(1+xwv^{-1}-tw,w)\,\mathrm{d}x\,\nu(\mathrm{d}w)\\ &\quad + q\int_0^{tv} f(1+x-tv,v)\,\mathrm{d}x\\ &= p\int_V k(w,v)\int_{1-tw}^1 f(x,w)\,\mathrm{d}x\,\nu(\mathrm{d}w)\\ &\quad + q\int_{1-tv}^1 f(x,v)\,\mathrm{d}x.\end{aligned}$$

Hence, by the Fubini theorem and (3.1) we get

$$\int_V\int_0^{tv}\tilde{f}(x-tv,v)\,\mathrm{d}x\,\nu(\mathrm{d}v) = (p+q)\int_V\int_{1-tv}^1 f(x,v)\,\mathrm{d}x\,\nu(\mathrm{d}v),$$

which proves that

$$\begin{aligned}\|T(t)f\|_{L^1(\Omega)} = &\int_V\int_0^{1-tv} f(x,v)\,\mathrm{d}x\,\nu(\mathrm{d}v)\\ &+ (p+q)\int_V\int_{1-tv}^1 f(x,v)\,\mathrm{d}x\,\nu(\mathrm{d}v).\end{aligned} \tag{3.13}$$

Finally, if $1 \geqslant p+q$, then we set

$$g = \mathbb{1}_{\{(x,v)\colon\, 0<x<1-tv,\ v\in V\}},$$

and if conversely $1 < p+q$, then we set

$$g = \mathbb{1}_{\{(x,v)\colon\, 1-tv<x<1,\ v\in V\}}.$$

By (3.13) it follows that

$$\|T(t)g\|_{L^1(\Omega)} = \max(1,p+q)\|g\|_{L^1(\Omega)},$$

which completes the proof. □

By Lemma 3.6 it is easy to prove the following result.

Corollary 3.7. *Given $t \geqslant 0$ we have*

$$\|T(t)\|_{L^1(\Omega)\to L^1(\Omega)} \leqslant \max(1,p+q)^{[tb]+1},$$

where $[tb]$ is the greatest integer smaller then or equal to tb.

References

1. A. Bobrowski, *Functional analysis for probability and stochastic processes*, Cambridge University Press, Cambridge, 2005. MR 2176612 (2006g:60001)
2. A. Bobrowski, *Generation of cosine families via Lord Kelvin's method of images*, J. Evol. Equ. **10** (2010), no. 3, 663–675. MR 2674063 (2012b:47114)
3. A. Bobrowski, *Lord Kelvin's method of images in semigroup theory*, Semigroup Forum **81** (2010), no. 3, 435–445. MR 2735660 (2011m:47075)
4. A. Bobrowski and A. Gregosiewicz, *A general theorem on generation of moments-preserving cosine families by Laplace operators in* $C[0,1]$, Semigroup Forum **88** (2014), no. 3, 689–701. MR 3212614
5. A. Bobrowski, A. Gregosiewicz, and M. Murat, *Functionals-preserving cosine families generated by laplace operators in* $C[0,1]$, Discrete and Continuous Dynamical Systems - Series B **20** (2015), no. 7, 1877–1895.
6. K. Bogdan and T. Żak, *On Kelvin transformation*, J. Theoret. Probab. **19** (2006), no. 1, 89–120.
7. M. Boulanouar, *A mathematical study for a Rotenberg model*, J. Math. Anal. Appl. **265** (2002), no. 2, 371–394. MR 1876146 (2002k:92020)
8. H. Cheng, *On the method of images for systems of closely spaced conducting spheres*, SIAM J. Appl. Math. **61** (2000/01), no. 4, 1324–1337.
9. K.-J. Engel and R. Nagel, *One-parameter semigroups for linear evolution equations*, Graduate Texts in Mathematics, vol. 194, Springer-Verlag, New York, 2000. MR 1721989 (2000i:47075)
10. V. I. Fabrikant, *Generalized method of images in the crack analysis*, Internat. J. Engrg. Sci. **35** (1997), no. 12-13, 1159–1184.
11. J. Goldstein, *Semigroups of linear operators and applications*, Oxford University Press, New York, 1985.
12. E. W. Hobson, *Synthetical Solutions in the Conduction of Heat*, Proc. London Math. Soc. **S1-19** (1887), no. 1, 279.
13. I. V. Lindell, K. H. Wallén, and A. H. Sihvola, *Electrostatic image theory for two intersecting conducting spheres*, J. Electromagn. Waves Appl. **17** (2003), no. 11, 1643–1660.
14. M. Rotenberg, *Transport theory for growing cell populations*, J. Theoret. Biol. **103** (1983), no. 2, 181–199. MR 713945 (85a:92031)
15. V. Volterra, *Note on the Application of the Method of Images to Problems of Vibrations*, Proc. London Math. Soc. **S2-2** (1905), no. 1, 327.

Differences between optimal routes for linear quadratic problems with fixed and optimally stopped horizon

Edward Kozłowski[1]

Abstract

The routing problem of linear system to hit the target is investigated in this paper. The classical linear quadratic control problem was replaced by the problem of determining the optimal trajectory (way, track, path). The general aim consists of minimization of composite cost function, which depends on route (as a set of landmarks) and horizon. To illustrate the influence of horizon a numerical examples are included.

1. Introduction

The different engineering applications of stochastic systems (e.g. control, navigation, stabilization, cost minimization, identification etc.) are widely presented in literature (see e.g. [2], [3], [4], [5], [13], [15], [19], [20]). In many cases these systems must be controlled to perfect perform the aim. Unfortunately to exactly achieve the goal first we must determine the control laws for systems. Sometimes, in order to achieve the goal the system should be moved after a certain path (trajectory). In this case the problem depends on determining the optimal trajectory.

The task presented in this paper consists of determining the optimal path on which the system achieves the lowest total costs, wich is a sum of costs of changes, energetic costs of controls and losses associated with not hitting the target. This task for fixed horizon was considered in [6]. Additionally it was proved, that the horizon has a large influence to total cost (see e.g. [5], [14], [15], [16], [17]). The construction of control laws for fixed horizon is usually presented in literature (see e.g. [1], [9], [10]). The problem arises

[1]Lublin University of Technology, Department of Quantitative Methods, Nadbystrzycka 38, 20-618 Lublin, Poland, *e-mail: e.kozlovski@pollub.pl*

2010 Mathematics Subject Classification. Primary 93E20, 60G40; Secondary 93E95, 90C20.

Keywords: optimal trajectory, optimal control, linear quadratic routing problem, Snell envelope.

when the horizon is unknown. Which of horizons is optimal? The control with optimal stopping is presented in the works [11], [12], [21].

The goal of this paper is determining the optimal trajectory (path, route as a set of landmarks) and horizon. The controlled stochastic system (object, robot) should move after this optimal route and must be stopped before reaching the biggest possible horizon to perfect hit the target. To determine the optimal trajectory and stopping moment (optimal horizon) the idea of dynamic programming was employed.

The paper is organized as follows. In section II the linear quadratic control problem is converted to linear quadratic routing problem. The solution of the routing problem with the fixed horizon is given in section III. Section IV provides the solution of linear quadratic routing problem with the optimally stopped horizon. The numerical simulations shown that to achieve the aim sometimes the system must be stopped earlier than the biggest possible horizon.

2. Exchange the LQC problem to LQR problem

In this part the linear quadratic control (LQC) problem with fixed horizon will be replaced by a linear quadratic routing (LQR) problem. As we shall see later, this problems are dual.

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space. Suppose that $w_1, w_2, ...$ are independent n-dimensional random vectors on this space, with normal $N(0, I_n)$ distribution. We assume that all the above-mentioned objects are stochastically independent. Let the stochastic linear system be described by a state equation

$$y_{i+1} = Ay_i + Bu_i + \sigma w_{i+1} \tag{2.1}$$

where $i = 0, ..., N-1$, $y_i \in \mathcal{R}^n$, $B \in \mathcal{R}^{n \times k}$, $\sigma \in \mathcal{R}^{n \times n}$ and an initial state is $\|y_0\| < \infty$. On $(\Omega, \mathcal{F}, P)$ we define a family of sub-σ-fields $\mathcal{Y}_j = \sigma\{y_i : i = 0, 1, ..., j\}$. The matrices $\|A\| < \infty$, $\|B\| < \infty$ and $\|\sigma\| < \infty$, where $\|\cdot\|$ denotes a matrix norm as $\|A\| = \max\limits_{\|x\| \leqslant 1} \|Ax\|$ (the system (2.1) is Boundary Input Boundary Output stable). The vector $u_j \in \mathcal{R}^l$ is $\mathcal{Y}_j$-measurableand called a control action. Let $u = (u_0, u_1, ..., u_{N-1})$ is an admissible control and the class of admissible controls is denoted by U. The task depends on remove the stochastic system (2.1) from initial state y_0 to the target ξ, which is unknown to the controller, and has apiori a normal distribution $N(m, Q)$.

The classical aim of control consists in optimization of performance criterion. Let the objective function represents total costs. The total cost is composed of costs of changes $\alpha \|y_{i+1} - y_i\|^2$, energetic costs of controls

$\beta \|u_i\|^2$ for $i = 0, ..., N-1$ and cost (loss) associated with not hitting the point (target) $\gamma \|y_N - \xi\|^2$, where $\alpha, \beta, \gamma \geqslant 0$. For the linear quadratic control problem with fixed horizon the aim of control is to minimize the total cost. Thus the task is to find

$$\inf_{u \in U} E \left\{ \sum_{i=0}^{N-1} \left(\alpha \|y_{i+1} - y_i\|^2 + \beta \|u_i\|^2 \right) + \gamma \|y_N - \xi\|^2 \right\}. \tag{2.2}$$

Let N denote the biggest possible horizon of control and $\tau : \Omega \longrightarrow \{0, 1, ..., N\}$ be a Markov moment. The class of Markov moments will be denoted by $\mathbf{T}$. The linear quadratic control problem with optimally stopped horizon can be presented as

$$\inf_{(y,\tau) \in \mathbf{Y} \times \mathbf{T}} E \left\{ \sum_{i=0}^{\tau-1} \left(\alpha \|y_{i+1} - y_i\|^2 + \beta \|u_i\|^2 \right) + \gamma \|y_\tau - \xi\|^2 \right\}. \tag{2.3}$$

We see, that the system should be carried out (controlled) at the cheapest cost. On the other hand to move the system (2.1) from y_0 to target ξ, we need to determine an optimal route (set of landmarks). To determine the route we create the substitute task.

Let $det\left(B^T B\right) \neq 0$. When we want to move the system (2.1) from state y_i to y_{i+1}, $i = 0, 1, ..., N-1$ then the control has a form

$$u_i = K y_{i+1} - L y_i + M w_{i+1}, \tag{2.4}$$

where

$$K = \left(B^T B\right)^{-1} B^T, \; L = KA, \; M = -K\sigma.$$

Let $y = (y_0, y_1, ..., y_{N-1})$ mean an route (path, trajectory) and $\mathbf{Y}$ denote the class of admissible routs. For a fixed horizon the task (2.2) may be replaced by

$$\inf_{y \in \mathbf{Y}} E \left(\sum_{i=0}^{N-1} \left(\alpha \|y_{i+1} - y_i\|^2 + \beta \|K y_{i+1} - L y_i + M w_{i+1}\|^2 \right) + \gamma \|y_N - \xi\|^2 \right). \tag{2.5}$$

The main aim is determining the optimal route y^* by solution the task (2.5). A direct solution of task (2.2) gives us the explicit formula of the optimal control, but the solution of task (2.5) gives the optimal trajectory.

The problem of determining the optimal route for random horizon can be presented in the following form

$$\begin{aligned} \inf_{(y,\tau) \in \mathbf{Y} \times \mathbf{T}} E \Bigg(& \sum_{i=0}^{\tau-1} \left(\alpha \|y_{i+1} - y_i\|^2 + \beta \|K y_{i+1} - L y_i + M w_{i+1}\|^2 \right) \\ & + \gamma \|y_\tau - \xi\|^2 \Bigg). \end{aligned} \tag{2.6}$$

3. Solution LQR problem for a fixed horizon

Below the method of determining the optimal route for system (2.1) with fixed horizon N will be presented. First we will determine the optimal path for the case, where the target is known. Next the obtained result will be modifyed for the case, where the target is random. To satisfy the aim (to move the system (2.1) from y_0 to point ξ) the system (2.1) should moved along the path $y_0, y_1, ..., y_N$. We determine this path by solving a task (2.5). As result we obtain a set of points (marks) $y = (y_0, ..., y_{N-1})$ for which the infimum is attained. The theorem below shows how determine the optimal path where the target ξ is known. Let

$$H = \alpha I + \beta K^T L,\ D = \alpha I + \beta K^T K,\ C = \alpha I + \beta L^T L$$

and $\mathcal{F}_j = \sigma(\xi) \vee \mathcal{Y}_j$.

Theorem 3.1. *Let*

$$\begin{aligned} \Phi_j^N &= C - H^T \left(\Phi_{j+1}^N + D\right)^{-1} H, && (3.1)\\ \Psi_j^N &= H^T \left(\Phi_{j+1}^N D\right)^{-1} \Psi_{j+1}^N, && (3.2)\\ \Upsilon_j^N &= \Upsilon_{j+1}^N - \left(\Psi_{j+1}^N\right)^T \left(\Phi_{j+1}^N D\right)^{-1} \Psi_{j+1}^N, && (3.3)\\ Z_j^N &= Z_{j+1}^N + tr\left(\left(\beta M^T M + \Phi_{j+1}^N + D\right)\sigma\sigma^T\right). && (3.4) \end{aligned}$$

where $\Phi_N^N = \Psi_N^N = \Upsilon_N^N = \gamma I,\ Z_N^N = 0$ *and* I *is an identity matrix. If* $det\left(\Phi_{j+1}^N + D\right) \neq 0$ *for* $j = 0, 1, ..., N-1$, *then the optimal state (mark, position) for the time* $j+1$ *based on information available to time* j *is*

$$E\left(y_{j+1} \middle| \mathcal{F}_j\right) = \left(\Phi_{j+1}^N + D\right)^{-1}\left(H y_j + \Psi_{j+1}^N \xi\right) \qquad (3.5)$$

and

$$\begin{aligned} \inf_{y \in \mathbf{Y}} E &\left(\sum_{i=0}^{N-1} \left(\alpha \left\|y_{i+1} - y_i\right\|^2 + \beta \left\|K y_{i+1} - L y_i + M w_{i+1}\right\|^2\right)\right. \\ &\left. + \gamma \left\|y_N - \xi\right\|^2\right) = W_0^N(\xi, y_0), \qquad (3.6) \end{aligned}$$

where

$$\begin{aligned} W_N^N(\xi, y_N) &= \gamma \left\|y_N - \xi\right\|^2, && (3.7)\\ W_j^N(\xi, y_j) &= y_j^T \Phi_j^N y_j - 2 y_j^T \Psi_j^N \xi + \xi^T \Upsilon_j^N \xi + Z_j^N. && (3.8) \end{aligned}$$

Proof. First we define the Bellman function (see e.g. [8]). At the time N the value of Bellman function $W_N^N(\xi, y_N)$ is given by (3.7), but at the times $i = 0, 1, 2, ..., N-1$ is defined as

$$\begin{aligned} W_i^N(\xi, y_i) &= \min_{y_{i+1}} E\Big\{\alpha \left\|y_{i+1} - y_i\right\|^2 \\ &\quad + \beta \left\|K y_{i+1} - L y_i + M w_{i+1}\right\|^2 + W_{i+1}(\xi, y_{i+1}) \Big| \mathcal{F}_i \Big\} \end{aligned} \tag{3.9}$$

for $j = 0, 1, ..., N-1$. From (3.9) for the time $N-1$ we have

$$\begin{aligned} W_{N-1}^N(\xi, y_{N-1}) &= \min_{y_N} E\big\{ y_N^T \left(\alpha I + \Phi_N^N + \beta K^T K\right) y_N \\ &\quad + y_{N-1}^T \left(\alpha I + \beta L^T L\right) y_{N-1} \\ &\quad - 2 y_N^T \left(\left(\alpha I + \beta K^T L\right) y_{N-1} + \Psi_N^N \xi\right) \\ &\quad + \xi^T \Upsilon_N^N \xi + \beta w_N^T M^T M w_N \mid \mathcal{F}_{N-1} \big\} \\ &= \min_{y_N} \big\{ E\left(y_N^T \mid \mathcal{F}_{N-1}\right) \left(\Phi_N^N + D\right) E\left(y_N \mid \mathcal{F}_{N-1}\right) \\ &\quad + y_{N-1}^T C y_{N-1} \\ &\quad - 2E\left(y_N^T \mid \mathcal{F}_{N-1}\right) \left(H y_{N-1} + \Psi_N^N \xi\right) \\ &\quad + \xi^T \Upsilon_N^N \xi + tr\left(\left(\beta M^T M + \Phi_N^N + D\right) \sigma \sigma^T\right) \big\}. \end{aligned}$$

The expected optimal state (position, mark) at time N based on information available to time $N-1$ is

$$E\left(y_N \middle| \mathcal{F}_{N-1}\right) = \left(\Phi_N^N + D\right)^{-1} \left(H y_{N-1} + \gamma \Psi_N^N \xi\right).$$

The value of Bellman function of time $N-1$ is equal

$$W_{N-1}^N(\xi, y_{N-1}) = y_{N-1}^T \Phi_{N-1}^N y_{N-1} - 2 y_{N-1}^T \Psi_{N-1}^N \xi + \xi^T \Upsilon_{N-1}^N \xi + Z_{N-1}^N,$$

where

$$\begin{aligned} \Phi_{N-1}^N &= C - H^T \left(\Phi_N^N + D\right)^{-1} H, \\ \Psi_{N-1}^N &= H^T \left(\Phi_N^N + D\right)^{-1} \Psi_N^N, \\ \Upsilon_{N-1}^N &= \Upsilon_N^N - \left(\Psi_N^N\right)^T \left(\Phi_N^N + D\right)^{-1} \Psi_N^N, \\ Z_{N-1}^N &= tr\left(\left(\beta M^T M + \Phi_N^N + D\right) \sigma \sigma^T\right). \end{aligned}$$

Let us assume, that equation (3.8) is true for $i+1$. From (3.8)-(3.9) and the properties of condition expectation we have

$$\begin{gathered}
W_j^N(\xi, y_j) = \min_{y_{j+1}} E\left\{\alpha \left\|y_{j+1} - y_j\right\|^2 + \beta \left\|Ky_{j+1} - Ly_j + Mw_{j+1}\right\|^2\right. \\
\left. + y_{j+1}^T \Phi_{j+1}^N y_{j+1} + y_{j+1}^T \Psi_{j+1}^N \xi + \xi^T \Upsilon_{j+1}^N \xi + Z_{j+1}^N \right| \mathcal{F}_j\} \\
= \min_{y_{j+1}} \{E\left(y_{j+1}^T \middle| \mathcal{F}_j\right) \left(\Phi_{j+1}^N + D\right) E\left(y_{j+1} \middle| \mathcal{F}_j\right) + y_j^T C y_j + \xi^T \Upsilon_{j+1}^N \xi \\
-2E\left(y_{j+1}^T \middle| \mathcal{F}_j\right)\left(Hy_j + \Psi_{j+1}^N \xi\right) + tr\left(\left(\beta M^T M + \Phi_{j+1}^N + D\right)\sigma\sigma^T\right) + Z_{j+1}^N
\end{gathered}$$

Thus, the expected optimal state (position) at time $j+1$ is

$$E\left(y_{j+1} \middle| \mathcal{F}_j\right) = \left(\Phi_{j+1}^N + D\right)^{-1}\left(Hy_j + \Psi_{j+1}^N \xi\right)$$

and the value of the Bellman function at time j is equal

$$\begin{gathered}
W_j^N(\xi, y_j) = -\left(Hy_j + \Psi_{j+1}^N \xi\right)\left(\Phi_{j+1}^N + D\right)^{-1}\left(Hy_j + \Psi_{j+1}^N \xi\right) \\
+y_j^T C y_j + \xi^T \Upsilon_{j+1}^N \xi + Z_{j+1}^N + tr\left(\left(\beta M^T M + \Phi_{j+1}^N + D\right)\sigma\sigma^T\right) \\
= y_j^T \Phi_j^N y_j - 2y_j^T \Psi_j^N \xi + \xi^T \Upsilon_j^N \xi + Z_j^N,
\end{gathered}$$

where $\Phi_j^N, \Psi_j^N, \Upsilon_j^N, Z_j^N$ are given by (3.1)–(3.4). □

Remark 3.2. The equation (3.5) gives the formula (recipe, rule) how to determine the optimal route (state- or landmarks) for time $j+1$ if the system (2.1) to time j traveled the way (path, track) $y_0, ..., y_j$. Additionally, from formulas (3.1)–(3.4) we have

$$\Phi_j^k = \Phi_{N-(k-j)}^N,\ \Psi_j^k = \Psi_{N-(k-j)}^N,\ \Upsilon_j^k = \Upsilon_{N-(k-j)}^N \text{ and } Z_j^k = Z_{N-(k-j)}^N$$

for $0 \leqslant j \leqslant k \leqslant N$. Hence from formula (3.8) we obtain

$$W_j^k(\xi, y) = W_{N-(k-j)}^N(\xi, y),$$

where $\xi, y \in \mathbb{R}^n$ and $0 \leqslant j \leqslant k \leqslant N$.

Below the recipies (law) of optimal route determinig will be presented for the case, where the target is unknown. Let the target ξ is random and has a normal distribution $N(m, Q)$. For $j = 0, 1, ..., N-1$ the expected optimal state (mark, position) for the time $j+1$ based on information available to time j is

$$E\left(y_{j+1} \middle| \mathcal{Y}_j\right) = E\left(E\left(y_{j+1} \middle| \mathcal{F}_j\right) \middle| \mathcal{Y}_j\right) = \left(\Phi_{j+1}^N + D\right)^{-1}\left(Hy_j + \Psi_{j+1}^N m\right) \tag{3.10}$$

and from (3.6) we have, that the expected cost of control of system (2.1)after time j is equal

$$\begin{aligned} V_j^N(y_j) &= E\left(W_j^N(\xi, y_j)\middle|\,\mathcal{Y}_j\right) = \\ &= y_j^T\Phi_j^N y_j + y_j^T\Psi_j^N m + m^T\Upsilon_j^N m + Z_j^N + tr\left(\Upsilon_j^N Q\right), \end{aligned} \tag{3.11}$$

where $\Phi_j^N, \Psi_j^N, \Upsilon_j^N, Z_j^N$ are given by (3.1)–(3.4). The value $tr\left(\Upsilon_j^N Q\right)$ denotes the cost of ignorance of target but the value Z_j^N presents the cost of eliminating external disturbances for the system (2.1) which will be controlled at moments from j to N.

Remark 3.3. Let us assume, that for the case with known target the aim is equal $m \in \mathbb{R}^n$ but for the case with unknown target the aim is random vector ξ with normal distribution $N(m, Q)$, where $E\xi = m$. The additional cost connected with ignorance of aim is a difference between total cost in case where target is unknown and total cost in case where target is known. Thus this cost is equal

$$V_0^N(y_0) - W_0^N(m, y_0) = tr\left(\Upsilon_0^N Q\right).$$

Remark 3.4. When the optimal route for the linear system (2.1) is known (planned), then from (2.4) the expected control conditioned on σ−field $\mathcal{Y}_j$ is

$$\begin{aligned} E\left(u_j\middle|\,\mathcal{Y}_j\right) &= -\left(B^T B\right)^{-1} B^T\left(E\left(y_{j+1}\middle|\,\mathcal{Y}_j\right) - Ay_j\right) \\ &= K\left(\Phi_{j+1}^N + D\right)^{-1}\left(Hy_j + \Psi_{j+1}^N\xi\right) - Ly_j. \end{aligned} \tag{3.12}$$

Example 3.5. The linear system with state equation (2.1) must be moved from initial state $y_0 = (90; 50)$ to known and unknown targets. We assume that the parameters α, β, γ are equal 0.1, 0.5, 1 accordingly and

$$Q = \begin{bmatrix} 2.2 & 0.3 \\ 0.4 & 1.9 \end{bmatrix}, \quad A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1.2 & 0.1 \\ 0.5 & 2 \end{bmatrix}, \quad \sigma = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.2 \end{bmatrix}.$$

For case with known target the aim is $m = (10; 20)$, but for case with unknown target the aim is a random vectorξ with normal distribution $N(m, Q)$. Let us determine the optimal routes for fixed horizons $N = 20$.

The figure (1)a presents the optimal planned trajectories. We see, that the sets of landmarks $E\left(y_j\middle|\,\mathcal{F}_0\right)$ and $E\left(y_j\middle|\,\mathcal{Y}_0\right)$, $j = 0, 1, ..., N$ for cases with fixed and random targets are identical, because $E\xi = m$. This marks are uniformly distributed along trajektory which connects the points y_0 and m.

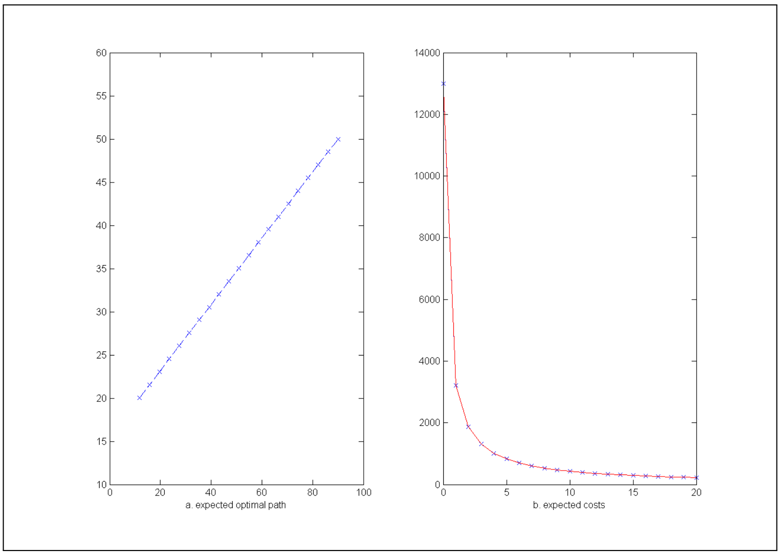


Figure 1: The optimal planned trajectory and expected total costs.

The figure (1)b presents the expected value of objective function $W_{N-t}^{N}(y_0)$ (the horizon of control is equal $N-(N-t)=t$) for $0 \leqslant t \leqslant N$. We see, if the horizon of control is increasing then the expected value of cost of control is decreasing.

The figure (2)a presents the costs connected with ignorance of target $tr\left(\Upsilon_{N-t}^{N}Q\right)$ for $0 \leqslant t \leqslant N$. This costs are decreasing while the horizon of control t is increasing. The figure (2)b presents the costs of eliminating external disturbances Z_{N-t}^{N} which increase with increasing horizon t.

4. Optimal route determining for stopped horizon

Sometimes if we stop the system (2.1) earlier than fixed horizon we may obtain the better result. Thus, independently from the optimal route determining it should also appoint a moment of stopping the system. The optimal stopping rules (see e.g. [21]) will be employed to determine the moment when the system (2.1) must be stopped. The theorem below presents how to determine the optimal route and moment.

Let $\tau : \Omega \longrightarrow \{0,1,...,N\}$ be a Markov moment in view of $\sigma-$field $\mathcal{Y}=(\mathcal{Y}_j)_{0\leqslant j\leqslant N}$. We determine the optimal route by solving a task (2.6) and we obtain a set of admissible points (marks) $y=(y_0,...,y_\tau)$ for which the

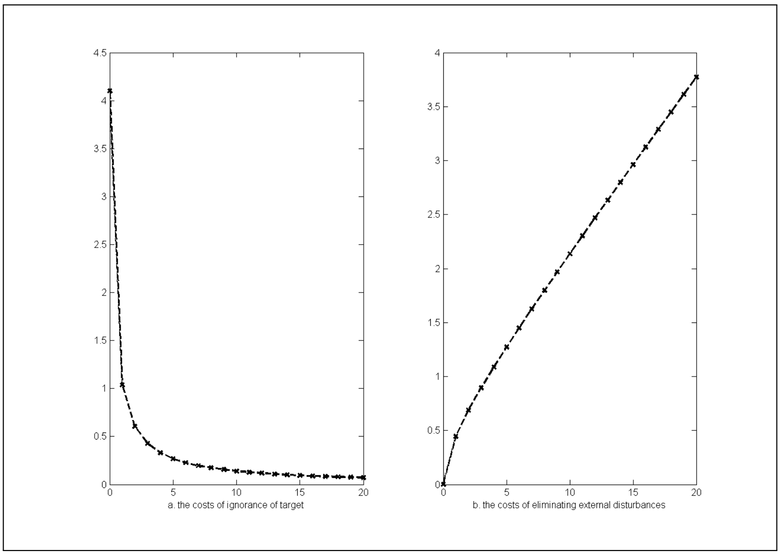


Figure 2: The additional costs of control.

infimum is attained. From (3.9) we have

$$\begin{aligned} V_k^j(y_k) &= \inf_{y_{k+1},\dots,y_j} E\left(\sum_{i=k}^{j-1}\left(\alpha\|y_{i+1}-y_i\|^2\right.\right. \\ &\left.\left.+\beta\|Ky_{i+1}-Ly_i+Mw_{i+1}\|^2\right)+\gamma\|y_j-\xi\|^2\middle|\mathcal{Y}_k\right), \quad (4.1)\end{aligned}$$

where $0 \leqslant k \leqslant j \leqslant N$. From above we can present task (2.6) in the following form

$$\inf_{\tau\in\mathbf{T}} V_0^\tau(y_0). \tag{4.2}$$

Theorem 4.1. *Let the stochastic system be described by equation (2.1). The optimal solution of task (2.6) is:*
a. the optimal stopping moment of controlled system (2.1)

$$\tau^* = \min\left\{0 \leqslant k \leqslant N : V_k^k(y_k) = \min_{k\leqslant j\leqslant N} V_k^j(y_k)\right\}, \tag{4.3}$$

where $V_k^j(y_k)$, $0 \leqslant k \leqslant j \leqslant N$ *is given by (3.11);*
b. if $k \in \{0, 1, ..., N-1\}$ *is not a stopping moment and det* $\left(\Phi_{k+1}^j + D\right) \neq 0$ *for* $0 \leqslant k \leqslant j \leqslant N$ *, then the optimal state (mark, position) at the time* $k+1$ *is*

$$E(y_{k+1}|\mathcal{Y}_k) = \left(\Phi_{k+1}^t + D\right)^{-1}\left(Hy_k + \Psi_{k+1}^t E(\xi|\mathcal{Y}_k)\right), \tag{4.4}$$

where moment $t \in \{k+1, ..., N\}$ *is determined as*

$$t = \min \left\{ k+1 \leqslant j \leqslant N : V_k^j (y_k) = \min_{k < s \leqslant N} V_k^s (y_k) \right\}$$

and Φ_{k+1}^t *is given by (3.1).*

Proof. Let us consider the task of optimal stopping in class $\mathcal{T}(N, N)$. Thus we may only stop the system (2.1), thus the stopping moment $\tau^0 = N$ and the Bellman's function is equal

$$\tilde{V}_N (y_N) = V_N^N (y_N) = \gamma E \left\| y_N - \xi \right\|^2.$$

In the class $\mathcal{T}(N-1, N)$ we may stop the system (2.1) at $N-1$ or control the system (2.1) optimally to moment N. If we stop the system (2.1) at moment $N-1$ then we have only heredity cost $\gamma E \left\| y_{N-1} - \xi \right\|^2$. If we control the system (2.1) at time $N-1$ and next we will stop this system at time N, than we have the costs of transformation and control

$$E \left(\alpha \left\| y_N - y_{N-1} \right\|^2 + \beta \left\| K y_N - L y_{N-1} + M w_N \right\|^2 \middle| \mathcal{Y}_{N-1} \right)$$

and expected heredity cost $\gamma E \left(\left\| y_N - \xi \right\|^2 \middle| \mathcal{Y}_{N-1} \right)$. Thus, if we will stop the system (2.1) at moment N then We will incur the total cost $V_{N-1}^N (y_{N-1})$ (see theorem 1). The optimal cost in class $\mathcal{T}(N-1, N)$ is

$$\tilde{V}_{N-1} (y_{N-1}) = \min \left\{ \gamma E \left\| y_{N-1} - \xi \right\|^2, V_{N-1}^N (y_{N-1}) \right\}$$

and the optimal stopping moment is

$$\tau^1 = \begin{cases} N-1, & \text{if } \tilde{V}_{N-1} (y_{N-1}) = \gamma E \left\| y_{N-1} - \xi \right\|^2, \\ \tau^0, & \text{if } \tilde{V}_{N-1} (y_{N-1}) > \gamma E \left\| y_{N-1} - \xi \right\|^2. \end{cases}$$

If $N-1$ is not a stopping moment, then we control the system (2.1) and the expected optimal state at time N is

$$E (y_N | \mathcal{Y}_{N-1}) = \left(\Phi_N^N + D \right)^{-1} \left(H y_{N-1} + \Psi_N^N E (\xi | \mathcal{Y}_{N-1}) \right).$$

Similarly we consider for classes $\mathcal{T}(N-2, N), ..., \mathcal{T}(0, N) = \mathcal{T}$. In the class $\mathcal{T}(k, N)$, $0 \leqslant k \leqslant N-1$ we may stop the system (2.1) at time k or optimally control the system (2.1) to possible times $k+1, ..., N$. If we stop the system (2.1) then we accept the cost of heredity $\gamma E \left(\left\| y_k - \xi \right\|^2 \middle| \mathcal{Y}_{k-1} \right)$ only. If at time k we control the system (2.1) we have the costs of transformation and controls. The optimal total cost in class $\mathcal{T}(k, N)$ is

$$\tilde{V}_k (y_k) = \min \left\{ \gamma E \left(\left\| y_k - \xi \right\|^2 \middle| \mathcal{Y}_{k-1} \right), V_k^{k+1} (y_k), ..., V_k^N (y_k) \right\}$$

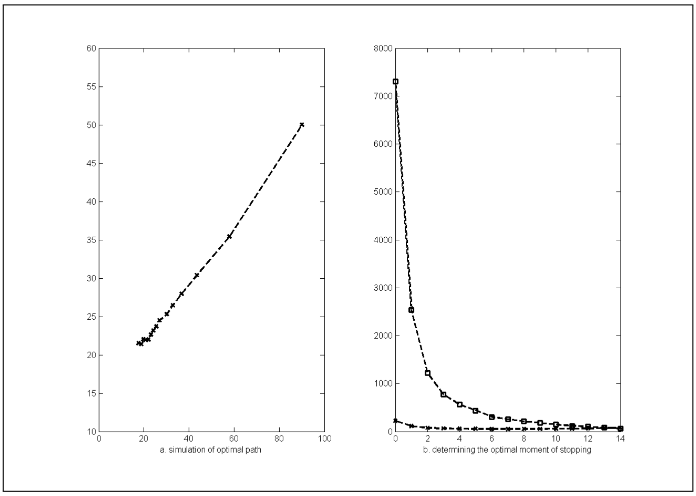


Figure 3: The optimal route with optimal stopping moment.

and the optimal stopping moment in class $\mathcal{T}(k, N)$ is

$$\tau^k = \begin{cases} k, & \text{if } \tilde{V}_k(y_k) = \gamma E\left(\|y_k - \xi\|^2 \middle| \mathcal{Y}_{k-1}\right), \\ \tau^{k-1}, & \text{if } \tilde{V}_k(y_k) > \gamma E\left(\|y_k - \xi\|^2 \middle| \mathcal{Y}_{k-1}\right). \end{cases}$$

If k is not a stopping moment, then the system (2.1) must be shifted from state y_k to expected optimal state

$$E\left(y_{k+1} \middle| \mathcal{Y}_k\right) = \left(\Phi_{k+1}^t + D\right)^{-1}\left(H y_k + \Psi_{k+1}^t E\left(\xi \middle| \mathcal{Y}_k\right)\right).$$

where $t = \min\left\{k < t \leqslant N : \tilde{V}_k(y_k) = V_k^t(y_k)\right\}$. From theory of optimal stopping rules (see e.g. [21]) the Markov moment given by

$$\tau^* = \min\left\{0 \leqslant k \leqslant N : \gamma E\left(\|y_k - \xi\|^2 \middle| \mathcal{Y}_{k-1}\right) = \tilde{V}_k(y_k)\right\}$$

is an optimal stopping moment, what proves the assertion. □

Remark 4.2. If $\gamma = 0$ then form theorem 4.1 the system (2.1) must be controlled to the end of fixed horizon.

Example 4.3. The linear system is defined by the state equation (2.1) and must be moved from initial state $y_0 = (90; 50)$ to the random target ξ. We assume, that the parameters α, β, γ and the matrices Q, A, B, σ are the same as in example 1.

Figure 3a shows the simulation of trajectory of stochastic system (2.1). The figure 3b presents the realization $\tilde{V}_k(y_k)$ (curve with mark 'cross') and

$\gamma E\left(\left\|y_k-\xi\right\|^2\middle|\mathcal{Y}_k\right)$ (curve with mark 'square') for $0\leqslant k\leqslant N$. In this picture we see, that for the moment $\tau^*=14$ the lowest expected cost $\tilde{V}_{\tau^*}\left(y_{\tau^*}\right)$ is identical as heredity cost $\gamma E\left(\left\|y_{\tau^*}-\xi\right\|^2\middle|\mathcal{Y}_{\tau^*}\right)$. From above the moment $\tau^*=14$ is an optimal moment of stopping. Comparing the figures 1a and 3a we see, that optimal paths for fixed and optimally stopped horizon are completely different. For fixed horizon the system (2.1) is moved uniformly along trajectory form initial state y_0 to target ξ. For optimally stopped horizon the system (2.1) is more moved at the beginning but slowly moved at the end of horizon.

5. Conclusion

The problem presented in this article depends on determining the optimal path (trajectory) to perfect track the target. To realise this aim the linear quadratic control problem was converted to the linear quadratic routing problem. The laws of detemining the optimal route and optimal horizon were given. The examples presented above show, that the horizon has a significant influence to result of realization of aim.

The extension of presented results can be used, for example, to the source seeking problem, the navigation planning, the perfect tracking etc.

References

1. M. Aoki, *Optimization of Stochastic Systems*, Academic Press, 1967.
2. B. Abouzaid, M.E. Achhab, V. Wertz, Feedback stabilization of infinite-dimensional linear systems with constraints on control and its rate, *European Journal of Control*, vol. 17 (2), pp. 183-190, 2011.
3. S. Azuma, M.S. Sakar, G.J. Pappas, Stochastic Source Seeking by Mobile Robots, *IEEE Transactions on Automatic Control*, vol. 57 (9), pp. 2308-2321, 2012.
4. T. Banek, E. Kozłowski, Adaptive control of system entropy, *Control and Cybernetics*, vol. 35 (2), pp. 279-289, 2006.
5. T. Banek, E. Kozłowski, Deterministic versus stochastic in the free horizon LQC problem, *Aktualne problemy automatyki i robotyki* ed. by K. Malinowski, J. Józefczyk, J. Światek, 94-101, Warszawa: Exit, 2014.
6. T. Banek, E. Kozłowski, LQC Multi-jumps follower as a model of company reorganization, *Proces control in management* ed. by T. Banek, 13-26, Torun, 2009.
7. T. Banek, E. Kozłowski, Adaptive control with randon horizon, *Annales Informatica*, vol. 3, pp. 5-14, 2005.
8. R. Bellman, *Adaptive Control Processes*, Princeton, 1961.
9. W. H. Fleming, R. Rishel, *Deterministic and Stochatic Optimal Control*, Berlin:Springer-Verlag, 1975.

10. L. Harris, R. Rishel, An algorithm for a solution of a stochastic adaptive linear quadratic optimal control problem, *IEEE Transactions on Automatic Control*, vol. 31 (12), pp. 1165-1170, 1986.
11. I. Karatzas , S.E. Shreve, Connections between optimal stopping and singular control I. Monotone follower problems, *SIAM Journal of Control Optimization*, vol. 22 (6), pp. 856–877, 1984.
12. I. Karatzas , S.E. Shreve, Connections between optimal stopping and singular control II. Reflected follower problems. *SIAM Journal of Control Optimization*, vol. 23 (3), pp. 433–451, 1985.
13. F. Kozin, Stability of stochastic dynamical systems, *Lecture Notes in Mathematics* vol. 294, pp. 186-229, 1972.
14. E. Kozłowski, The linear quadratic stochastic optimal control problem with random horizon at finite number of events intependent of state system, *Systems Science*, vol. 36 (3), pp. 5-11, 2010.
15. E. Kozłowski, Identyfication of linear system in random time, *International Journal of Computer and Information Technology*, vol. 1 (2), pp. 103-108, 2012.
16. E. Kozłowski, Stabilization of linear systems in random horizon via control, *Control and Cybernetics*, vol. 42(2), pp. 527-541, 2013.
17. E. Kozłowski, Optimal route and control for the LQC problem, *Probability in action* ed. by T. Banek, E. Kozłowski, pp.41-50, Politechnika Lubelska 2014.
18. R.SH. Liptser, A.N. Shiryaev, *Statistics of Stochastic Processes*, New York:Springer-Verlag, 1978.
19. C. Manzie, M. Krstic, Extremum seeking with stochastic perturbation, *IEEE Transactions on Automatic Control* vol. 54 (3), pp. 580-585, 2009.
20. G. N. Saridis, *Stochastic processes, estimation and control: the entropy approach*, John Wiley and Sons, 1995.
21. A.N. Shiryaev, *Statistical analysis of sequential processes. Optimal stopping rules*, New York: Springer Verlag, 1978.

Asymptotics of the products of sums of independent and non-identically distributed random variables

Tomasz Krajka[1], Przemysław Matuła[2], Habib Naderi[3]

Abstract

We study weak convergence of products of sums of independent and non-identically distributed random variables. Some examples concerning the rate of convergence are also presented in this setting.

1. Introduction

Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of i.i.d., positive and square–integrable random variables with mean μ, variance σ^2 and coefficient of variation $\gamma = \sigma/\mu$. The study of the limiting behavior of products of sums $S_n = \sum_{k=1}^n X_k$ originated in the paper of Arnold and Villaseñor [1], who proved the following convergence of sums of record values based on a sequence of i.i.d. standard exponential r.v.'s

$$\frac{\sum_{i=1}^n \log S_i - n\log n + n}{\sqrt{2n}} \xrightarrow{d} \mathcal{N}(0,1) \text{ as } n\to\infty, \tag{1.1}$$

here and in the sequel $\mathcal{N}\left(\mu,\sigma^2\right)$ denotes the normal variable with mean μ and variance σ^2. Let us observe that

$$\frac{\sum_{i=1}^n \log S_i - n\log n + n}{\sqrt{2n}} = \log\left(\frac{\prod_{i=1}^n S_i \cdot e^n}{n^n}\right)^{\frac{1}{\sqrt{2n}}}$$

[1]Lublin University of Technology, ul. Nadbystrzycka 38d, 20-618 Lublin, Poland, *e-mail: t.krajka@pollub.pl*

[2]Marie Curie-Skłodowska University, pl. M.C.-Skłodowskiej 1, 20-031 Lublin, Poland, *e-mail: matula@hektor.umcs.lublin.pl*

[3]Faculty of Mathematics, University of Sistan and Baluchestan, Zahedan, Iran, *e-mail: h_naderisp@math.usb.ac.ir*

2010 Mathematics Subject Classification. Primary 60F05, 60G50 ; Secondary 60F15.

Keywords: independent random variables, products of sums, central limit theorem.

and by applying the Stirling formula

$$\lim_{n\to\infty} \frac{n!}{\sqrt{2\pi n} n^n e^{-n}} = 1$$

and

$$\lim_{n\to\infty} \left(\frac{1}{\sqrt{2\pi n}}\right)^{\frac{1}{\sqrt{2n}}} = 1$$

we see that (1.1) is equivalent to

$$\left(\prod_{i=1}^{n} \frac{S_i}{i}\right)^{\frac{1}{\sqrt{n}}} \xrightarrow{d} \exp\left(\mathcal{N}\left(0,2\right)\right) \text{ as } n \to \infty$$

Rempała and Wesołowski [16] found that this result has deeper meaning and proved that if $(X_n)_{n\in\mathbb{N}}$ is a sequence of i.i.d. positive and square–integrable random variables, then

$$\left(\frac{\prod_{i=1}^{n} S_i}{n!\mu^n}\right)^{\frac{1}{\gamma\sqrt{n}}} \xrightarrow{d} \exp\left(\mathcal{N}\left(0,2\right)\right), \text{ as } n \to \infty. \tag{1.2}$$

This result was extended and generalized by several authors. Let us briefly recall some of these results. Natural directions for generalizations of (1.2) is relaxing the assumption of independence by considering sequences of dependent r.v.'s, considering r.v.'s which are non-identically distributed or weakening the moment requirement i.e. square-integrability of r.v.'s. The other interesting questions are the rate of convergence in (1.2) and functional or almost sure version of this result. Randomly indexed and self-normalized products were investigated as well. In recent years also large deviation and precise asymptotics were studied.

Qi [15] and Lu and Qi [8] relaxed the assumption of square integrability of the r.v.'s and considered sequences belonging to the domain of attraction of a stable law with index greater or equal to 1. They obtained analogous results to (1.2) with the stable law in the limit. This direction of research in the functional version carried on Kosiński in [4]. The convergence in the space $D[0,1]$ of processes constructed from products of sums for the first time were studied by Zhang and Huang [19] and later for non-identically distributed r.v.'s by Matuła and Stępień [10, 13]. The non-i.i.d. case considered also Krajka and Rychlik [6] even in the randomly indexed setting. The so-called almost sure version of the central limit theorem focused the attraction of many researchers in the context of products of sums, let us only mention Gonchigdanzan [3] and Matuła and Stępień [12]. There are only two papers devoted to the study of the rate of convergence in (1.2), these are results of Matuła and Stępień [11] and Krajka and Rychlik [5]. Dependent r.v.'s such as positively or negatively dependent, mixing sequences were studied for example in [12], [20] and [18]. A general approach to dependent sequences via strong approximation was discussed in [9]. Among the most recent results let us mention the paper of Tan [17] on precise asymptotics and Zhu [21] on large deviation for products of sums.

The goal of our paper is to generalize (1.2) to the case of independent but non-identically distributed r.v.'s and to combine the results of [6], [10] and [13] into one theorem.

For a sequence $(X_n)_{n\in\mathbb{N}}$ be a sequence of independent, and square–integrable random variables defined on some probability space $(\Omega, \mathfrak{F}, P)$, let us introduce the following notation:

$$\mu_n = EX_n,\ \tau_n^2 = \mathrm{Var}(X_n),\ S_n = \sum_{k=1}^{n} X_k,\ \sigma_n^2 = \mathrm{Var}(S_n) = \sum_{k=1}^{n} \tau_k^2, \quad \text{for } n \in \mathbb{N}.$$

The main results concerning weak convergence will be stated and proved in Section 2, in Section 3 we quote a result concerning the rate of convergence and in Section 4 we present some illustrative examples.

2. Weak convergence

In what follows $(a_n)_{n\in\mathbb{N}\cup\{0\}}$ is an increasing sequence such that $a_0 = 0$ and $\lim_{n\to\infty} a_n = \infty$.

Theorem 2.1. *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of independent and square-integrable r.v.'s and $(\alpha_{i,n})_{1\leqslant i\leqslant n, n\in\mathbb{N}}$ an array of positive numbers. Assume that the following conditions are satisfied:*

$$X_n - \mathbb{E}X_n > a_{n-1} - a_n,\ \textit{almost surely for all } n \in \mathbb{N}, \tag{2.1}$$

$$\sum_{i=1}^{\infty} \frac{\mathbb{E}\,|X_i - \mathbb{E}X_i|^p}{a_i^p} < \infty,\ \textit{for some } p \in (0,2\rangle\,, \tag{2.2}$$

$$\sum_{i=1}^{n} \frac{\alpha_{i,n}}{a_i^2}\sigma_i^2 \to 0,\ \textit{as } n \to \infty, \tag{2.3}$$

$$\sum_{i=1}^{n} (A_i^n)^2\, \tau_i^2 \to \sigma^2,\ \textit{as } n \to \infty,\ \textit{for some } \sigma > 0, \tag{2.4}$$

where $A_i^n = \sum_{k=i}^{n} \frac{\alpha_{k,n}}{a_k}$,

$$\textit{for any } i_0 \in \mathbb{N},\ \max_{1\leqslant i\leqslant i_0} \alpha_{i,n} \to 0,\ \textit{as } n \to \infty, \tag{2.5}$$

$$\lim_{n\to\infty}\sum_{i=1}^{n}(A_i^n)^2\,\mathbb{E}\,(X_i-\mathbb{E}X_i)^2\,\mathbb{I}\,[A_i^n\,|X_i-\mathbb{E}X_i|\geqslant\varepsilon]=0,\ \textit{for any }\varepsilon>0. \tag{2.6}$$

Then

$$T_n:=\prod_{i=1}^{n}\left(\frac{S_i-\mathbb{E}S_i}{a_i}+1\right)^{\alpha_{i,n}}\xrightarrow{d}\exp\left(\mathcal{N}(0,\sigma^2)\right),\ \textit{as } n\to\infty. \tag{2.7}$$

Proof. We essentially follow the lines of [10] and [13] and apply the expansion of the logarithm

$$\log(1+x)=x+R(x),\ \text{for } |x|\leqslant 1/2,\ \text{where } |R(x)|\leqslant 2x^2$$

to prove that

$$\log T_n=\sum_{i=1}^{n}\frac{\alpha_{i,n}}{a_i}\,(S_i-\mathbb{E}S_i)+A_n \tag{2.8}$$

with $A_n\xrightarrow{P}0$.

Let us put $C_i=(S_i-\mathbb{E}S_i)\,/a_i$. From (2.1) we have $C_i+1>0$ almost surely and therefore we easily get

$$\log\left(\frac{S_i-\mathbb{E}S_i}{a_i}+1\right)=\log\left(C_i+1\right)$$
$$=C_i+R(C_i)\mathbb{I}\,[|C_i|\leqslant 1/2]+\left(\log\left(C_i+1\right)-C_i\right)\mathbb{I}\,[|C_i|>1/2]\,.$$

Thus

$$\begin{aligned}\log T_n &= \sum_{i=1}^{n}\alpha_{i,n}\log\left(C_i+1\right)=\\ &= \sum_{i=1}^{n}\alpha_{i,n}C_i+\sum_{i=1}^{n}\alpha_{i,n}R(C_i)\mathbb{I}\,[|C_i|\leqslant 1/2]\\ &\quad+\sum_{i=1}^{n}\alpha_{i,n}\left(\log\left(C_i+1\right)-C_i\right)\mathbb{I}\,[|C_i|>1/2]\\ &= \sum_{i=1}^{n}\frac{\alpha_{i,n}}{a_i}\,(S_i-\mathbb{E}S_i)+A_n'+A_n''\ ,\ \text{say.}\end{aligned}$$

Let us observe that

$$\mathbb{E}\left|A_n'\right|\leqslant 2\sum_{i=1}^{n}\alpha_{i,n}\mathbb{E}\left(\frac{S_i-\mathbb{E}S_i}{a_i}\right)^2=2\sum_{i=1}^{n}\alpha_{i,n}\frac{\sigma_i^2}{a_i^2}\to 0$$

by assumption (2.3). By (2.2) the strong law of large numbers holds (see [14]) i.e. $C_n \to 0$ almost surely. Thus, for almost all $\omega \in \Omega$, there exists $i_0(\omega)$ such that for all $i \geqslant i_0(\omega)$ there holds $\mathbb{I}\left[|C_i| > 1/2\right] = 0$. Since that

$$\begin{aligned} A''_n &= \sum_{i=1}^{i_0} \alpha_{i,n} \left(\log\left(C_i + 1\right) - C_i\right) \mathbb{I}\left[|C_i| > 1/2\right] \\ &\leqslant \max_{1 \leqslant i \leqslant i_0} a_{i,n} \sum_{i=1}^{i_0} \left(\log\left(C_i + 1\right) - C_i\right) \mathbb{I}\left[|C_i| > 1/2\right] \to 0 \end{aligned}$$

by (2.5). It means that $A''_n \to 0$ almost surely and (2.8) is proved.

It remains to prove that

$$Z_n := \sum_{i=1}^{n} \frac{\alpha_{i,n}}{a_i} \left(S_i - \mathbb{E}S_i\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2\right).$$

Let us introduce a triangle array of r.v.'s independent in each row

$$Y_{i,n} = A_i^n \left(X_i - \mathbb{E}X_i\right), \ 1 \leqslant i \leqslant n, \ n \in \mathbb{N}.$$

Then

$$Z_n = \sum_{i=1}^{n} Y_{i,n}$$

and $\mathbb{E}Y_{i,n} = 0$, $\mathrm{Var}\left(Y_{i,n}\right) = \left(A_i^n\right)^2 \tau_i^2$. By assumption (2.4)

$$\mathrm{Var}\left(Z_n\right) = \sum_{i=1}^{n} \left(A_i^n\right)^2 \tau_i^2 \to \sigma^2$$

and the conclusion follows from (2.6) and the Lindeberg's theorem for triangle arrays of independent random variables (see Theorem 27.2 in [2]). □

Remark 2.2. Obviously (2.6) holds if the following Lyapunov's condition is satisfied for some $\delta > 0$

$$\lim_{n \to \infty} \sum_{i=1}^{n} \left(A_i^n\right)^{2+\delta} \mathbb{E}\left|X_i - \mathbb{E}X_i\right|^{2+\delta} = 0. \tag{2.9}$$

Remark 2.3. It is quite natural to use the normalization $a_i = \mathbb{E}S_i$. In this case (2.1) is equivalent to $X_n > 0$ almost surely and we have a simple form of the factors $\frac{S_i - \mathbb{E}S_i}{a_i} + 1 = \frac{S_i}{\mathbb{E}S_i}$.

In the case $\alpha_{i,n} = \beta_n$ i.e. when $\alpha_{i,n}$ does not depend on i, we get Theorem 2.1 of [6] (the part concerning nonrandom number of factors in the product). In this case (2.5) is satisfied provided $0 < \beta_n \to 0$. Let us recall this result as a corollary.

Corollary 2.4. *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of independent and square-integrable r.v.'s and $0 < \alpha_{i,n} = \beta_n \to 0$ as $n \to \infty$. If (2.1)–(2.4) and (2.6) are satisfied, then*

$$\left(\prod_{i=1}^{n}\left(\frac{S_i - \mathbb{E}S_i}{a_i} + 1\right)\right)^{\beta_n} \xrightarrow{d} \exp\left(\mathcal{N}(0,\sigma^2)\right), \text{ as } n \to \infty. \tag{2.10}$$

In [10] the weak convergence of the normalized products of the form

$$\left(\prod_{i=1}^{n}\left(\frac{S_i}{\mathbb{E}S_i}\right)^{\frac{\tau_{i+1}^2\mathbb{E}S_i}{\sigma_i^2}}\right)^{1/\sigma_n}$$

was studied, i.e. the weights $\alpha_{i,n} = \frac{\tau_{i+1}^2\mathbb{E}S_i}{\sigma_i^2\sigma_n}$ and normalization $a_i = \mathbb{E}S_i$ were used in this case. This problem was generalized in [13], where the following weights were considered

$$\alpha_{i,n} = f\left(\frac{\sigma_i^2}{\sigma_n^2}\right)\frac{\tau_i^2\mathbb{E}S_i}{\sigma_n^3}.$$

We shall show that the one-dimensional case of [10] and [13] can be derived from our Theorem 2.1. We shall consider the family of nonnegative functions $f : [0,1] \to \mathbb{R}^+$ which are continuously differentiable on $(0,1]$ and satisfy

$$\begin{gathered}\int_0^1\left(\int_x^1 f(y)dy\right)^2 dx < \infty,\\ \int_0^1 f(x)\sqrt{x}dx < \infty,\\ f(x)x^{3/2} \to 0, \text{ as } x \to 0^+.\end{gathered} \tag{2.11}$$

Corollary 2.5. *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of independent, positive and square-integrable r.v.'s and the function f satisfies (2.11). Assume that*

$$\sum_{i=1}^{\infty}\frac{\tau_i^2}{(\mathbb{E}S_i)^2} < \infty \text{ and } \mathbb{E}S_n \to \infty, \text{ as } n \to \infty \tag{2.12}$$

and

$$\lim_{n\to\infty}\frac{1}{\sigma_n^2}\sum_{i=1}^{n}\mathbb{E}\left(X_i - \mathbb{E}X_i\right)^2\mathbb{I}\left[|X_i - \mathbb{E}X_i| \geqslant \varepsilon\sigma_n\right] = 0, \text{ for all } \varepsilon > 0. \tag{2.13}$$

Then

$$\prod_{i=1}^{n}\left(\frac{S_i}{\mathbb{E}S_i}\right)^{\alpha_{i,n}} \xrightarrow{d} \exp\left(\mathcal{N}(0,\sigma^2)\right), \text{ as } n \to \infty, \tag{2.14}$$

where $\alpha_{k,n} = f\left(\frac{\sigma_k^2}{\sigma_n^2}\right)\frac{\tau_k^2\mathbb{E}S_k}{\sigma_n^3}$ *and* $\sigma^2 = \int_0^1\left(\int_x^1 f(y)dy\right)^2 dx$

Proof. We shall verify that the assumptions of Theorem 2.1 are satisfied. From the positivity of the random variables and the choice $a_i = \mathbb{E}S_i$ (2.1) follows, (2.2) is (2.12) for $p=2$. Under conditions (2.11) imposed on f, the convergence in (2.3) was proved in formula (21) in [13]. From Lindeberg's condition (2.13) the Feller condition follows

$$\lim_{n\to\infty} \max_{1\leqslant i\leqslant n} \frac{\tau_i^2}{\sigma_n^2} = 0 \tag{2.15}$$

and (2.4) may be considered as an integral sum

$$\sum_{i=1}^{n} (A_i^n)^2 \tau_i^2 = \sum_{i=1}^{n} \left(\sum_{k=i}^{n} f\left(\frac{\sigma_k^2}{\sigma_n^2}\right) \frac{\tau_k^2}{\sigma_n^2} \right)^2 \frac{\tau_i^2}{\sigma_n^2} \to \int_0^1 \left(\int_x^1 f(y)dy \right)^2 dx = \sigma^2.$$

Since $\alpha_{k,n} = f\left(\frac{\sigma_k^2}{\sigma_n^2}\right)\left(\frac{\sigma_k^2}{\sigma_n^2}\right)^{3/2} \frac{\tau_k^2 \mathbb{E}S_k}{\sigma_k^3}$, we see that (2.5) is a consequence of continuity of f and convergence $f(x)x^{3/2} \to 0$.

It remains to prove that the classical Lindeberg's condition (2.13) implies (2.6). Let us define $m_n(t) = \max\left\{i : \sigma_i^2 \leqslant t\sigma_n^2\right\}$ and observe that $\lim_{n\to\infty} \sigma^2_{m_n(t)}/\sigma_n^2 = \lim_{n\to\infty} \sigma^2_{m_n(t)+1}/\sigma_n^2 = t$ (see also (21) in [10]). For given $\varepsilon > 0$ we can find $\delta > 0$ such that $\int_0^\delta \left(\int_x^1 f(y)dy\right)^2 dx < \varepsilon$ and by continuity of f on the closed interval $\langle \delta, 1\rangle$ the following integral is finite $\eta := \int_\delta^1 f(x)dx < \infty$. Let us split the sum in (2.6) into two parts

$$\begin{aligned}
&\sum_{i=1}^{n} (A_i^n)^2 \mathbb{E}(X_i - \mathbb{E}X_i)^2 \mathbb{I}\left[A_i^n |X_i - \mathbb{E}X_i| \geqslant \varepsilon\right] \\
= &\sum_{i=1}^{m_n(\delta)} (A_i^n)^2 \mathbb{E}(X_i - \mathbb{E}X_i)^2 \mathbb{I}\left[A_i^n |X_i - \mathbb{E}X_i| \geqslant \varepsilon\right] \\
&+ \sum_{i=m_n(\delta)+1}^{n} (A_i^n)^2 \mathbb{E}(X_i - \mathbb{E}X_i)^2 \mathbb{I}\left[A_i^n |X_i - \mathbb{E}X_i| \geqslant \varepsilon\right] \\
\leqslant &\sum_{i=1}^{m_n(\delta)} (A_i^n)^2 \tau_i^2 \\
&+ \sum_{i=1}^{n} \left(A^n_{m_n(\delta)+1}\right)^2 \mathbb{E}(X_i - \mathbb{E}X_i)^2 \mathbb{I}\left[A^n_{m_n(\delta)+1} |X_i - \mathbb{E}X_i| \geqslant \varepsilon\right].
\end{aligned}$$

We have

$$\sum_{i=1}^{m_n(\delta)} (A_i^n)^2 \tau_i^2 = \sum_{i=1}^{m_n(\delta)} \left(\sum_{k=i}^{n} f\left(\frac{\sigma_k^2}{\sigma_n^2}\right) \frac{\tau_k^2}{\sigma_n^2} \right)^2 \frac{\tau_i^2}{\sigma_n^2} \to \int_0^\delta \left(\int_x^1 f(y)dy \right)^2 dx,$$

thus for sufficiently large n we get $\sum_{i=1}^{m_n(\delta)} (A_i^n)^2 \tau_i^2 < 2\varepsilon$.

Furthermore $\sum_{k=m_n(\delta)+1}^{n} f\left(\frac{\sigma_k^2}{\sigma_n^2}\right)\frac{\tau_k^2}{\sigma_n^2} \to \int_\delta^1 f(x)dx$, therefore for sufficiently large n we get $\sum_{k=m_n(\delta)+1}^{n} f\left(\frac{\sigma_k^2}{\sigma_n^2}\right)\frac{\tau_k^2}{\sigma_n^2} < 2\eta$ and in consequence

$$A_{m_n(\delta)+1}^n = \sum_{k=m_n(\delta)+1}^{n} f\left(\frac{\sigma_k^2}{\sigma_n^2}\right)\frac{\tau_k^2}{\sigma_n^3} < \frac{2\eta}{\sigma_n}.$$

Thus, for sufficiently large n, we get

$$\begin{aligned}&\sum_{i=1}^{n} (A_i^n)^2 \,\mathbb{E}\,(X_i - \mathbb{E}X_i)^2\, \mathbb{I}\left[A_i^n\, |X_i - \mathbb{E}X_i| \geqslant \varepsilon\right] \\ \leqslant\ & 2\varepsilon + \frac{4\eta^2}{\sigma_n^2}\sum_{i=1}^{n} \mathbb{E}\,(X_i - \mathbb{E}X_i)^2\, \mathbb{I}\left[|X_i - \mathbb{E}X_i| \geqslant \sigma_n\varepsilon/2\eta\right]\end{aligned}$$

and (2.6) follows. □

Another approach to non-identically distributed r.v.'s is to normalize them by their expectations i.e. by considering $X_k' = X_k/\mathbb{E}X_k$. Then $\mathrm{Var}\,(X_k') = \gamma_k^2$, where $\gamma_k = \tau_k/\mathbb{E}X_k$ is the coefficient of variation of X_k. Furthermore we put $S_n' = \sum_{k=1}^n X_k'$ and $\Gamma_n^2 = \mathrm{Var}\,(S_n') = \sum_{k=1}^n \gamma_k^2$. Under this notation we have the following corollary to our main Theorem 2.1.

Corollary 2.6. *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of independent, positive and square-integrable r.v.'s satisfying the following conditions.*

$$\sum_{k=1}^{\infty} \frac{\gamma_k^2}{k^2} < \infty, \tag{2.16}$$

$$\lim_{n\to\infty} \frac{1}{\Gamma_n^2}\sum_{i=1}^{n} \mathbb{E}\left(\frac{X_i}{\mathbb{E}X_i} - 1\right)^2 \mathbb{I}\left[\left|\frac{X_i}{\mathbb{E}X_i} - 1\right| \geqslant \varepsilon\Gamma_n\right] = 0, \text{ for all } \varepsilon > 0, \tag{2.17}$$

$$\Gamma_n \to \infty, \text{ as } n \to \infty. \tag{2.18}$$

Then

$$\left(\prod_{i=1}^{n} \frac{S_i'}{i}\right)^{1/\Gamma_n} \xrightarrow{d} \exp\left(\mathcal{N}(0,2)\right), \text{ as } n \to \infty. \tag{2.19}$$

Proof. We apply Theorem 2.1 to the r.v.'s $X_k' = X_k/\mathbb{E}X_k$ with $a_n = n$. From (2.16) we get (2.2), furthermore by Kronecker's lemma

$$\frac{1}{n^2}\sum_{k=1}^{n} \gamma_k^2 = \left(\frac{\Gamma_n}{n}\right)^2 \to 0. \tag{2.20}$$

In our case (2.3) takes the form

$$\frac{1}{\Gamma_n}\sum_{i=1}^{n}\frac{\Gamma_i^2}{i^2}\to 0. \tag{2.21}$$

By applying the summation by parts formula (or by changing the order of summation) we get

$$\frac{1}{\Gamma_n}\sum_{i=1}^{n}\frac{\Gamma_i^2}{i^2}\leqslant\frac{\text{const}}{\Gamma_n}\sum_{i=1}^{n}\frac{\gamma_i^2}{i}.$$

In view of (2.18) we can apply the Stolz theorem

$$\frac{\sum_{i=1}^{n+1}\frac{\gamma_i^2}{i}-\sum_{i=1}^{n}\frac{\gamma_i^2}{i}}{\Gamma_{n+1}-\Gamma_n}=\frac{\gamma_{n+1}^2}{n+1}\cdot\frac{\Gamma_{n+1}+\Gamma_n}{\Gamma_{n+1}^2-\Gamma_n^2}=\frac{\Gamma_{n+1}+\Gamma_n}{n+1}\to 0$$

by (2.20), thus (2.21) follows. From the Lindeberg's condition (2.17) follows the Feller condition for the sequence $(X_n')_{n\in\mathbb{N}}$ i.e.

$$\max_{1\leqslant i\leqslant n}\frac{\gamma_i^2}{\Gamma_n^2}\to 0.$$

Therefore we can handle with (2.4) as with an integral sum

$$\begin{aligned}\sum_{i=1}^{n}\left(\sum_{k=i}^{n}\frac{1}{\Gamma_n\cdot k}\right)^2\gamma_i^2 &= \sum_{i=1}^{n}\left(\sum_{k=i}^{n}\frac{1}{\frac{k}{n}}\cdot\frac{1}{n}\right)^2\frac{\gamma_i^2}{\Gamma_n^2}\to\\ &\to\int_0^1\left(\int_x^1\frac{1}{y}dy\right)^2dx=2.\end{aligned}$$

The condition (2.5) is obvious, while (2.6) follows from the Lindeberg's condition (2.17) similarly as in the proof of Corollary 2.5. □

3. Rate of convergence

In what follows we shall write $b_n\sim a_n$ if there exist some constants $0<c_1\leqslant c_2<\infty$ such that

$$c_1a_n\leqslant b_n\leqslant c_2a_n.$$

Let us quote Theorem 3.1 of [7] which estimates upper bound of the speed of convergence in Theorem 2.1 in case when $\alpha_{i,n}$ is dependent on n only, and the limiting distribution is $e^{\mathcal{N}(0,1)}$. This result is an extension of the paper [5].

Theorem 3.1. *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of independent random variables, such that $\mathbb{E}X_n = \mu_n$, $\mathbb{E}(X_n-\mu_n)^2 = \tau_n^2$, $n \geqslant 1$. Let $(a_n)_{n\in\mathbb{N}\cup\{0\}}$ be a non-decreasing and divergent to infinity sequence of positive reals (with $a_0 = 0$) such that $\frac{a_{n+1}}{a_n} = O(1)$, as $n\to\infty$. For any $k \geqslant 1$, $\delta > 0$ let us denote*

$$\varphi_k(\delta) = P[X_k - \mu_k \leqslant (\delta-1)(a_k - a_{k-1})].$$

Moreover let $(\gamma_n)_{n\in\mathbb{N}}$ be a positive sequence and $A_k^n = \sum_{i=k}^n \frac{1}{a_i}$. If for some $2 < r \leqslant 3$, $\mathbb{E}|X_n|^r < \infty$, $n \geqslant 1$, and

$$\frac{a_n^r}{L_n+\sigma_n^r}\sum_{j=n}^{\infty}\frac{\mathbb{E}|\overline{X}_j|^r + \sigma_{j+1}^r - \sigma_j^r}{a_j^r} = O(1), \text{ as } n\to\infty,$$

as well as

$$\frac{\mathbb{E}|\overline{X}_{n+1}|^r + \sigma_{n+1}^r - \sigma_n^r}{L_n+\sigma_n^r}\cdot\frac{1}{(a_{n+1}/a_n)^r - 1} = O(1),$$

and $\frac{L_n+\sigma_n^r}{a_n^r}\downarrow 0$, as $n\to\infty$, where

$$L_n = \sum_{j=1}^n \mathbb{E}|\overline{X}_j|^r, \quad \sigma_n^2 = \sum_{j=1}^n \tau_j^2.$$

Then, for any sequence of positive reals δ_n, and $m\in\mathbb{N}$, $1\leqslant m < n$, we have

$$\begin{aligned}\Delta_n := \sup_{x\in\mathbb{R}}\left|P\left[\left(\prod_{k=1}^n \frac{S_k - \mathbb{E}S_k + a_k}{a_k}\right)^{\gamma_n} < x\right] - P\left[\exp(\mathcal{N}(0,1)) < x\right]\right| \\ \leqslant C\left\{\frac{L_m+\sigma_m^r}{a_m^r} + \gamma_n^{\frac{1}{2}}\left(\sum_{k=m+1}^n \frac{\sigma_k^2}{a_k^2}\right)^{\frac{1}{2}}\right. \qquad (3.1)\\ \left. +\psi_{n,m}(\delta_n) + \kappa_m(\delta_n) + \frac{|\max\{\varrho_n,\varrho_n^{-1}\}-1|}{\sqrt{2\pi e}} + \frac{\sum_{j=1}^n (A_j^n)^r\mathbb{E}|\overline{X}_j|^r}{(\sum_{k=1}^n (A_k^n)^2\tau_k^2)^{\frac{r}{2}}}\right\},\end{aligned}$$

where $\varrho_n = \gamma_n\sqrt{\mathrm{Var}(\sum_{k=1}^n A_k^n(\overline{X}_k))}$, $\overline{X}_n = X_n - \mu_n$, $n\geqslant 1$, and

$$\psi_{n,m}(\delta_n) = \frac{\gamma_n^{\frac{1}{2}}}{1\wedge\delta_n}\left(\sum_{k=1}^m \frac{\sigma_k^2}{a_k^2}\right)^{\frac{1}{2}},$$

$$\kappa_m(\delta_n) = 1 - \prod_{k=1}^m (1-\varphi_k(\delta_n)) \leqslant \sum_{k=1}^m \varphi_k(\delta_n).$$

Moreover if $X_n - \mu_n > a_{n-1} - a_n$ almost surely for $n \geqslant 1$, and δ_n be any sequence of positive reals and t be the largest natural number such that $a_t < \delta_n$, then

$$\psi_{n,m}(\delta_n) = \gamma_n^{\frac{1}{2}} \left(\sum_{k=1}^{m\wedge t} \frac{\sigma_k^2}{a_k^2} + \sum_{k=t+1}^{m} \frac{\sigma_k^2}{\delta_n} \mathbb{I}[t < m] \right)^{\frac{1}{2}}$$
$$\kappa_m(\delta_n) = P[X_1 \leqslant \delta_n + \mu_1 - a_1].$$

We shall apply this theorem to estimate the speed of convergence in some examples of sequences of non-identically distributed independent random variables. In Section 4 we shall show that the above conditions may easily be verified in some special situations.

Remark 3.2. This theorem is an estimation of the rate of weak convergence to the limiting distribution $e^{\mathcal{N}(0,1)}$. In the case of convergence to the distribution $e^{\mathcal{N}\left(0,\sigma^2\right)}$, the speed of convergence will be the same. It is easy to see that as we take $\gamma_n' = \frac{\gamma_n}{\sigma}$, then we get

$$\sup_{x\in\mathbb{R}} \left| P\left[\left(\prod_{k=1}^{n} \frac{S_k - ES_k + a_k}{a_k} \right)^{\gamma_n} < x \right] - P[\exp(\mathcal{N}\left(0,\sigma^2\right)) < x] \right| =$$
$$\sup_{x\in\mathbb{R}} \left| P\left[\left(\prod_{k=1}^{n} \frac{S_k - ES_k + a_k}{a_k} \right)^{\gamma_n'} < x^{\frac{1}{\sigma}} \right] - P[\exp(\mathcal{N}\left(0,1\right)) < x^{\frac{1}{\sigma}}] \right|.$$

4. Examples

In the examples given below $C > 0$ denotes a constant which may be different in the consecutive inequalities.

At the beginning we shall modify the Example 4.2 of [6]. We shall use shifted Poisson distribution in order to have (2.1) satisfied with sharp inequality (in [6] weak inequality was used what makes a technical problem in the definition of the product and its logarithm when $C_i + 1 = 0$). Such example was also considered in [13].

Example 4.1. Let $(\xi_k)_{k\in\mathbb{N}}$ be a sequence of independent r.v.'s with Poisson distribution $\xi_k \sim \mathrm{Po}(k)$. We define $X_k = \xi_k + 1$, then $X_k \geqslant 1$ almost surely, furthermore

$$\begin{aligned}
\mathbb{E}X_k &= k+1, \\
\tau_k^2 &= \mathrm{Var}(X_k) = k, \\
\mathbb{E}S_n &= 2 + \ldots + (n+1) = n(n+3)/2, \\
\sigma_n^2 &= \mathrm{Var}(S_n) = 1 + \ldots + n = n(n+1)/2, \\
\mathbb{E}\left(X_k - \mathbb{E}X_k\right)^4 &= \mathbb{E}\left(\xi_k - \mathbb{E}\xi_k\right)^4 = 3k^2 + k.
\end{aligned}$$

We shall illustrate how Corollary 2.4 works with $a_i = \mathbb{E}S_i$ and $\alpha_{i,n} = \frac{1}{\sqrt{\log n}}$. Clearly (2.1) holds by Remark 2.3 and (2.2) is satisfied with $p = 2$. Moreover

$$\sum_{i=1}^{n} \frac{\alpha_{i,n}}{a_i^2}\sigma_i^2 = \frac{1}{\sqrt{\log n}} \sum_{i=1}^{n} \frac{2(i+1)}{i(i+3)^2} \to 0, \text{ as } n \to \infty$$

and we have (2.3). Now let us observe that

$$\begin{aligned} \frac{1}{i+3} - \frac{1}{n+3} &= \int_i^n \frac{dx}{(x+3)^2} \leqslant \sum_{k=i}^{n} \frac{1}{k(k+3)} \\ &\leqslant \frac{1}{i(i+3)} + \int_i^n \frac{dx}{x^2} \leqslant \frac{1}{i^2} + \frac{1}{i} - \frac{1}{n} \end{aligned}$$

therefore

$$\sum_{i=1}^{n} (A_i^n)^2 \tau_i^2 = \frac{1}{\log n} \sum_{i=1}^{n} \left(\sum_{k=i}^{n} \frac{2}{k(k+3)} \right)^2 i \to 4, \text{ as } n \to \infty.$$

It remains to check the Lyapunov condition (2.9) with $\delta = 2$

$$\sum_{i=1}^{n} (A_i^n)^4 \, \mathbb{E}\left(X_i - \mathbb{E}X_i\right)^4 \leqslant \frac{16}{\log^2 n} \sum_{i=1}^{n} \frac{1}{i^4}\left(3i^2 + i\right) \to 0, \text{ as } n \to \infty.$$

Thus we have proved that

$$\left(\prod_{i=1}^{n} \left(\frac{2S_i}{i(i+3)} \right) \right)^{1/\sqrt{\log n}} \xrightarrow{d} \exp\left(\mathcal{N}(0,4)\right), \text{ as } n \to \infty. \tag{4.1}$$

Now, using Theorem 3.1, we shall estimate the rate of convergence in (4.1). According to Remark 3.2 we take $\gamma_n' = \frac{\gamma_n}{2} = \frac{1}{2\sqrt{\log n}}$. For $r = 3$ we have $\mathbb{E}|\overline{X}_k|^3 \sim k^{\frac{3}{2}}$, and in consequence $L_n \sim n^{\frac{5}{2}}$. Let us check the assumptions:

$$\frac{a_{n+1}}{a_n} = \frac{(n+1)(n+4)}{n(n+3)} = O(1),$$

$$\begin{aligned} &\frac{a_n^r}{L_n + \sigma_n^r} \sum_{j=n}^{\infty} \frac{\mathbb{E}|\overline{X}_j|^r + \sigma_{j+1}^r - \sigma_j^r}{a_j^r} \\ \sim &\frac{n^3(n+3)^3}{\sqrt{n}^5 + \sqrt{n(n+1)}^3} \sum_{j=n}^{\infty} \frac{\sqrt{j}^3 + \sqrt{(j+1)(j+2)}^3 - \sqrt{j(j+1)}^3}{(j(j+3))^3} \\ &\sim n^3 \sum_{j=n}^{\infty} \frac{\sqrt{j}^3 + j^2}{j^6} \sim n^3 \sum_{j=n}^{\infty} \frac{1}{j^4} = O(1), \end{aligned}$$

$$\frac{\mathbb{E}|\overline{X}_{n+1}|^r + \sigma_{n+1}^r - \sigma_n^r}{L_n + \sigma_n^r} \cdot \frac{1}{(a_{n+1}/a_n)^r - 1}$$
$$\sim \frac{n^{\frac{3}{2}} + \sqrt{(n+1)(n+2)}^3 - \sqrt{n(n+1)}^3}{n^{\frac{5}{2}} + \sqrt{n(n+1)}^3} \cdot \frac{1}{(1 + \frac{2n+4}{n^2+3n})^3 - 1} = O(1),$$

$$\frac{L_n + \sigma_n^r}{a_n^r} \sim \frac{n^{5/2} + \left(\frac{n(n+1)}{2}\right)^{3/2}}{\frac{n^3(n+3)^3}{8}} \downarrow 0.$$

Thus the assumptions of Theorem 3.1 are satisfied. Now we shall estimate all summands in (3.1):

$$\frac{L_m + \sigma_m^r}{a_m^r} \leqslant \frac{8}{m^3},$$

$$\begin{aligned} \gamma_n^{\frac{1}{2}} \left(\sum_{k=m+1}^{n} \frac{\sigma_k^2}{a_k^2} \right)^{\frac{1}{2}} &\leqslant C \frac{1}{\sqrt[4]{\log n}} \left(\sum_{k=m+1}^{n} \frac{2k(k+1)}{k^2(k+3)^2} \right)^{\frac{1}{2}} \\ &\leqslant C \frac{1}{\sqrt[4]{\log n}} \left(\sum_{k=m+1}^{n} \frac{1}{k^2} \right)^{\frac{1}{2}} \\ &\leqslant C \frac{1}{\sqrt[4]{\log n}\sqrt{m}} \end{aligned}$$

$$\psi_{n,m}(\delta_n) = \frac{1}{\sqrt[4]{\log n}} \cdot \frac{1}{\delta_n \wedge 1} \left(\sum_{k=1}^{m} \frac{\sigma_k^2}{a_k^2} \right)^{\frac{1}{2}} \leqslant C \frac{1}{\sqrt[4]{\log n}\delta_n}, \text{ for } \delta_n < 1,$$

$$\begin{aligned} \kappa_m(\delta_n) &\leqslant \sum_{k=1}^{m} \varphi_k(\delta_n) = \sum_{k=1}^{m} P[X_k - k - 1 \leqslant (\delta_n - 1)(k+1)] \\ &= \sum_{k=1}^{m} P[X_k \leqslant \delta_n(k+1)] = 0, \end{aligned}$$

if $\delta_n(k+1) < 1$, and in consequence $\delta_n < \frac{1}{m+1} \leqslant \frac{1}{k+1}$.

$$\varrho_n = \gamma_n' \sqrt{\operatorname{Var}\left(\sum_{k=1}^n A_k^n \overline{X}_k\right)} = \frac{1}{2\sqrt{\log n}} \left(\sum_{k=1}^n (A_k^n)^2 \mathbb{E}\overline{X}_k^2\right)^{\frac{1}{2}}$$

$$\begin{aligned}
&= \frac{1}{2\sqrt{\log n}} \left(\sum_{k=1}^n \left(\sum_{i=k}^n \frac{2}{i(i+3)}\right)^2 k\right)^{\frac{1}{2}} \\
&= \frac{1}{2\sqrt{\log n}} \left(\sum_{k=1}^n \left(\frac{2}{3}\sum_{i=k}^n \left(\frac{1}{i} - \frac{1}{i+3}\right)\right)^2 k\right)^{\frac{1}{2}} \\
&= \frac{1}{3\sqrt{\log n}} \left(\sum_{k=1}^n k\left(\frac{1}{k} + \frac{1}{k+1} + \frac{1}{k+2} - \frac{1}{n+1} - \frac{1}{n+2} - \frac{1}{n+3}\right)^2\right)^{\frac{1}{2}} \\
&\leqslant 1 + \frac{C}{\sqrt{\log n}},
\end{aligned}$$

if $C > 0$. Otherwise we use the inequality $\frac{1}{1-\frac{C}{\sqrt{\log n}}} \leqslant 1 + \frac{2C}{\sqrt{\log n}}$, for sufficiently large n. In consequence

$$\frac{|\max\{\varrho_n, \varrho_n^{-1}\}|}{\sqrt{2\pi e}} \leqslant \frac{C}{\sqrt{\log n}}$$

$$\begin{aligned}
\frac{\sum_{k=1}^n (A_k^n)^3 \mathbb{E}|\overline{X}_k|^r}{\left(\sum_{k=1}^n (A_k^n)^2 \tau_k^2\right)^{\frac{3}{2}}} &= \frac{\sum_{k=1}^n k\left(\frac{1}{k} + \frac{1}{k+1} + \frac{1}{k+2} - \frac{1}{n+1} - \frac{1}{n+2} - \frac{1}{n+3}\right)^3}{\left(\sum_{k=1}^n k\left(\frac{1}{k} + \frac{1}{k+1} + \frac{1}{k+2} - \frac{1}{n+1} - \frac{1}{n+2} - \frac{1}{n+3}\right)^2\right)^{\frac{3}{2}}} \\
&\leqslant \frac{C}{(\sqrt{\log n})^3}.
\end{aligned}$$

Finally we have

$$\begin{aligned}
\Delta_n &= \sup_{x\in\mathbb{R}} \left| P\left[\left(\prod_{k=1}^n \frac{S_k - \mathbb{E}S_k + a_k}{a_k}\right)^{\gamma_n'} < x\right] - P[\exp(\mathcal{N}(0,1)) < x]\right| \\
&\leqslant C\left(\frac{8}{m^3} + \frac{1}{\sqrt[4]{\log n}\sqrt{m}} + \frac{1}{\sqrt[4]{\log n}\,\delta_n} + \frac{1}{\sqrt{\log n}} + \frac{1}{(\log n)^{\frac{3}{2}}}\right) \\
&\leqslant C\left(\frac{8}{m^3} + \frac{1}{\sqrt[4]{\log n}\,\delta_n}\right),
\end{aligned}$$

and $\delta_n < \frac{1}{m+1}$. Taking $\delta_n = \frac{1}{m+2}$, we have

$$\Delta_n \leqslant C\left(\frac{8}{m^3} + \frac{m+2}{\sqrt[4]{\log n}}\right),$$

minimizing this bound with $m = \sqrt[16]{\log n}$ and $\delta_n = \frac{1}{\sqrt[16]{\log n}}$, we get

$$\Delta_n \leqslant \frac{C}{(\log n)^{3/16}}.$$

To complete the applications of Corollary 2.4 we present an example where different normalization than $a_i = \mathbb{E}S_i$ is used.

Example 4.2. Consider a sequence $(X_k)_{k\in\mathbb{N}}$ of i.i.d. standard exponential r.v.'s i.e. $X_k \sim \mathrm{Exp}(1)$. In this case

$$\mathbb{E}X_k = 1,\ \tau_k^2 = 1,\ \mathbb{E}\left(X_k - \mathbb{E}X_k\right)^4 = 9,\ \mathbb{E}S_n = n,\ \sigma_n^2 = n.$$

We take $a_i = i\sqrt{i}$ and $\alpha_{i,n} = \frac{1}{\sqrt{\log n}}$. Let us observe that $\mathbb{E}X_n + a_{n-1} - a_n = 1 + (n-1)^{3/2} - n^{3/2} \leqslant 0$, thus (2.1) is satisfied. Furthermore (2.2) holds with $p = 2$, also (2.3) is easily verified. Similarly as in the previous example we check that (2.4) takes the form

$$\frac{1}{\log n}\sum_{i=1}^{n}\left(\sum_{k=i}^{n}\frac{1}{k\sqrt{k}}\right)^2 \longrightarrow 4.$$

To verify the Lyapunov's condition (2.9) observe that

$$(A_i^n)^4 = \left(\sum_{k=i}^{n}\frac{1}{\sqrt{\log n}k\sqrt{k}}\right)^4 \leqslant \frac{1}{\log^2 n}\cdot\frac{\mathrm{const}}{i^2}.$$

Thus from Corollary 2.4 we get

$$\left(\prod_{i=1}^{n}\left(\frac{S_i}{i\sqrt{i}} - \frac{1}{\sqrt{i}} + 1\right)\right)^{1/\sqrt{\log n}} \xrightarrow{d} \exp\left(\mathcal{N}(0,4)\right),\ \text{as } n\to\infty. \tag{4.2}$$

It may be compared with the classical result of Arnold, Villaseñor, Rempała and Wesołowski (1.2)

$$\left(\prod_{i=1}^{n}\left(\frac{S_i}{i}\right)\right)^{1/\sqrt{n}} \xrightarrow{d} \exp\left(\mathcal{N}(0,2)\right),\ \text{as } n\to\infty.$$

For another comparison let us also recall Example 3 of [13] which can be also derived from our Corollary 2.5 with the function $f(x) = x^p$ where $p > -3/2$. We have the following extension of the above convergence

$$\left(\prod_{i=1}^{n}\left(\frac{S_i}{i}\right)^{(i/n)^{p+1}}\right)^{1/\sqrt{n}} \xrightarrow{d} \exp\left(\mathcal{N}\left(0,\frac{2}{(p+2)\,(2p+3)}\right)\right),\ \text{as } n\to\infty.$$

Now we shall estimate the rate of convergence in (4.2). We take $\gamma'_n = \gamma_n/2 = \frac{1}{2\sqrt{\log n}}$. For $r = 3$ we have $\mathbb{E}|\overline{X}_k|^3 = \frac{12}{e} - 2$ and obviously $L_n \sim n$. At first let us check the assumptions of Theorem 3.1 :

$$\frac{a_{n+1}}{a_n} = O(1),$$

$$\begin{aligned} &\frac{a_n^r}{L_n + \sigma_n^r} \sum_{j=n}^{\infty} \frac{\mathbb{E}|\overline{X}_j|^r + \sigma_{j+1}^r - \sigma_j^r}{a_j^r} \\ &\sim \frac{n^{\frac{9}{2}}}{n + n^{\frac{3}{2}}} \sum_{j=n}^{\infty} \frac{C + (j+1)^{\frac{3}{2}} - j^{\frac{3}{2}}}{j^{\frac{9}{2}}} \sim \frac{n^{\frac{9}{2}}}{n^{\frac{3}{2}}} \sum_{j=n}^{\infty} \frac{C + \sqrt{j}}{j^{\frac{9}{2}}} = O(1). \end{aligned}$$

It is easy to see that $\left(\frac{a_{n+1}}{a_n}\right)^3 - 1 \sim \frac{1}{n}$, therefore

$$\frac{\mathbb{E}|\overline{X}_{n+1}|^r + \sigma_{n+1}^r - \sigma_n^r}{L_n + \sigma_n^r} \cdot \frac{1}{(a_{n+1}/a_n)^r - 1} = \frac{C + (n+1)^{\frac{3}{2}} - n^{\frac{3}{2}}}{Cn + n^{\frac{3}{2}}} \cdot n = O(1),$$

$$\frac{L_n + \sigma_n^r}{a_n^r} \sim \frac{n + n^{\frac{3}{2}}}{n^{\frac{9}{2}}} \downarrow 0.$$

So the assumptions of Theorem 3.1 are verified. Now we shall estimate all summands in (3.1):

$$\frac{L_m + \sigma_m^r}{a_m^r} \leqslant \frac{C}{m^3},$$

$$\gamma_n^{\frac{1}{2}} \left(\sum_{k=m+1}^{n} \frac{\sigma_k^2}{a_k^2} \right)^{\frac{1}{2}} = \frac{1}{\sqrt[4]{\log n}} \left(\sum_{k=m+1}^{n} \frac{1}{k^2} \right)^{\frac{1}{2}} \leqslant C \frac{1}{\sqrt[4]{\log n}\sqrt{m}},$$

$$\psi_{n,m}(\delta_n) = \frac{1}{\sqrt[4]{\log n}} \cdot \frac{1}{\delta_n \wedge 1} \left(\sum_{k=1}^{m} \frac{1}{k^2} \right)^{\frac{1}{2}} \leqslant C \frac{1}{\sqrt[4]{\log n}\,\delta_n}, \text{ for } \delta_n < 1,$$

$$\begin{aligned} \kappa_m(\delta_n) &\leqslant \sum_{k=1}^{m} \varphi_k(\delta_n) = \sum_{k=1}^{m} P[X_k - 1 \leqslant (\delta_n - 1)(k^{\frac{3}{2}} - (k-1)^{\frac{3}{2}})] \\ &= P[X_1 \leqslant \delta_n] = 1 - e^{-\delta_n} \leqslant \delta_n, \end{aligned}$$

if $\delta_n < \frac{2\sqrt{2}-2}{2\sqrt{2}-1}$, as $1 - e^{-x} \leqslant x$.

$$\varrho_n = \gamma_n' \sqrt{\operatorname{Var}\left(\sum_{k=1}^{n} A_k^n \overline{X}_k\right)} = \frac{1}{2\sqrt{\log n}} \left(\sum_{k=1}^{n} (A_k^n)^2 \mathbb{E}\overline{X}_k^2\right)^{\frac{1}{2}}$$

$$= \frac{1}{2\sqrt{\log n}} \left(\sum_{k=1}^{n} \left(\sum_{i=k}^{n} \frac{1}{i^{\frac{3}{2}}}\right)^2\right)^{\frac{1}{2}} \sim \frac{1}{2\sqrt{\log n}} \left(\sum_{k=1}^{n} \left(\frac{2}{\sqrt{k}}\right)^2\right)^{\frac{1}{2}}$$

$$\sim \frac{1}{2\sqrt{\log n}} \left(\frac{1}{4}\log n + C\right)^{\frac{1}{2}} \sim 1 + \frac{C}{\sqrt{\log n}},$$

if $C > 0$. Otherwise we use the inequality $\frac{1}{1-\frac{C}{\sqrt{\log n}}} \leqslant 1 + \frac{2C}{\sqrt{\log n}}$, for sufficiently large n. In consequence

$$\frac{|\max\{\varrho_n, \varrho_n^{-1}\}|}{\sqrt{2\pi e}} \leqslant \frac{C}{\sqrt{\log n}}$$

$$\begin{aligned} \frac{\sum_{j=1}^{n} (A_j^n)^3 \mathbb{E}|\overline{X}_j|^r}{(\sum_{k=1}^{n} (A_k^n)^2 \tau_k^2)^{\frac{3}{2}}} &= \frac{C\sum_{j=1}^{n} (A_j^n)^3}{(\sum_{k=1}^{n} (A_k^n)^2)^{\frac{3}{2}}} = C \frac{\sum_{j=1}^{n} (\sum_{i=j}^{n} \frac{1}{i\sqrt{i}})^3}{(\sum_{k=1}^{n} (\sum_{i=k}^{n} \frac{1}{i\sqrt{i}})^2)^{\frac{3}{2}}} \\ &\sim \frac{\sum_{j=1}^{n} \frac{1}{j\sqrt{j}}}{(\sum_{k=1}^{n} \frac{1}{k})^{\frac{3}{2}}} \sim \frac{1}{\log n \sqrt{\log n}}. \end{aligned}$$

Finally we have

$$\begin{aligned} \Delta_n &= \sup_{x \in \mathbb{R}} \left| P\left[\left(\prod_{k=1}^{n} \frac{S_k - \mathbb{E}S_k + a_k}{a_k}\right)^{\gamma_n'} < x\right] - P[\exp(\mathcal{N}(0,1)) < x] \right| \\ &\leqslant C\left(\frac{1}{m^3} + \frac{1}{\sqrt[4]{\log n}\sqrt{m}} + \frac{1}{\sqrt[4]{\log n}\delta_n} + \delta_n + \frac{1}{\sqrt{\log n}} + \frac{1}{(\log n)^{\frac{3}{2}}}\right) \\ &\leqslant C\left(\frac{1}{m^3} + \frac{1}{\sqrt[4]{\log n}\delta_n} + \delta_n\right). \end{aligned}$$

Taking $m = \sqrt[12]{\log n}$ we have

$$\Delta_n \leqslant C\left(\frac{1}{\sqrt[4]{\log n}\delta_n} + \delta_n\right)$$

minimizing this bound with $\delta_n = \frac{1}{\sqrt[8]{\log n}}$, finally we get

$$\Delta_n \leqslant C\frac{1}{(\log n)^{1/8}}.$$

In the next example we consider random variables which take also negative values and such that the obtained convergence result cannot be deduced neither form [6] nor [13].

Example 4.3. Let $(X_k)_{k\in\mathbb{N}}$ be a sequence of independent r.v.'s such that

$$P\left(X_n = \pm\frac{n}{2}\right) = \frac{1}{2}$$

for each $n \in \mathbb{N}$. In this case

$$\mathbb{E}X_k = 0,\ \tau_k^2 = \frac{k^2}{4},\ \mathbb{E}X_k^4 = \frac{k^4}{16},\ \mathbb{E}S_n = 0 \text{ and } \sigma_n^2 = \frac{n(n+1)(2n+1)}{24}.$$

We shall apply our Theorem 2.1 with $\alpha_{i,n} = \frac{i}{n\sqrt{n}}$ and $a_i = i^2$. Since $\mathbb{E}X_n + a_{n-1} - a_n = -2n+1$, then (2.1) is satisfied, (2.2) holds with $p = 2$ and (2.3) takes the form

$$\sum_{i=1}^{n} \frac{\alpha_{i,n}}{a_i^2}\sigma_i^2 = \frac{1}{n\sqrt{n}} \sum_{i=1}^{n} \frac{(i+1)(2i+1)}{24i^2} \leqslant \frac{1}{n\sqrt{n}} \sum_{i=1}^{n} \frac{2i \cdot 3i}{24i^2} = \frac{1}{4\sqrt{n}} \to 0.$$

To calculate the limiting variance σ^2 in (2.4) observe that

$$\begin{aligned}\sum_{i=1}^{n} (A_i^n)^2 \tau_i^2 &= \frac{1}{n^3}\sum_{i=1}^{n}\left(\sum_{k=i}^{n}\frac{1}{k}\right)^2 \frac{i^2}{4} = \frac{1}{4}\sum_{i=1}^{n}\left(\sum_{k=i}^{n}\frac{1}{\frac{k}{n}}\cdot\frac{1}{n}\right)^2 \left(\frac{i}{n}\right)^2 \frac{1}{n} \\ &\to \frac{1}{4}\int_0^1 \left(\int_x^1 \frac{1}{y}dy\right)^2 x^2 dx = \frac{1}{4}\int_0^1 x^2 \ln^2 x dx = \frac{1}{54}.\end{aligned}$$

The assumption (2.5) is easily verified and it remains to prove (2.6) by checking the Lyapunov condition with $\delta = 2$. We have

$$\begin{aligned}&\sum_{i=1}^{n} (A_i^n)^4\, \mathbb{E}X_i^4 \\ &= \frac{1}{n^6}\sum_{i=1}^{n}\left(\sum_{k=i}^{n}\frac{1}{k}\right)^4 \frac{i^4}{16} = \frac{1}{16n}\sum_{i=1}^{n}\left(\sum_{k=i}^{n}\frac{1}{\frac{k}{n}}\cdot\frac{1}{n}\right)^4 \left(\frac{i}{n}\right)^4 \frac{1}{n} \to 0\end{aligned}$$

since $\sum_{i=1}^{n}\left(\sum_{k=i}^{n}\frac{1}{\frac{k}{n}}\cdot\frac{1}{n}\right)^4\left(\frac{i}{n}\right)^4\frac{1}{n} \to \int_0^1 x^4 \ln^4 x dx = \frac{24}{3125}$. Finally by Theorem 2.1 we get

$$\left(\prod_{i=1}^{n}\left(\frac{S_i}{i^2}+1\right)^i\right)^{1/n\sqrt{n}} \xrightarrow{d} \exp\left(\mathcal{N}\left(0, \frac{1}{54}\right)\right),\ \text{as } n \to \infty.$$

To present the applicability of our results let us finally consider an example with non-identically uniformly distributed r.v.s.

Example 4.4. Let $(X_k)_{k\in\mathbb{N}}$ be a sequence of independent r.v.'s with uniform distribution on the interval $(0,k\rangle$ i.e. $X_k \sim \mathbb{U}\,(0,k)$. Hence

$$\mathbb{E}X_k = \frac{k}{2},\ \tau_k^2 = \frac{k^2}{12},\ \mathbb{E}\,(X_k - \mathbb{E}X_k)^4 = \frac{k^4}{80},$$
$$\mathbb{E}S_n = \frac{n(n+1)}{4} \text{ and } \sigma_n^2 = \frac{n(n+1)(2n+1)}{72}.$$

At first we take $a_i = \mathbb{E}S_i$ and $\alpha_{i,n} = \frac{1}{\sqrt{n}}$. The conditions (2.1)-(2.3) and (2.5) are easily verified. We check (2.4)

$$\begin{aligned}\sum_{i=1}^n (A_i^n)^2\,\tau_i^2 &= \sum_{i=1}^n \left(\frac{4}{\sqrt{n}}\sum_{k=i}^n \frac{1}{k(k+1)}\right)^2 \frac{i^2}{12}\\ &= \frac{4}{3n}\sum_{i=1}^n \left(\frac{1}{i} - \frac{1}{n+1}\right)^2 i^2 = \frac{4}{3n}\sum_{i=1}^n \left(1 - \frac{2i}{n+1} + \frac{i^2}{(n+1)^2}\right)\\ &= \frac{4}{3n}\cdot\frac{n(2n+1)}{6(n+1)} \to \frac{4}{9}.\end{aligned}$$

Lyapunov's condition reads as follows

$$\begin{aligned}\sum_{i=1}^n (A_i^n)^4\,\mathbb{E}\,(X_i - \mathbb{E}X_i)^4 &= \sum_{i=1}^n \left(\frac{4}{\sqrt{n}}\sum_{k=i}^n \frac{1}{k(k+1)}\right)^4 \frac{i^4}{80}\\ &= \frac{16}{5n^2}\sum_{i=1}^n \left(\frac{1}{i} - \frac{1}{n+1}\right)^4 i^4 \leqslant \frac{16}{5n} \to 0.\end{aligned}$$

Thus we get

$$\left(\prod_{i=1}^n \frac{4S_i}{i(i+1)}\right)^{1/\sqrt{n}} \xrightarrow{d} \exp\left(\mathcal{N}\left(0,\frac{4}{9}\right)\right), \text{ as } n\to\infty. \tag{4.3}$$

Another choice of normalizing constants is $a_i = \mathbb{E}S_i$ and $\alpha_{i,n} = \frac{\tau_i^2\mathbb{E}S_i}{\sigma_i^2\sigma_n}$. This case was described in Corollary 2.5 with $f(x) = 1/x$. The limiting distribution is then $\exp\left(\mathcal{N}\,(0,2)\right)$. This normalization may be simplified according to our Theorem 2.1, we take $a_i = i^2$ and $\alpha_{i,n} = \frac{i}{n\sqrt{n}}$ as in the Example 4.3. We have $\mathbb{E}X_n + a_{n-1} - a_n = -\frac{3}{2}n + 1 < 0$, thus (2.1) holds. We check (2.4)

$$\sum_{i=1}^n (A_i^n)^2\,\tau_i^2 = \frac{1}{n^3}\sum_{i=1}^n \left(\sum_{k=i}^n \frac{1}{k}\right)^2 \frac{i^2}{12} \to \frac{1}{12}\int_0^1 x^2\ln^2 x dx = \frac{1}{162}.$$

The other conditions are verified similarly and we get

$$\left(\prod_{i=1}^{n}\left(\frac{S_i - i(i+1)/4}{i^2}+1\right)^{i}\right)^{1/n\sqrt{n}} \xrightarrow{d} \exp\left(\mathcal{N}\left(0,\frac{1}{162}\right)\right), \text{ as } n\to\infty.$$

Now we shall estimate the rate of convergence in (4.3). We take $\gamma_n' = \frac{3\gamma_n}{2} = \frac{3}{2\sqrt{n}}$. For $r=3$ we have $\mathbb{E}|\overline{X}_k|^3 = \frac{k^3}{32}$ and in consequence $L_n \sim n^4$. At first we check the assumptions of Theorem 3.1:

$$\frac{a_{n+1}}{a_n} = \frac{(n+1)(n+2)}{n(n+1)} = O(1),$$

$$\begin{aligned}
&\frac{a_n^r}{L_n+\sigma_n^r}\sum_{j=n}^{\infty}\frac{\mathbb{E}|\overline{X}_j|^r+\sigma_{j+1}^r-\sigma_j^r}{a_j^r}\\
\sim\ &\frac{n^3(n+1)^3}{n^4+n^{\frac{9}{2}}}\sum_{j=n}^{\infty}\frac{j^3+(j+1)^{\frac{9}{2}}-j^{\frac{9}{2}}}{(j(j+1))^3} \sim \frac{n^6}{n^{\frac{9}{2}}}\sum_{j=n}^{\infty}\frac{j^3+j^{\frac{7}{2}}}{j^6}\\
\sim\ & n^{\frac{3}{2}}\sum_{j=n}^{\infty}\frac{1}{j^{\frac{5}{2}}} = O(1),
\end{aligned}$$

$$\begin{aligned}
&\frac{\mathbb{E}|\overline{X}_{n+1}|^r+\sigma_{n+1}^r-\sigma_n^r}{L_n+\sigma_n^r}\cdot\frac{1}{(a_{n+1}/a_n)^r-1} \sim \frac{n^3+(n+1)^{\frac{9}{2}}-n^{\frac{9}{2}}}{n^4+n^{\frac{9}{2}}}\cdot\\
&\cdot\frac{1}{(1+\frac{2}{n})^3-1} \sim \frac{n^3+n^{\frac{7}{2}}}{n^4+n^{\frac{9}{2}}}\cdot\frac{1}{\frac{1}{n}+\frac{1}{n^2}+\frac{1}{n^3}} = O(1),
\end{aligned}$$

$$\frac{L_n+\sigma_n^r}{a_n^r} \leqslant C\frac{n^4+n^{\frac{9}{2}}}{n^3(n+1)^3} \leqslant C\frac{n^{\frac{9}{2}}}{n^6} \leqslant \frac{C}{n\sqrt{n}} \downarrow 0.$$

The assumptions of Theorem 3.1 are satisfied. Now we shall estimate all summands in (3.1):

$$\frac{L_m+\sigma_m^r}{a_m^r} \leqslant \frac{C}{m\sqrt{m}},$$

$$\begin{aligned}
\gamma_n^{\frac{1}{2}}\left(\sum_{k=m+1}^{n}\frac{\sigma_k^2}{a_k^2}\right)^{\frac{1}{2}} &\leqslant C\frac{1}{\sqrt[4]{n}}\left(\sum_{k=m+1}^{n}\frac{k(k+1)(2k+1)}{k^2(k+1)^2}\right)^{\frac{1}{2}}\\
&\leqslant C\frac{1}{\sqrt[4]{n}}\left(\sum_{k=m+1}^{n}\frac{1}{k}\right)^{\frac{1}{2}} \leqslant C\frac{\sqrt{\log n}}{\sqrt[4]{n}}
\end{aligned}$$

$$\psi_{n,m}(\delta_n) = \frac{1}{\sqrt[4]{n}} \cdot \frac{1}{\delta_n \wedge 1} \left(\sum_{k=1}^{m} \frac{1}{k} \right)^{\frac{1}{2}} \leqslant C \frac{1}{\sqrt[4]{n}\delta_n} \sqrt{\log m}, \text{ for } \delta_n < 1,$$

$$\begin{aligned} \kappa_m(\delta_n) &\leqslant \sum_{k=1}^{m} \varphi_k(\delta_n) = \sum_{k=1}^{m} P\left[X_k - \frac{k}{2} \leqslant (\delta_n - 1)\frac{k}{2} \right] \\ &= \sum_{k=1}^{m} P\left[X_k \leqslant \delta_n \frac{k}{2} \right] = \frac{m\delta_n}{2}, \end{aligned}$$

$$\varrho_n = \gamma_n' \sqrt{\operatorname{Var}(\sum_{k=1}^{n} A_k^n \overline{X}_k)} = \frac{3}{2\sqrt{n}} \left(\sum_{k=1}^{n} (A_k^n)^2 \mathbb{E}\overline{X}_k^2 \right)^{\frac{1}{2}}$$

$$\begin{aligned} &= \frac{3}{2\sqrt{n}} \left(\sum_{k=1}^{n} \left(\sum_{i=k}^{n} \frac{4}{i(i+1)} \right)^2 \frac{k^2}{12} \right)^{\frac{1}{2}} \\ &= \frac{3}{2\sqrt{n}} \left(\sum_{k=1}^{n} \left(4 \left(\frac{1}{k} - \frac{1}{n+1} \right) \right)^2 \frac{k^2}{12} \right)^{\frac{1}{2}} \\ &\leqslant 1 + \frac{C}{\sqrt{n}}, \end{aligned}$$

if $C > 0$. Otherwise we use the inequality $\frac{1}{1-\frac{C}{\sqrt{n}}} \leqslant 1 + \frac{2C}{\sqrt{n}}$, for sufficiently large n and in consequence

$$\frac{|\max\{\varrho_n, \varrho_n^{-1}\}|}{\sqrt{2\pi e}} \leqslant \frac{C}{\sqrt{n}}$$

$$\begin{aligned} \frac{\sum_{j=1}^{n} (A_j^n)^3 \mathbb{E}|\overline{X}_j|^r}{(\sum_{k=1}^{n} (A_k^n)^2 \tau_k^2)^{\frac{3}{2}}} &= \frac{\sum_{j=1}^{n} (4(\frac{1}{j} - \frac{1}{n+1}))^3 \frac{j^3}{32}}{(\sum_{k=1}^{n} (4(\frac{1}{k} - \frac{1}{n+1}))^2 \frac{k^2}{12})^{\frac{3}{2}}} \\ = \quad & \frac{2\sum_{j=1}^{n} (1 - \frac{3j}{n+1} + \frac{3j^2}{(n+1)^2} - \frac{j^3}{(n+1)^3})}{(\sum_{k=1}^{n} (1 - \frac{2k}{n+1} + \frac{k^2}{(n+1)^2}) \frac{1}{3})^{\frac{3}{2}}} \leqslant \frac{C}{\sqrt{n}}. \end{aligned}$$

Finally we have

$$\begin{aligned} \Delta_n &= \sup_{x \in \mathbb{R}} \left| P\left[\left(\prod_{k=1}^{n} \frac{S_k - \mathbb{E}S_k + a_k}{a_k} \right)^{\gamma_n'} < x \right] - P[\exp(\mathcal{N}(0,1)) < x] \right| \\ &\leqslant C \left(\frac{1}{m\sqrt{m}} + \frac{\sqrt{\log n}}{\sqrt[4]{n}} + \frac{\sqrt{\log m}}{\sqrt[4]{n}\delta_n} + \frac{m\delta_n}{2} + \frac{1}{\sqrt{n}} \right) \\ &\leqslant C \left(\frac{1}{m\sqrt{m}} + \frac{\sqrt{\log n}}{\sqrt[4]{n}} + \frac{\sqrt{\log m}}{\sqrt[4]{n}\delta_n} + \frac{m\delta_n}{2} \right). \end{aligned}$$

Minimizing this bound with $m = \sqrt[16]{n}$ and $\delta_n = \frac{2}{m^{\frac{5}{2}}}$, we get

$$\Delta_n \leqslant C \frac{\sqrt{\log n}}{n^{3/32}}.$$

References

1. B. C. Arnold, J. A. Villaseñor, *The asymptotic distribution of sums of records*, Extremes **1:3** (1998), 351–363.
2. P. Billingsley, *Probability and measure 3rd ed.*, Wiley (1995).
3. K. Gonchigdanzan, *A note on the almost sure limit theorem for U-statistic*, Period. Math. Hungar. *50:1-2* (2005), 149–153.
4. K. M. Kosiński, *On the functional limits for sums of a function of partial sums*, Statist. Probab. Lett. **79** (2009), 1522–1527.
5. T. Krajka, Z. Rychlik, *The speed of convergence of random products of sums of independent random variables*, Probab. Math. Statist. **30:1** (2010), 73–85.
6. T. Krajka, Z. Rychlik, *Limit theorems for products of sums of independent random variables*, Probab. Math. Statist. **30:1** (2010), 73–85.
7. T. Krajka, *Rozkłady graniczne iloczynów sum i maksimów zmiennych losowych*, Ph.D Thesis, Marie Curie-Skłodowska University, Lublin (2013).
8. X. W. Lu, Q. C. Qi, *A note on asymptotic distribution of products of sums*, Statist. Probab. Lett. **68** (2004), 407–413.
9. P. Matuła, I. Stępień, *On the application of strong approximation to weak convergence of products of sums for dependent random variables*, Condensed Matter Physics **11** No. 4(56) (2008), 749–754.
10. P. Matuła, I. Stępień, *Weak convergence of products of sums of independent and non–identically distributed random variables*, J. Math. Anal. Appl. **353** (2009), 49–54.
11. P. Matuła, I. Stępień, *A remark on the rate of convergence for products of sums*, Far East J. Math. Sci. **34:3** (2009), 361-368.
12. P. Matuła, I. Stępień, *Weak and almost sure convergence for products of sums of associated random variables*, ISRN Probability and Statistics, Volume 2012, Article ID 107096, 14 pages.
13. P. Matuła, I. Stępień, *Functional limit theorem for products of sums of independent and nonidentically distributed random variables*, J. Math., Volume 2013, Article ID 640137, 6 pages.
14. V. V. Petrov, *Limit theorems of probability theory*, Clarendon Press, Oxford (1995).
15. Q. C. Qi, *Limit distribution for products of sums*, Statist. Probab. Lett. **62** (2003), 93–100.
16. G. Rempała, J. Wesołowski, *Asymptotics for products of sums and U–statistics*, Elect. Comm. in Probab. **7** (2002), 47–54.
17. Z. Tan, *Precise asymptotics for products of sums and U-statistics*, Braz. J. Probab. Stat. **27** (2013), 20–30.
18. D. Ye, Q. Wu, *Almost sure central limit theorem for product of partial sums of strongly mixing random variables*, J. Inequal. Appl. 2011, Article ID 576301, 9 pages.

19. L.-X. Zhang, W. Huang, *A note on the invariance principle of the products of sums of random variables*, Elect. Comm. in Probab. **12** (2007), 51–56.
20. Y. Zhang, X.-Y. Yang, Z.-S. Dong, *An almost sure central limit theorem for products of partial sums under association*, J. Math. Anal. Appl. **355** (2009), 708–716.
21. L. Zhu, *Large deviations for products of sums of random variables*, Statist. Probab. Lett. **89** (2014), 17–22.

Characterization of certain distributions by transformed kth record values

Iwona Malinowska[1]

Abstract

In this paper, we give the characterization of the general class of continuous distributions based on independent transforms kth lower and upper record values. Specific distributions considered as particular cases of the general class of distributions are inverse exponential, inverse Weibull, inverse Pareto, negative exponential, negative Weibull, negative Pareto, negative power, Gumbel, Exponentiated-Weibull, Burr X, lognormal, Chen distribution.

1. Introduction

Let $\{X_n,\, n \geqslant 1\}$ be a sequence of independent identically distributed (i.i.d.) random variables with cumulative distribution function (cdf) $F(x)$ and probability density function (pdf) $f(x)$. The jth order statistic of a sample $(X_1, ..., X_n)$ is denoted by $X_{j:n}$. For a fixed $k \geqslant 1$ we define the sequence $L_k(n)$, $n \geq 1$, of kth lower record times of $\{X_n,\, n \geqslant 1\}$ as follows:

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

$n \geqslant 1$. The sequence $\{Z_n^{(k)}, n \geqslant 1\}$ with $Z_n^{(k)} = X_{k:L_k(n)+k-1}$, $n \geq 1$, is called the sequence of kth lower record values of $\{X_n,\, n \geqslant 1\}$. Note that

[1]Department of Mathematics, Lublin University of Technology, Nadbystrzycka 38a, 20-618 Lublin, Poland, *e-mail: i.malinowska@pollub.pl*

2010 Mathematics Subject Classification. 62F10, 62F15, 62J99.

Keywords: kth record values, characterization, inverse exponential distribution, inverse Weibull distribution, inverse Pareto distribution, negative exponential distribution, negative Weibull distribution, negative Pareto distribution, negative power distribution, Gumbel distribution, Exponentiated-Weibull distribution, lognormal distribution, Burr X distribution.

$Z_1^{(k)} = \max\{X_1, ..., X_k\}$ and $Z_n^{(1)} = X_{L(n)},\ n \geqslant 1$, are lower record values. It is known that

$$f_{Z_1^{(k)},...,Z_n^{(k)}}(z_1, z_2, ...z_n) = k^n\,(F(z_n))^{k-1}\,f(z_n)\prod_{i=1}^{n-1}\frac{f(z_i)}{F(z_i)}, \quad z_1 > ... > z_n, \tag{1.1}$$

(cf. [14]). Hence the pdf of $Z_n^{(k)}$ and $(Z_m^{(k)}, Z_n^{(k)})$, $m < n$, are as follows:

$$f_{Z_n^{(k)}}(x) = \frac{k^n}{(n-1)!}(H(x))^{n-1}(F(x))^{k-1}f(x), \quad n \geqslant 1, \tag{1.2}$$

$$\begin{aligned} f_{Z_m^{(k)},Z_n^{(k)}}(x,y) = &\frac{k^n}{(m-1)!(n-m-1)!}\Big(H(y) - H(x)\Big)^{n-m-1} \\ &\cdot (H(x))^{m-1}h(x)(F(y))^{k-1}f(y),\ x > y,\ n \geqslant 2, \end{aligned} \tag{1.3}$$

respectively, where $H(\cdot) = -\ln(F(\cdot))$ and $h(\cdot) = -H'(\cdot)$.

Now we recall the definition of k th upper record values. With the above notation, for a fixed $k \geqslant 1$ we define the sequence $U_k(n)$, $n \geq 1$, of kth upper record times of $\{X_n,\, n \geqslant 1\}$ as follows:

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

$n \geqslant 1$. The sequence $\{Y_n^{(k)}, n \geqslant 1\}$ with $Y_n^{(k)} = X_{U_k(n):U_k(n)+k-1},\ n \geq 1$, is called the sequence of kth upper record values of $\{X_n,\, n \geqslant 1\}$ (cf. Dziubdziela and Kopociński [8]). Note that $Y_1^{(k)} = \min\{X_1, ..., X_k\}$, and $Y_n^{(1)} = X_{U(n)},\ n \geqslant 1$, with $U(n) = \min\left\{j > U(n-1) : X_j > X_{U(n-1)}\right\}$ are upper record values . It is known that joint pdf of $Y_1^{(k)}, \ldots, Y_n^{(k)}$ are given by

$$\begin{aligned} &f_{Y_1^{(k)},...,Y_n^{(k)}}(x_1, \ldots, x_n) \\ &= \begin{cases} k^n \prod_{i=1}^{n-1}\dfrac{f(x_i)}{1-F(x_i)}(1-F(x_n))^{k-1}f(x_n), & x_1 < \ldots < x_n, \\ 0, & \text{else.} \end{cases} \end{aligned}$$

Hence the pdf of $Y_n^{(k)}$ and $(Y_m^{(k)}, Y_n^{(k)})$, $m < n$, are as follows:

$$f_{Y_n^{(k)}}(x) = \frac{k^n}{(n-1)!}(\overline{H}(x))^{n-1}(1-F(x))^{k-1}f(x), \quad n \geqslant 1,$$

$$\begin{aligned} f_{Y_m^{(k)},Y_n^{(k)}}(x,y) = &\frac{k^n}{(m-1)!(n-m-1)!}\left(\overline{H}(y) - \overline{H}(x)\right)^{n-m-1} \\ &\cdot (\overline{H}(x))^{m-1}\overline{h}(x)(1-F(y))^{k-1}f(y),\ x > y,\ n \geqslant 2, \end{aligned}$$

respectively, where $\overline{H}(\cdot) = -\ln(1 - F(\cdot))$ and $\overline{h}(\cdot) = \overline{H'}(\cdot)$.

Record values and associated statistics are of great importance in several real-life problems involving weather, economic and sport date. The formal study of record values started with Chandler [5] and has now spread in various directions. The properties of record values have been extensively studied in the literature. In particular, the problem of characterizing a distribution in terms of record values is an important problem which has recently attracted the attention of many researchers. Various characterization of distribution based on record values have been discussed by many authors e.g. Ahsanullah [2] and Arnold et al. [3], Balakrishnan [4] and Nevzorov [13], Dembińska and Wesołowski [7], Pawlas and Szynal [14], Malinowska and Szynal [11].

The aim of this article is to provide a characterization of general class of distribution by using the suitable transformations of kth records in a sequence of independent, identically distributed random variables.This paper generalized results obtained by Juhás and Skrivánková [9].

2. Main results

Theorem 2.1. *Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of i.i.d. random variables with absolutely continuous distribution function $F(x)$ on (a,b). Moreover, let $g : (a,b) \to (0,\infty)$ be a differentiable function with $g'(x) < 0$ for all $x \in (a,b)$ and $\lim\limits_{x\to a^+} g(x) = \infty$, $\lim\limits_{x\to b^-} g(x) = 0$. Then the distribution of $X_1, X_2, \ldots$ is of the form*

$$F(x) = e^{-cg(x)},\ c > 0,\ x \in (a,b)$$

if and only if random variables

$$g(Z_n^{(k)}) \quad \text{and } g(Z_{n+1}^{(k)}) - g(Z_n^{(k)}),\ n \geqslant 1$$

are pairwise independent.

Remark

From assumptions of Theorem 2.1 we have the following facts: function g is injection and $\lim\limits_{x\to\infty} g^{-1}(x) = a$, $\lim\limits_{x\to 0^+} g^{-1}(x) = b$.

Proof. Suppose that $F(x) = e^{-cg(x)}$, $x \in (a,b)$, $c > 0$.
Then $f(x) = -cg'(x)e^{-cg(x)}$, $H(x) = cg(x)$ and $h(x) = -cg'(x)$. It is clear from (1.3) that

$$f_{Z_n^{(k)}, Z_{n+1}^{(k)}}(x,y) = \frac{k^{n+1}}{\Gamma(n)} \left[cg(x)\right]^{n-1} c^2 g'(x) g'(y) \cdot e^{-ckg(y)},\ x, y \in (a,b).$$

Consider the transformation

$$\begin{aligned} t &: \begin{pmatrix} Z_n^{(k)} \\ Z_{n+1}^{(k)} \end{pmatrix} \longrightarrow \begin{pmatrix} g(Z_n^{(k)}) \\ g(Z_{n+1}^{(k)}) - g(Z_n^{(k)}) \end{pmatrix} = \begin{pmatrix} U \\ V \end{pmatrix}; \\ \tau &: \begin{pmatrix} U \\ V \end{pmatrix} \longrightarrow \begin{pmatrix} g^{-1}(U) \\ g^{-1}(U+V) \end{pmatrix}. \end{aligned} \tag{2.1}$$

This transformation has the determinate of the form

$$D_\tau = (g^{-1})'(u)(g^{-1})'(u+v). \tag{2.2}$$

Then the pdf of (U, V) is as follows:

$$\begin{aligned} f_{U,V}(u,v) &= \frac{k^{n+1}}{\Gamma(n)} \left[H(g^{-1}(u))\right]^{n-1} h(g^{-1}(u)) \left[F(g^{-1}(u+v))\right] \\ &\qquad h(g^{-1}(u+v))|(g^{-1})'(u)(g^{-1})'(u+v)| \\ &= \frac{(k \cdot c)^{n+1}}{\Gamma(n)} u^{n-1} e^{-ck(u+v)}, \;\; u > 0,\, v > 0,\, c > 0. \end{aligned}$$

In view of (1.2) we have

$$f_{Z_n^{(k)}}(x) = -\frac{k^n}{\Gamma(n)} (cg(x))^{n-1} e^{-ckg(x)} cg'(x), \;\; x \in (a,b); \; c > 0$$

so the pdf of $U = g(Z_n^{(k)})$ is given by formula

$$f_U(u) = \frac{k^n c^n}{\Gamma(n)} u^{n-1} e^{-cku}, \quad u > 0,\, c > 0. \tag{2.3}$$

Integrating the joint pdf $f_{U,V}(u,v)$ according to u we obtain pdf of V

$$f_V(v) = kce^{-ckv}, \;\; v > 0. \tag{2.4}$$

In the view of (2.3) and (2.4), we can notice that $f_{U,V}(u,v) = f_U(u) \cdot f_V(v)$. So U and V are independent random variables. The necessary condition is proved.

In order to prove sufficient condition we assume, that random variable U and V are independent. Consider the transformation (2.1) with determinant (2.2). Then the joint pdf of U and V can be written in general form

$$\begin{aligned} f_{U,V}(u,v) &= \frac{k^{n+1}}{\Gamma(n)} \left[H(g^{-1}(u))\right]^{n-1} h(g^{-1}(u)) \left[F(g^{-1}(u+v))\right] \\ &\qquad \cdot h(g^{-1}(u+v))|(g^{-1})'(u)(g^{-1})'(u+v)|. \end{aligned}$$

We know that $U = g(Z_n^{(k)})$ thus for its density function holds

$$f_U(u) = \frac{k^n}{\Gamma(n)} \left[H(g^{-1}(u))\right]^{n-1} h(g^{-1}(u)) \left[F(g^{-1}(u))\right]^k |(g^{-1})'(u)|, \; u > 0.$$

Because U and V are independent then the density of V is

$$f_V(v) = k \frac{\left[F(g^{-1}(u+v))\right]^{k-1}}{\left[F(g^{-1}(u))\right]^k} f(g^{-1}(u+v)) |(g^{-1})'(u+v)|. \tag{2.5}$$

In order to obtain the $F_V(v^*)$ we integrate equation (2.5)

$$\begin{aligned}
\int_0^{v^*} f_V(v)dv &= k \int_0^{v^*} \frac{\left[F(g^{-1}(u+v))\right]^{k-1}}{\left[F(g^{-1}(u))\right]^k} f(g^{-1}(u+v)) \\
&\qquad |(g^{-1})'(u+v)| dv \\
&= \frac{-k}{\left[F(g^{-1}(u))\right]^k} \int_0^{v^*} \left[F(g^{-1}(u+v))\right]^{k-1} f(g^{-1}(u+v)) \\
&\qquad (g^{-1})'(u+v) dv \\
\begin{pmatrix} \text{substituting} \\ t = g^{-1}(u+v) \end{pmatrix} &= \frac{-k}{\left[F(g^{-1}(u))\right]^k} \int_{g^{-1}(u)}^{g^{-1}(u+v^*)} \left[F(t)\right]^{k-1} f(t) dt \\
&= \frac{1}{\left[F(g^{-1}(u))\right]^k} \left(\left[F(g^{-1}(u))\right]^k - \left[F(g^{-1}(u+v^*))\right]^k \right).
\end{aligned}$$

Consider the limit case where $u \to 0^+$ so $g^{-1}(u) \to b$ and $F(g^{-1}(u)) \to 1$. Then

$$F_V(v^*) = 1 - \left[F(g^{-1}(v^*))\right]^k$$

and it holds

$$\left[F(g^{-1}(v^*))\right]^k \cdot \left[F(g^{-1}(u))\right]^k = \left[F(g^{-1}(u+v^*))\right]^k, \quad v^* > 0, u > 0.$$

Denote $F_1(x) = F(g^{-1}(x)), x > 0$, then

$$\left[F_1(v^*)\right]^k \cdot \left[F_1((u)\right]^k = \left[F_1(u+v^*)\right]^k. \tag{2.6}$$

The equation (2.6) is equivalent to

$$F_1(v^*) \cdot F_1(u) = F_1(u+v^*). \tag{2.7}$$

Nontrivial solution of the equation (2.7) is $F_1(x) = e^{cx}$ (Cauchy functional equation) (cf. [1]) where c is an arbitrary constant and $F(x) = e^{cg(x)}$. Since F is distribution function, we have $F(x) = e^{-cg(x)}$, where $c > 0$. Thus the proof is finished. $\square$

Remark

Interval (a, b) which is mentioned in Theorem 2.1 can have also one of the following forms: (a, ∞), $(-\infty, b)$ or $(-\infty, \infty)$.

Examples

Many of distributions can be characterized by the suitable choice of function g and interval (a, b). Some of them are presented below.

(i) Let NExp$(\lambda,\, \nu)$ denote the negative exponential distribution with

$$F(x) = \exp\left(\lambda(x-\nu)\right), \quad x < \nu;\ \lambda > 0,\ \nu \in \mathbf{R}. \tag{2.8}$$

If we take $g(x) = -\frac{1}{c}\left(\lambda(x-\nu)\right)$, $c > 0$, $x \in (-\infty, \nu)$, $\lambda > 0, \nu \in \mathbb{R}$, then the independence of variables

$$-\frac{1}{c}\left(\lambda(Z_n^{(k)} - \nu)\right), \quad -\frac{\lambda}{c}\left(Z_{n+1}^{(k)} - Z_n^{(k)}\right)$$

characterizes NExp$(\lambda,\, \nu)$ distribution.

(ii) Let IExp(θ) denote the inverse exponential distribution with

$$F(x) = \exp\left(-\theta/x\right), \quad x > 0;\ \theta > 0, \text{ (cf. Klugman et al. [10]).} \tag{2.9}$$

If we take $g(x) = \frac{1}{c}\frac{\theta}{x}$, $c > 0$, $x \in (0, \infty)$, $\theta > 0$, then the independence of variables

$$\frac{\theta}{c}\left(Z_n^{(k)}\right)^{-1}, \quad \frac{\theta}{c}\left(\left(Z_{n+1}^{(k)}\right)^{-1} - \left(Z_n^{(k)}\right)^{-1}\right)$$

characterizes IE(θ) distribution.

(iii) Let NPar$(\theta, \nu\,, \delta)$ denote the negative Pareto distribution with

$$F(x) = \left(\frac{\delta-\nu}{\delta-x}\right)^{\theta}, \quad x < \nu;\ \theta > 0,\ \nu, \delta \in \mathbf{R},\ \nu < \delta. \tag{2.10}$$

If we take $g(x) = -\frac{1}{c}\ln\left(\frac{\delta-\nu}{\delta-x}\right)^{\theta}$, $c > 0$, $x \in (-\infty, \nu)$, $\theta > 0$, $\nu, \delta \in \mathbf{R}$, $\nu < \delta$, then the independence of variables

$$-\frac{\theta}{c}\ln\left(\frac{\delta-\nu}{\delta - Z_n^{(k)}}\right), \quad \frac{\theta}{c}\ln\left(\frac{\delta - Z_{n+1}^{(k)}}{\delta - Z_n^{(k)}}\right)$$

characterizes NPar$(\theta, \nu\,, \delta)$ distribution.

(iv) Let NPow$(\theta, \alpha\,, \beta)$ denote the inverse power distribution with

$$F(x) = \left(\frac{x-\alpha}{\beta-\alpha}\right)^{\theta}, \quad \alpha < x < \beta;\, \theta > 0,\, \alpha, \beta \in \mathbf{R},\, \alpha < \beta. \tag{2.11}$$

If we take $g(x) = -\frac{1}{c}\ln\left(\frac{x-\alpha}{\beta-\alpha}\right)^{\theta}$, $c > 0$, $x \in (\alpha, \beta)$; $\theta > 0$, $\alpha, \beta \in \mathbf{R}$, $\alpha < \beta$, then the independence of

$$-\frac{\theta}{c}\ln\left(\frac{Z_n^{(k)}-\alpha}{\beta-\alpha}\right), \quad -\frac{\theta}{c}\ln\left(\frac{Z_{n+1}^{(k)}-\alpha}{Z_n^{(k)}-\alpha}\right)$$

characterizes NPow$(\theta, \alpha\,, \beta)$ distribution.

(v) Let Gum$(\beta,\, \gamma)$ denote the Gumbel distribution with

$$F(x) = \exp\left[-e^{-\beta(x-\gamma)}\right], \quad x \in \mathbf{R};\, \beta > 0,\, \gamma \in \mathbf{R}. \tag{2.12}$$

If we take $g(x) = \frac{1}{c}\exp[-\beta\,(x-\gamma)]$, $c > 0$, $x \in (-\infty, \infty)$; $\beta > 0$, $\gamma \in \mathbf{R}$, then the independence of variables

$$\frac{1}{c}\exp\left(-\beta(Z_n^{(k)}-\gamma)\right)$$

and

$$\frac{\theta}{c}\left[\exp\left(-\beta(Z_{n+1}^{(k)}-\gamma)\right) - \exp\left(-\beta(Z_n^{(k)}-\gamma)\right)\right]$$

characterizes Gum$(\beta,\, \gamma)$ distribution.

(vi) Let Fre$(\theta,\, \delta,\, \mu)$ denote the Fréchet distribution with

$$F(x) = \exp\left(-\left(\frac{\delta-\mu}{x-\mu}\right)^{\theta}\right),\ x > \mu;\, \theta > 0,\, \mu, \delta \in \mathbf{R},\, \mu < \delta. \tag{2.13}$$

If we take $g(x) = \frac{1}{c}\left(\frac{\delta-\mu}{x-\mu}\right)^{\theta}$, $c > 0$, $x \in (\mu, \infty)$; $\theta > 0$, $\mu, \delta \in \mathbf{R}$, $\mu < \delta$, then the independence of variables

$$\frac{1}{c}\left(\frac{\delta-\mu}{Z_n^{(k)}-\mu}\right)^{\theta}, \quad \frac{1}{c}\left[\left(\frac{\delta-\mu}{Z_{n+1}^{(k)}-\mu}\right)^{\theta} - \left(\frac{\delta-\mu}{Z_n^{(k)}-\mu}\right)^{\theta}\right]$$

characterizes Fre$(\theta,\, \delta,\, \mu)$ distribution.

(vii) Let NWeib(θ, μ, γ) denote the negative Weibull distribution with

$$F(x) = \exp\left(-\left(\frac{\mu - x}{\mu - \gamma}\right)^{\theta}\right), \quad x < \mu;\ \theta > 0,\ \mu, \gamma \in \mathbf{R},\ \mu > \gamma. \tag{2.14}$$

If we take $g(x) = \frac{1}{c}\left(\frac{\mu-x}{\mu-\gamma}\right)^{\theta}$, $c > 0$, $x \in (-\infty, \mu)$; $\theta > 0$, $\mu, \gamma \in \mathbf{R}$, $mu > \gamma$, then the independence of variables

$$\frac{1}{c}\left(\frac{\mu - Z_n^{(k)}}{\mu - \gamma}\right)^{\theta}, \quad \frac{1}{c}\left[\left(\frac{\mu - Z_{n+1}^{(k)}}{\mu - \gamma}\right)^{\theta} - \left(\frac{\mu - Z_n^{(k)}}{\mu - \gamma}\right)^{\theta}\right]$$

characterizes NWeib(θ, μ, γ) distribution.

(viii) Let IWeib(θ, τ) denote the Inverse Weibull distribution with

$$F(x) = \exp\left(-\left(\frac{\theta}{x}\right)^{\tau}\right), \quad x > 0;\ \theta > 0,\ \tau > 0, \tag{2.15}$$

(cf. Klugman et al. [10]). If we take $g(x) = \frac{1}{c}\left(\frac{\theta}{x}\right)^{\tau}$, $c > 0$, $x \in (0, \infty)$; $\theta > 0$, $\tau > 0$, then the independence of variables

$$\frac{1}{c}\left(\frac{\theta}{Z_n^{(k)}}\right)^{\tau}, \quad \frac{\theta^{\tau}}{c}\left[\left(Z_{n+1}^{(k)}\right)^{-\tau} - \left(Z_n^{(k)}\right)^{-\tau}\right]$$

characterizes IWeib(θ, τ) distribution.

(ix) Let ExpWeib(θ, α) denote the Exponentiated-Weibull distribution with

$$F(x) = \left(1 - \exp\left(-x^{\alpha}\right)\right)^{\theta}, \quad x > 0;\ \theta > 0,\ \alpha > 0, \tag{2.16}$$

(cf. Manal and Fathy [12]). If we take $g(x) = -\frac{1}{c}\ln\left(1 - \exp(-x^{\alpha})\right)^{\theta}$, $c > 0$, $x \in (0, \infty)$; $\theta > 0$, $\alpha > 0$, then the independence of variables

$$-\frac{\theta}{c}\left(\ln\left(1 - \exp\left(-(Z_n^{(k)})^{\alpha}\right)\right)\right), \quad \frac{-\theta}{c}\left[\ln\frac{1 - \exp\left(-\left(Z_{n+1}^{(k)}\right)^{-\alpha}\right)}{1 - \exp\left(-\left(Z_n^{(k)}\right)^{-\alpha}\right)}\right]$$

characterizes ExpWeib(θ, α) distribution.

(x) Let BuX(θ) denote the Burr X distribution with

$$F(x) = \left(1 - \exp\left(-x^2\right)\right)^{\theta}, \quad x \in \mathbf{R};\ \theta > 0. \tag{2.17}$$

If we take $g(x) = -\frac{\theta}{c}\ln\left(1-\exp(-x^2)\right)^{\theta}$, $c>0$, $x\in(-\infty,\infty)$; $\theta>0$, then the independence of variables

$$-\frac{\theta}{c}\left(\ln\left(1-\exp\left(-(Z_n^{(k)})^2\right)\right)\right), \quad \frac{-\theta}{c}\left[\ln\frac{1-\exp\left(-\left(Z_{n+1}^{(k)}\right)^2\right)}{1-\exp\left(-\left(Z_n^{(k)}\right)^2\right)}\right]$$

characterizes BuX(θ) distribution.

(xi) Let LogNor$(\mu,\,\gamma)$ denote the lognormal distribution with

$$F(x) = \Phi\left(\frac{\ln x-\mu}{\sigma}\right), \quad x\in(0,\infty);\ \mu\in\mathbf{R},\ \sigma>0. \tag{2.18}$$

If we take $g(x) = -\frac{1}{c}\ln\Phi\left(\frac{\ln x-\mu}{\sigma}\right)$, $c>0$, $x\in(0,\infty)$; $\mu\in\mathbf{R}$, $\sigma>0$ then the independence of variables

$$-\frac{1}{c}\ln\Phi\left(\frac{\ln Z_n^{(k)}-\mu}{\sigma}\right), \quad \frac{1}{c}\ln\left[\frac{\Phi\left(\frac{\ln Z_n^{(k)}-\mu}{\sigma}\right)}{\Phi\left(\frac{\ln Z_{n+1}^{(k)}-\mu}{\sigma}\right)}\right]$$

characterizes LogNor$(\mu,\,\delta)$ distribution.

(xii) Let Chen$(\lambda,\,\beta)$ denote the Chen distribution with

$$F(x) = 1-\exp\left(\lambda(1-e^{x^\beta})\right), \quad x>0;\ \lambda>0,\ \beta>0,\ (\text{cf. Chen [6]}). \tag{2.19}$$

If we take $g(x) = -\frac{1}{c}\ln\left(1-\exp\left(\lambda(1-e^{x^\beta})\right)\right)$, $c>0$, $x\in(0,\infty)$; $\lambda>0$, $\beta>0$, then the independence of

$$-\frac{1}{c}\ln\left(1-\exp\left(\lambda(1-e^{(Z_n^{(k)})^\beta})\right)\right), \ \frac{1}{c}\ln\left[\frac{1-\exp\left(\lambda(1-e^{(Z_n^{(k)})^\beta})\right)}{1-\exp\left(\lambda(1-e^{(Z_{n+1}^{(k)})^\beta})\right)}\right]$$

characterizes Chen$(\lambda,\,\beta)$ distribution.
In case $\beta=1$ we give characterization for Gompertz distribution.

Theorem 2.2. *Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of i.i.d. random variable with absolutely continuous distribution function $F(x)$ on (a,b). Moreover, let*

$g\colon (a,b)\to(0,\infty)$ be a differentiable function with $g'(x)<0$ for all $x\in(a,b)$ and $\lim\limits_{x\to a^+} g(x)=\infty$, $\lim\limits_{x\to b^-} g(x)=0$. Then the distribution of $X_1, X_2, \dots$ is of the form

$$F(x)=e^{-cg(x)},\ c>0,\ x\in(a,b)$$

if and only if random variables

$$g(Z_1^{(k)}),\ g(Z_2^{(k)})-g(Z_1^{(k)}),\ \dots.,g(Z_n^{(k)})-g(Z_{n-1}^{(k)}),\ n\geqslant 2$$

are independent.

Proof. Suppose that $F(x)=e^{-cg(x)}$, $x\in(a,b)$, $c>0$, then the joint pdf of $Z_1^{(k)}$, ...$Z_n^{(k)}$ given by (1.1) takes the form

$$f_{Z_1^{(k)},\dots,Z_n^{(k)}}(z_1,z_2,\dots z_n)=k^n(-c)^n e^{-kcg(x)}\prod_{i=1}^{n} g'(z_i),\quad z_1>\dots>z_n.$$

Now we consider the transformation

$$\begin{aligned} t:&\begin{pmatrix} Z_1^{(k)}\\ Z_2^{(k)}\\ \vdots\\ Z_n^{(k)}\end{pmatrix}\longrightarrow\begin{pmatrix} g(Z_1^{(k)})\\ g(Z_2^{(k)})-g(Z_1^{(k)})\\ \vdots\\ g(Z_n^{(k)})-g(Z_{n-1}^{(k)})\end{pmatrix}=\begin{pmatrix} U_1\\ U_2\\ \vdots\\ U_n\end{pmatrix};\\ \tau:&\begin{pmatrix} U_1\\ U_2\\ \vdots\\ U_n\end{pmatrix}\longrightarrow\begin{pmatrix} g^{-1}(U_1)\\ g^{-1}(U_1+U_2)\\ \vdots\\ g^{-1}(U_1+\dots+U_n)\end{pmatrix}.\end{aligned}\tag{2.20}$$

This transformation has the determinate of the form

$$D_\tau=(g^{-1})'(u_1)(g^{-1})'(u_1+u_2)\dots(g^{-1})'(u_1+\dots+u_m).$$

Since the joint pdf of $U_1, \dots..U_n$ is given by

$$f_{U_1,U_2,\dots.,U_n}(u_1,u_2,\dots,u_n)=k^n(c)^n e^{-ck(u_1+\dots+u_2)},\ \ u_1>0,\ \dots u_n>0,\ c>0.$$

and the marginal probability density functions are as follow

$$f_{U_1}(u_1)=kce^{-cku_1},\ f_{U_2}(u_2)=kce^{-cku_2},\ \dots,\ f_{U_n}(u_n)=kce^{-cku_n}.$$

Thus $U_1,\dots,U_n$ are independent random variables. The necessary condition is proved.

Assume now that variable $U_1, ...U_n$ are independent. Consider the transformation given by (2.20). Then we get

$$f_{U_n}(u_n) = -\frac{k[F(g^{-1}(u_1+u_2+...+u_n))]^{k-1}}{[F(g^{-1}(u_1+u_2+...+u_{n-1}))]^k} \qquad (2.21)$$
$$f(g^{-1}(u_1+u_2+...+u_n))(g^{-1})'(u_1+u_2+...+u_n).$$

By integration of (2.21) we obtain that

$$F_{U_n}(u_{n^*}) = -\int_0^{u_{n^*}} \frac{k[F(g^{-1}(u_1+u_2+...+u_n))]^{k-1}}{[F(g^{-1}(u_1+u_2+...+u_{n-1}))]^k} f(g^{-1}(u_1+...+u_n))$$
$$\cdot (g^{-1})'(u_1+u_2+...+u_n)du_n$$
$$= \frac{[F(g^{-1}(u_1+u_2+...+u_{n-1}))]^k - [F(g^{-1}(u_1+u_2+...+u_{n-1}+u_{n^*}))]^k}{[F(g^{-1}(u_1+u_2+...+u_{n-1}))]^k}.$$

Limit cases $u_1 \to 0^+$, $u_2 \to 0^+$, $...u_{n-1} \to 0^+$, lead to the same functional equation as in proof of the Theorem 2.1. So cumulative distribution function has the form $F(x) = e^{-cg(x)}$, $c > 0$. □

Similar characterizations can be given in terms of the kth upper record values.

Theorem 2.3. *Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of i.i.d. random variable with absolutely continuous distribution function $F(x)$ on (a,b). Moreover, let $g:(a,b) \to (0,\infty)$ be a differentiable function with $g'(x) > 0$ for all $x \in (a,b)$ and $\lim_{x\to a^+} g(x) = 0$, $\lim_{x \to b^-} g(x) = \infty$. Then the distribution of $X_1, X_2, ...$ is of the form*

$$F(x) = 1 - e^{-cg(x)},\ c > 0,\ x \in (a,b)$$

if and only if random variables

$$g(Y_n^{(k)}) \ \ \textit{and}\ g(Y_{n+1}^{(k)}) - g(Y_n^{(k)}),\ n \geqslant 1$$

are pairwise independent.

The use of Theorem 2.3 is illustrated by the following examples.

Examples

(i) Let $g(x) = -\frac{1}{c}\ln\left[1 - e^{\lambda(x-\nu)}\right]$, $c > 0$, $x \in (-\infty, \nu)$, $\lambda > 0, \nu \in \mathbb{R}$. The independence of variables

$$-\frac{1}{c}\ln\left[1 - \exp\left(\lambda(Y_n^{(k)} - \nu)\right)\right],\quad \frac{1}{c}\ln\left[\frac{1-\exp\left(\lambda(Y_n^{(k)}-\nu)\right)}{1-\exp\left(\lambda(Y_{n+1}^{(k)}-\nu)\right)}\right]$$

characterizes NExp(λ, ν) distribution given by (2.8).

(ii) Let $g(x) = -\frac{1}{c}\ln\left[1 - e^{-\left(\frac{\theta}{x}\right)}\right]$, $c > 0$, $x \in (0, \infty)$, $\theta > 0$. The independence of variables

$$-\frac{1}{c}\ln\left[1 - \exp\left(-\frac{\theta}{Y_n^{(k)}}\right)\right], \quad \frac{1}{c}\ln\left[\frac{1 - \exp\left(-\frac{\theta}{Y_n^{(k)}}\right)}{1 - \exp\left(-\frac{\theta}{Y_{n+1}^{(k)}}\right)}\right]$$

then characterizes IExp(θ) distribution given by (2.9).

(iii) Let $g(x) = -\frac{1}{c}\ln\left[1 - \left(\frac{x-\alpha}{\beta-\alpha}\right)^{\theta}\right]$, $c > 0$, $x \in (\alpha, \beta), \theta > 0$, $\alpha, \beta \in \mathbb{R}$, $\alpha < \beta$. The independence of variables

$$-\frac{1}{c}\ln\left[1 - \left(\frac{Y_n^{(k)} - \alpha}{\beta - \alpha}\right)^{\theta}\right], \quad \frac{1}{c}\ln\left[\frac{(\beta-\alpha)^{\theta} - \left(Y_n^{(k)} - \alpha\right)^{\theta}}{(\beta-\alpha)^{\theta} - \left(Y_{n+1}^{(k)} - \alpha\right)^{\theta}}\right]$$

characterizes NPow(θ, α, β) distribution given by (2.11).

(iv) Let $g(x) = -\frac{1}{c}\ln\left[1 - \left(\frac{\delta-\nu}{\delta-x}\right)^{\theta}\right]$, $c > 0$, $x \in (-\infty, \nu)$, $\mathbb{R}$, $\nu < \delta$, $\theta > 0$. The independence of variables

$$-\frac{1}{c}\ln\left[1 - \left(\frac{\delta - \nu}{\delta - Y_n^{(k)}}\right)^{\theta}\right], \quad \frac{1}{c}\ln\left[\frac{1 - \left(\frac{\delta-\nu}{\delta-Y_n^{(k)}}\right)^{\theta}}{1 - \left(\frac{\delta-\nu}{\delta-Y_{n+1}^{(k)}}\right)^{\theta}}\right].$$

characterizes NPar(θ, ν, δ) distribution given by (2.10).

(v) Let $g(x) = -\frac{1}{c}\ln\left[1 - e^{-e^{-\beta(x-\gamma)}}\right]$, $c > 0$, $x \in \mathbf{R}$, $\beta > 0$, $\gamma \in \mathbb{R}$. The independence of variables

$$-\frac{1}{c}\ln\left[1 - \exp\left(-e^{-\beta(Y_n^{(k)}-\gamma)}\right)\right], \quad \frac{1}{c}\ln\left[\frac{1 - \exp\left(-e^{-\beta(Y_n^{(k)}-\gamma)}\right)}{1 - \exp\left(-e^{-\beta(Y_{n+1}^{(k)}-\gamma)}\right)}\right]$$

characterizes Gum(β, γ) distribution given by (2.12).

(vi) Let $g(x) = -\frac{1}{c}\ln\left[1-\exp\left(-\left(\frac{\delta-\mu}{x-\mu}\right)^{\theta}\right)\right]$, $c>0$, $x\in(\mu,\infty)$; $\theta>0$, $\mu,\delta\in\mathbf{R}$, $\mu<\delta$. The independence of variables

$$-\frac{1}{c}\ln\left[1-\exp\left(-\left(\frac{\delta-\mu}{Y_n^{(k)}-\mu}\right)^{\theta}\right)\right],\ \frac{1}{c}\ln\left[\frac{[1-\exp\left(-\left(\frac{\delta-\mu}{Y_n^{(k)}-\mu}\right)^{\theta}\right)}{[1-\exp\left(-\left(\frac{\delta-\mu}{Y_{n+1}^{(k)}-\mu}\right)^{\theta}\right)}\right]$$

characterizes Fréchet distribution given by (2.13).

(vii) Let $g(x) = -\frac{1}{c}\ln\left[1-\exp\left(-\left(\frac{\mu-x}{\mu-\gamma}\right)^{\theta}\right)\right]$, $c>0$, $x\in(-\infty,\mu)$; $\theta>0$, $\mu,\gamma\in\mathbf{R}$, $\mu>\gamma$. The independence of variables

$$-\frac{1}{c}\ln\left[1-\exp\left(-\left(\frac{\mu-Y_n^{(k)}}{\mu-\gamma}\right)^{\theta}\right)\right],\ \frac{1}{c}\ln\left[\frac{[1-\exp\left(-\left(\frac{\mu-Y_n^{(k)}}{\mu-\gamma}\right)^{\theta}\right)}{[1-\exp\left(-\left(\frac{\mu-Y_{n+1}^{(k)}}{\mu-\gamma}\right)^{\theta}\right)}\right]$$

characterizes NWeib(θ, μ, γ) distribution given by (2.14).

(viii) Let $g(x) = -\frac{1}{c}\ln\left[1-\exp\left(-\left(\frac{\theta}{x}\right)^{\tau}\right)\right]$, $c>0$, $x\in(0,\infty)$; $\theta>0$, $\tau>0$. The independence of variables

$$-\frac{1}{c}\ln\left[1-\exp\left(-\left(\frac{\theta}{Y_n^{(k)}}\right)^{\tau}\right)\right],\ \frac{1}{c}\ln\left[\frac{1-\exp\left(-\left(\frac{\theta}{Y_n^{(k)}}\right)^{\tau}\right)}{1-\exp\left(-\left(\frac{\theta}{Y_{n+1}^{(k)}}\right)^{\tau}\right)}\right]$$

characterizes IWeib(θ, τ) distribution given by (2.15).

(ix) Let $g(x) = -\frac{1}{c}\ln\left[1-(1-\exp(-x^{\alpha}))^{\theta}\right]$, $c>0$, $x\in(0,\infty)$; $\theta>0$, $\alpha>0$. The independence of variables

$$-\frac{1}{c}\ln\left[1-\left(1-\exp\left(-\left(Y_n^{(k)}\right)^{\alpha}\right)\right)^{\theta}\right],$$

and

$$\frac{1}{c}\ln\left[\frac{1-\left(1-\exp\left(-\left(Y_n^{(k)}\right)^{\alpha}\right)\right)^{\theta}}{1-\left(1-\exp\left(-\left(Y_{n+1}^{(k)}\right)^{\alpha}\right)\right)^{\theta}}\right]$$

characterizes ExpWeib(θ, α) distribution given by (2.16).

(x) Let $g(x) = -\frac{1}{c}\ln\left[1-\left(1-\exp\left(-x^2\right)\right)^{\theta}\right]$, $c > 0$, $x \in (0,\infty)$; $\theta > 0$. The independence of variables

$$-\frac{1}{c}\ln\left[1-\left(1-\exp\left(-\left(Y_n^{(k)}\right)^2\right)\right)^{\theta}\right],$$

and

$$\frac{1}{c}\ln\left[\frac{1-\left(1-\exp\left(-\left(Y_n^{(k)}\right)^2\right)\right)^{\theta}}{1-\left(1-\exp\left(-\left(Y_{n+1}^{(k)}\right)^2\right)\right)^{\theta}}\right]$$

characterizes BuX(θ) distribution given by (2.17).

(xi) Let $g(x) = -\frac{1}{c}\ln\left[1-\Phi\left(\frac{\ln x-\mu}{\sigma}\right)\right]$, $c > 0$, $x \in (0,\infty)$; $\mu \in R$, $\sigma > 0$. The independence of variables

$$-\frac{1}{c}\ln\left[1-\Phi\left(\frac{\ln Y_n^{(k)}-\mu}{\sigma}\right)\right], \quad \frac{1}{c}\ln\left[\frac{1-\Phi\left(\frac{\ln Y_n^{(k)}-\mu}{\sigma}\right)}{1-\Phi\left(\frac{\ln Y_{n+1}^{(k)}-\mu}{\sigma}\right)}\right]$$

characterises LogNor(μ, δ) distribution given by (2.18).

(xii) Let $g(x) = -\frac{\lambda}{c}\left[1-e^{x^{\beta}}\right]$, $c > 0$, $x \in (0,\infty)$; $\lambda > 0$, $\beta > 0$. The independence of variables

$$-\frac{\lambda}{c}\left[1-e^{Y_n^{(k)\beta}}\right], \quad \frac{\lambda}{c}\left[\exp(Y_{n+1}^{(k)\ \beta}) - \exp(Y_n^{(k)\beta})\right]$$

characterizes Chen(λ, β) distribution given by (2.19).

Theorem 2.4. *Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of i.i.d. random variable with absolutely continuous distribution function $F(x)$ on (a,b). Moreover, let $g:(a,b)\to(0,\infty)$ be a differentiable function with $g'(x) > 0$ for all $x \in (a,b)$ and $\lim\limits_{x\to a^+} g(x) = 0$, $\lim\limits_{x\to b^-} g(x) = \infty$. Then the distribution of $X_1, X_2, \ldots$ is of the form*

$$F(x) = 1 - e^{-cg(x)},\ c > 0,\ x \in (a,b)$$

if and only if random variables

$$g(Y_1^{(k)}),\ g(Y_2^{(k)}) - g(Y_1^{(k)}), \ldots,\ g(Y_n^{(k)}) - g(Y_{n-1}^{(k)}),\ n \geqslant 2$$

are independent.

Remark
Above Theorems are generalizations of the results given in [9].

Acknowledgements. The author is grateful to referee for useful comments which improved the presentation of the paper.

References

1. J. Aczel, *Lectures on Functional EQuations and Their Applications.* Academic Press, New York, USA, 1966.
2. M. Ahsanullah, *Record Values - Theory and Applications.* University Press of America, Lanham, Maryland, USA, 2004.
3. B.C. Arnold, H.N. Balakrishnan, A. Nagaraja, A *First Course in order Statistics.* John Wiley & Sons, New York, 1992.
4. N. Balakrishnan, Record values from Rayleigh and Weibull distribution and associated inference. *Statistics and Probability Letter.* 866:1471-1482 (1993).
5. K.N. Chandler,The distribution and frequency of record values. *Journal of royal Statistical Society, B.* **14**:220-228 (1952).
6. Z. Chen, A new two-parameteer lifetime distribution with bathtub space or incrreaisng failur rate function. *National Institute of Standards and technology, J.Res., Special publications.* **49**:155-161, (2000).
7. A. Dembińska and J. Wesołowski, *Linearity of regresion for non-adjacent record values.* Journal of Statistical Planning and Inference. **90**, 195-205, (2000).
8. W. Dziubdziela and B. Kopociński, Limiting properties of the k-th record values. *Appl. Math.* 15:187-190, (1976).
9. M. Juhás and V. Skřivánková *Characterization of general classes of distributions based on independent property of transformed record values.* Applied Mathematics and Computation **226**, 44-50, (2014).
10. S.A. Klugman, H.H. Panjer, G.E. Willmot, *Loss Models From Date to Decisions.* John Wiley & Sons, INC., New York, USA, 1998.
11. I. Malinowska and D. Szynal, *On characterization of certain distributions of kth lower (upper) record values.* Applied Mathematics and Computation **202**, 338-347, (2008).
12. M.N. Manal and H.F. Fathy, *On the Exponentiated Weibull Distribution.* Commun. Statist. Theory Meth. **32(7)**,1317-1336, (2003).
13. V.B. Nevzorov, *Records.* Mathematical Theory, American Mathematical Society, Providence, 2001.
14. P. Pawlas and D. Szynal, *Relations for single and product moments of kth record values from exponential and Gumbel distributions.* J. Appl. Statist. Sci. **7(1)**,53-62, (1998).

The Schreier-Sims algorithm and random permutations

Ernest Nieznaj[1]

Abstract

We show in a collection of examples how to generate a random element of a subgroup of the group of permutations. We apply to this the Schreier-Sims algorithm that is based on the Otto Schreier theorem.

1. Introduction

Suppose that G is a subgroup of S_n generated by a set of permutations W, i.e. $G = \langle W \rangle$ where $W \subset S_n$. The goal of this paper is to show how to choose at random an element of G. In two cases this problem is quite easy to solve and one uses only general facts about permutations, see Section 2. In the first case if W consists of all transpositions then $\langle W \rangle = S_n$ and Lemma 2.4 holds. In the second case if W is the set of all 3-cycles then $\langle W \rangle = A_n$, see Lemma 2.7.

In general however we have to use the Otto Schreier theorem, see Theorem 4.1 in Section 4. The main result of this article is Theorem 4.4. In Section 5 we study in detail several examples. Section 3 is a short introduction to Section 4.

2. Preliminaries

We give a brief summary of the group of permutations. Let X be a finite set. By $Sym(X)$ we denote a group of all permutations of X. In particular, when $X = \{1, 2, \ldots, n\}$ we write S_n. It is clear that $|S_n| = n!$ for $n \geqslant 1$. The identity of S_n we denote by I, i.e. $I(i) = i$ for $i = 1, \ldots, n$. The composition of $\alpha, \beta \in S_n$ we define by

$$(\alpha \circ \beta)(i) = (\alpha\beta)(i) = \beta(\alpha(i)), \quad i = 1, \ldots, n.$$

[1]Department of Mathematics, Technical University of Lublin, Nadbystrzycka 38A, 20-618 Lublin, Poland, *e-mail: e.nieznaj@pollub.pl*

2010 Mathematics Subject Classification. Primary 20B30, 20B35; Secondary 20B40.
Keywords: The Schreier-Sims algorithm, permutation groups, strong generators.

This left-to-right definition of composition of functions is more convenient for permutations than right-to-left definition in calculus. Hence for

$$\alpha = \begin{pmatrix} 1 & 2 & \dots & n \\ \alpha(1) & \alpha(2) & \dots & \alpha(n) \end{pmatrix}, \quad \beta = \begin{pmatrix} 1 & 2 & \dots & n \\ \beta(1) & \beta(2) & \dots & \beta(n) \end{pmatrix},$$

we have

$$\alpha\beta = \begin{pmatrix} 1 & 2 & \dots & n \\ \beta(\alpha(1)) & \beta(\alpha(2)) & \dots & \beta(\alpha(n)) \end{pmatrix}.$$

The inverse to α is gven by

$$\alpha^{-1} = \begin{pmatrix} \alpha(1) & \alpha(2) & \dots & \alpha(n) \\ 1 & 2 & \dots & n \end{pmatrix}.$$

Clearly $(\alpha\beta)^{-1} = \beta^{-1}\alpha^{-1}$ and by induction $(\alpha_1 \dots \alpha_n)^{-1} = \alpha_n^{-1} \dots \alpha_1^{-1}$, $n \geqslant 3$.

Definition 2.1. Let $X = \{1, \dots, n\}$ and $\alpha \in S_n$. We define

$$Fix(\alpha) := \{i \in X : \alpha(i) = i\}, \quad Act(\alpha) := \{i \in X : \alpha(i) \neq i\}.$$

Clearly $Fix(\alpha)$ is a set of fixed points of α. $Act(\alpha)$ is also called the support of α and denoted by $supp(\alpha)$.

Observe that $|Fix(\alpha)| \leqslant n - 2$ or $|Fix(\alpha)| = n$. Since $|Act(\alpha)| = n - |Fix(\alpha)|$ then $|Act(\alpha)| = 0$ or $2 \leqslant |Act(\alpha)| \leqslant n$.

Definition 2.2. Let $A = \{i_1, \dots, i_k\}$ be a subset of X and $\alpha \in S_n$. If $\alpha(i_1) = i_2$, $\alpha(i_2) = i_3, \dots, \alpha(i_k) = i_1$ and $\alpha(i) = i$ for $i \in X \backslash A$ then α is called a k-cycle. A 2-cycle is usually referred to as a transposition.

Every permutation may by written as a product of disjoint cycles, see e.g. [1], [4], [6]. In the example below we show how to do it. Recall that $\alpha, \beta \in S_n$ are disjoint or independent if $Act(\alpha) \cap Act(\beta) = \emptyset$.

Example 2.3. Let us consider the permutation

$$\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ 3 & 2 & 4 & 8 & 12 & 11 & 7 & 1 & 10 & 5 & 6 & 9 \end{pmatrix}$$

We have $\alpha(1) = 3$, $\alpha(3) = 4$, $\alpha(4) = 8$ and $\alpha(8) = 1$. We write this cycle in the form $(1, 3, 4, 8)$. Since $\alpha(2) = 2$ so 2 is a fixed point of α. Next we have $\alpha(5) = 12$, $\alpha(12) = 9$, $\alpha(9) = 10$ and $\alpha(10) = 5$. Finally $\alpha(6) = 12$ and $\alpha(12) = 6$. Therefore we write α as $(1, 3, 4, 8)(5, 12, 9, 10)(6, 11)$.

A k-cycle can be written as follows

$$(i_1, i_2, \dots, i_k) = (i_1, i_2)(i_1, i_3)\dots(i_1, i_{k-1})(i_1, i_k), \quad k \geqslant 2. \tag{2.1}$$

We conclude from (2.1) that every $\alpha \in S_n$ can be written as a product of transpositions, that is $S_n = \langle\{(i, j) : i, j \in \{1, 2, \dots, n\}\}\rangle$, $n \geqslant 2$. However this can be done in many ways, e.g.

$$(2, 4, 5) = (2, 4)(2, 5) = (1, 3)(1, 2)(1, 4)(3, 5)(1, 5)(2, 4).$$

Observe that $(i_1, i_2)^2 = I$ and in consequence $(i_1, i_2)^{-1} = (i_1, i_2)$.

Lemma 2.4. *Every $\alpha \in S_n$ can be written uniquely as the product*

$$\alpha = \alpha_2\alpha_3\dots\alpha_n, \quad \alpha_i \in L_i, \tag{2.2}$$

where $L_i = \{(1, i), (2, i), \dots, (i-1, i), I\}$, $i = 2, 3, \dots, n$.

Note that $|L_i| = i$ and $L_i \cap L_j = I$ for $i \neq j$. The proof of Lemma 2.4 is in the Appendix. In the example below we outline its idea.

Example 2.5. Let $\alpha = (2, 7, 4, 5, 3)(1, 8, 6) \in S_8$. We will explain in detail how to write α in the form (2.2). Since $\alpha(8) = 6$ we multiply α by $(6, 8) \in L_8$ and get

$$\alpha \cdot (6, 8) = (2, 7, 4, 5, 3)(1, 6),$$

with $8 \in Fix(\alpha \cdot (6, 8))$. Now we take 7 and see that $(\alpha \cdot (6, 8))(7) = 4$. Multiplying this permutation by $(4, 7) \in L_7$ we obtain

$$\alpha \cdot (6, 8)(4, 7) = (2, 4, 5, 3)(1, 6).$$

In this moment $\{7, 8\} \in Fix(\alpha \cdot (6, 8)(4, 7))$. We repeat this procedure until the identity appears on the right hand side of an equation. After computation we have

$$\alpha \cdot (6, 8)(4, 7)(1, 6)(3, 5)(3, 4)(2, 3) = I,$$

and in consequence $\alpha = (2, 3)(3, 4)(3, 5)(1, 6)(4, 7)(6, 8)$. Hence $\alpha_2 = I$, $\alpha_3 = (2, 3)$, $\alpha_4 = (3, 4)$, $\alpha_5 = (3, 5)$, $\alpha_6 = (1, 6)$, $\alpha_7 = (4, 7)$ and $\alpha_8 = (6, 8)$.

One can use Lemma 2.4 to generate a random element of S_n. Namely, from every L_i we choose independently a transposition with probability $1/i$ for $i = 2, \dots, n$. Then multiplying those transpositions as in (2.2) we obtain a random permutation. The probability of getting a particular permutation equals

$$p = \prod_{i=2}^{n} \frac{1}{|L_i|} = \prod_{i=2}^{n} \frac{1}{i} = \frac{1}{n!} = \frac{1}{|S_n|}.$$

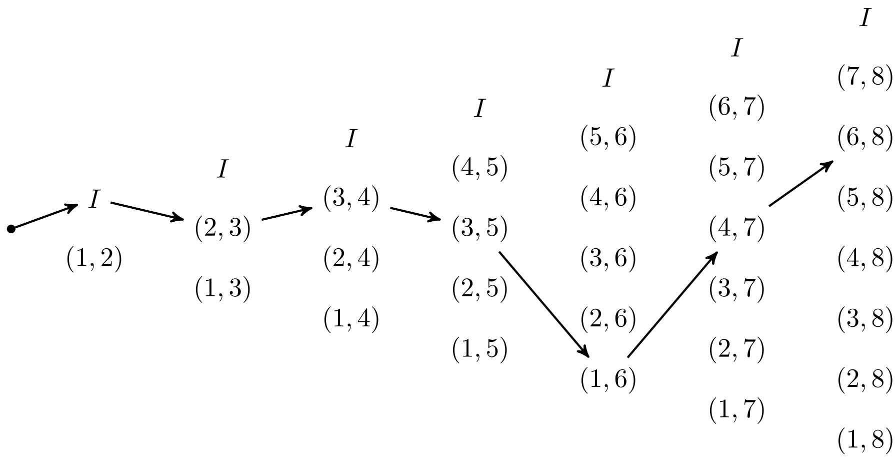


Figure 1: The procedure for generating a random permutation from S_8, see Example 2.1.

As we have already mentioned an element of S_n can be written as a product of transpositions in different ways. However one may prove that the number of transpositions which occur is either always even or always odd, see Chapter 6, [1]. Therefore $\alpha \in S_n$ is called an even permutation if it can be expressed as the product of even number of transpositions. Similarly $\alpha \in S_n$ is called an odd permutation if it is not an even permutation.

The subset of even permutations in S_n forms a subgroup of order $n!/2$ and is called the alternating group denoted by A_n (or $Alt(X)$ for a general set), see Theorem 6.4, [1].

Our question is: how to generate a random permutation from A_n? First observe that by (2.1) a 3-cycle is an even permutation and A_n is generated by all 3-cycles, i.e.

$$A_n = \langle \{(i,j,k) : i,j,k \in \{1,2,\ldots,n\}\} \rangle, \quad n \geqslant 3.$$

Indeed, we have

$$\begin{cases} (i,j)(i,j) = I, \\ (i,j)(i,k) = (i,j,k) \\ (i,j)(k,l) = (i,j,k)(k,i,l). \end{cases}$$

For example the commutator $[\alpha,\beta] := \alpha\beta\alpha^{-1}\beta^{-1}$ is always an even permutation for any $\alpha,\beta \in S_n$.

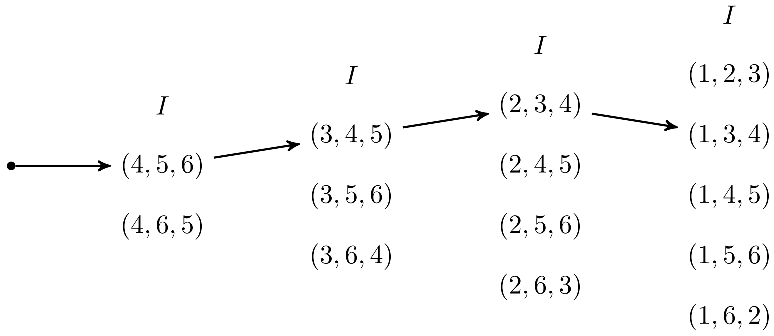


Figure 2: The procedure for generating a random permutation from A_6, see Example 2.8.

Example 2.6. If $\alpha = (1,9,2,3,8,5)(4,6)$ then

$$\alpha = (1,9)(1,2)(1,3)(1,8)(1,5)(4,6) = (1,9,2)(1,3,8)(1,5,4)(4,1,6).$$

On the other hand, e.g. $\alpha = (1,4,5)(2,4,9)(5,6,8)(3,8,4)$.

Lemma 2.7. *Every $\alpha \in A_n$ can be written uniquely as the product*

$$\alpha = \alpha_{n-2}\alpha_{n-3}\ldots\alpha_1, \quad \alpha_i \in K_i, \tag{2.3}$$

where

$$K_i = \{(i,i+1,i+2),(i,i+2,i+3),\ldots,(i,n-1,n),(i,n,i+1),I\},$$

$i = 1,\ldots,n-2$.

Note that $|K_i| = n-i+1$ and $K_i \cap K_j = I$ for $i \neq j$. The proof of Lemma 2.7 is explained in the Appendix.

Example 2.8. Let $\alpha = (1,3,2,4)(5,6) \in A_6$. We begin with 1. Since $\alpha(1) = 3$ we multiply α by $(1,4,3) = (1,3,4)^{-1}$, where $(1,3,4) \in K_4$. We have $\alpha \cdot (1,4,3) = (2,3)(5,6)$, with $1 \in Fix(\alpha \cdot (1,4,3))$. After several steps we obtain

$$\alpha \cdot (1,4,3)(2,4,3)(3,5,4)(4,6,5) = I,$$

and as a consequence of this

$$\alpha = [(1,4,3)(2,4,3)(3,5,4)(4,6,5)]^{-1} = (4,5,6)(3,4,5)(2,3,4)(1,3,4).$$

Choosing at random $\alpha_{n-2} \in K_{n-2}$, …, $\alpha_1 \in K_1$ and multiplying them as in (2.3) we will get an even permutation with probability

$$p = \prod_{i=1}^{n-2} \frac{1}{|K_i|} = \prod_{i=1}^{n-2} \frac{1}{n-i+1} = \frac{2}{n!} = \frac{1}{|A_n|}.$$

3. Group actions, orbits and stabilizers

Let G and H be groups. A function $\varphi : G \to H$ is said to be a homomorphism if $\varphi(ab) = \varphi(a)\varphi(b)$ for all $a, b \in G$. If φ is also a bijection then it is called an isomorphism. Denote neutral element in G and H by e_G, e_H respectively. The kernel of φ, defined by $Ker(\varphi) := \{a \in G : \varphi(a) = e_G\}$ is a normal subgroup of G, see Theorem 16.1, [1].

Definition 3.1. An action of a group G on a set X is a homomorphism from G to $Sym(X)$. That is $\varphi : G \to Sym(X)$ satisfies

(i) $\varphi_e = I$,

(ii) $\varphi_{gh} = \varphi_g \varphi_h, \quad g, h \in G.$

If $G = \{g_1, \ldots, g_n\}$ then by $G(x)$ or x^G we denote the orbit of $x \in X$, i.e.

$$G(x) := x^G = \{\varphi_{g_1}(x), \varphi_{g_2}(x), \ldots, \varphi_{g_n}(x)\}.$$

We have $X = \bigcup_{i=1}^m \mathcal{O}_i$ and $\mathcal{O}_i \cap \mathcal{O}_j = \emptyset$, $i \neq j$. We say that an action φ is faithful if $Ker(\varphi) = \{e_G\}$.

Definition 3.2. A group G acting on a set X is said to be transitive on X if $x^G = X$ for every $x \in X$.

Example 3.3. Let $G = \{e, a, b, ab\}$ be a Klein group and X a set containing six elements $x_1, \ldots, x_6$. For example X may be a set of vertices of a hexagon. Define $\varphi : G \to Sym(X)$ by: $\varphi_e = I$, $\varphi_a = (x_2, x_4)(x_3, x_6)$, $\varphi_b = (x_1, x_5)(x_3, x_6)$ and $\varphi_{ab} = (x_1, x_5)(x_2, x_4)$. Clearly φ is faithful. There are three orbits of this action: $\mathcal{O}_1 = \{x_1, x_5\}$, $\mathcal{O}_2 = \{x_2, x_4\}$, $\mathcal{O}_3 = \{x_3, x_6\}$ hence G is not transitive. Define the second action $\psi : G \to Sym(X)$ by: $\psi_e = I$, $\psi_a = \psi_b = (x_1, x_5)$, and therefore $\psi_{ab} = I$. We have $Ker(\psi) = \{I, ab\}$. There are 5 orbits of this action: $\mathcal{O}_1 = \{x_1, x_5\}$, $\mathcal{O}_2 = \{x_2\}$, $\mathcal{O}_3 = \{x_3\}$, $\mathcal{O}_4 = \{x_4\}$, $\mathcal{O}_5 = \{x_6\}$.

Definition 3.4. Let $\mathcal{A}$ be a subset of X. The pointwise stabilizer of $\mathcal{A}$ in G is

$$G_{(\mathcal{A})} := \{g \in G : g(x) = x \text{ for all } x \in \mathcal{A}\}.$$

and the setwise stabilizer of $\mathcal{A}$ in G is $G_{\{\mathcal{A}\}} := \{g \in G : g(\mathcal{A}) = \mathcal{A}\}$.

Obviously $G_{(\mathcal{A})} \subset G_{\{\mathcal{A}\}}$. If $\mathcal{A} = \{x\}$ then $G_{(x)} = G_{\{x\}}$ and we denote this subset of G by G_x. For $\mathcal{A} = \{x_1, \ldots, x_d\}$ we will also denote by $G_{x_1,\ldots,x_d}$ the pointwise stabilizer of $\mathcal{A}$.

Remark 3.5. Both $G_{(\mathcal{A})}$ and $G_{\{\mathcal{A}\}}$ are subgroups of G and one can prove that $G_{(\mathcal{A})}$ is a normal subgroup of $G_{\{\mathcal{A}\}}$, see e.g. [6].

Example 3.6. Let $X = \{1, 2, 3, 4, 5, 6, 7\}$ and $G = \langle \alpha \rangle$, where

$$\alpha = (1, 4, 5, 2)(3, 6, 7).$$

Since $\alpha^{12} = I$ then G is isomorphic to a cyclic group of order 12. For $\mathcal{A} = \{1, 5\}$ we have $G_{(\mathcal{A})} = \langle \alpha^4 \rangle = \{I, \alpha^4, \alpha^8\}$, with $\alpha^4 = (3, 6, 7)$. Observe that $G_{(\mathcal{A})} = G_{(\mathcal{B})}$, where $\mathcal{B} = \{1, 2, 4, 5\}$. Since $\alpha^2 = (1, 5)(2, 4)(3, 7, 6)$ then $G_{\{\mathcal{A}\}} = \langle \alpha^2 \rangle = \{I, \alpha^2, \alpha^4, \alpha^6, \alpha^8, \alpha^{10}\}$.

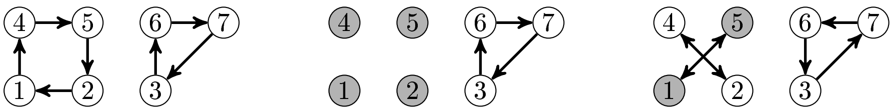


Figure 3: The illustration of Example 3.6.

The connection between orbits and stabilizers of elements of X is formulated in the following theorem.

Theorem 3.7 (Theorem 1.4A, [2])**.** *Suppose that G is a group acting on a set X and $x_1, x_2 \in X$. If $x_2 = \varphi_g(x_1)$, then $G_{x_1} = gG_{x_2}g^{-1}$ and $|G : G_x| = |x^G|$ for $x \in X$. Hence if G is finite then*

$$|G| = |x^G| \cdot |G_x|, \quad \forall\, x \in X. \tag{3.1}$$

4. The Schreier-Sims algorithm

If H is a subgroup of a group G and a is an arbitrary element of G then the set $aH = \{ah : h \in H\}$ is called a left coset of H. Similarly , the set $Ha = \{ha : h \in H\}$ is called a right coset of H. A set of consisting of right cosets (or left cosets) of H forms a partition of G, see e.g. [1], [4]. Since $a \in Ha$ we call a the representative of Ha. Suppose that G may be written in the form

$$G = Ha_1 \cup Ha_2 \cup \ldots \cup Ha_r$$

where $a_1 = e$ and $Ha_i \cap Ha_j = \emptyset$ for $i \neq j$. Then the set $T = \{a_1, \ldots, a_r\}$ is called a right transversal for G mod H or a set of right coset representatives for H in G. If H and T are given then

$$\overline{g} := Hg \cap T, \quad g \in G. \tag{4.1}$$

The Schreier-Sims algorithm is based on the following Schreier theorem, see for example [2] or [7].

Theorem 4.1 (Theorem 3.6.A, [2]). *Let H be a subgroup of a finite group G and let T be a set of right coset representatives for H in G. Assume additionally that $I \in T$. If W is a set of generators of G, then*

$$V := \{tw(\overline{tw})^{-1} : w \in W, t \in T\} \tag{4.2}$$

is a set of generators for H.

Remark 4.2. Note that $|V| = |W| \cdot |T|$ so V is usually a large set. However in many cases it may be reduced to much smaller set, see for example the chapter devoted to the Schreier-Sims algorithm in [7].

Definition 4.3. A subset $\mathcal{B} = \{x_1, x_2, \ldots, x_d\}$ of X is a base for G if $G_{(\mathcal{B})} = I$.

Now we are in a position to describe the algorithm.

In the first step let $G = \langle W \rangle$. Choose $x_1 \in X$ that lies in the support of some element of W. Then compute the orbit $\mathcal{O}_1$ of x_1 and the set of right coset representatives T_1 for G_{x_1}. Using W, T_1 and Theorem 4.1 one can compute a set V_1 of generators of G_{x_1}. Then try to reduce V_1 to a smaller set.

Next take G_{x_1} and V_1 and repeat this procedure from the first step until, for some $d \geqslant 1$ we get $G_{x_1,\ldots,x_d} = I$. As a result we obtain a chain of stabilizers

$$G \supset G_{x_1} \supset G_{x_1,x_2} \supset \ldots \supset G_{x_1,\ldots,x_d} = I,$$

a set of transversals $T_1, \ldots, T_d$ and orbits $\mathcal{O}_1, \ldots, \mathcal{O}_d$. Note that $|\mathcal{O}_i| = |T_i|$, $i = 1, \ldots, d$. From (3.1) we conclude that $|G| = \prod_{i=1}^{d} |\mathcal{O}_i|$.

Theorem 4.4. *Every $g \in G$ can be written uniquely as the product*

$$g = t_d t_{d-1} \ldots t_2 t_1, \quad t_i \in T_i, \quad i = 1, \ldots, d. \tag{4.3}$$

Proof. Let x_1 be a first element of X for which $g(x_1) \neq x_1$. Denote $x_2 = g(x_1)$. Since $x_2 \in x_1^G$ then we can find, say $t_1 \in T_1$, such that $t_1(x_1) = x_2$. Note that t_1 may be chosen in a unique way. Then observe that $gt_1^{-1}(x_1) = x_1$, i.e. $x_1 \in Fix(gt_1^{-1})$. Now let x_2 be a first element of X for which $gt_1^{-1}(x_2) \neq x_2$, denote $x_3 = gt_1(x_2)$ and continue this process described above. After d steps we get

$$gt_1^{-1}t_2^{-1} \ldots t_d^{-1} = I, \tag{4.4}$$

that is just (4.3). $\square$

Remark 4.5. We have the solution to the following problem: how to get a random permutation of $G = \langle W \rangle$? Answer: choose at random $t_1 \in T_1, \dots, t_d \in T_d$ and multiply them as in (4.3). In this way we get an element of G with probability

$$p = \prod_{i=1}^{d} \frac{1}{|T_i|} = \prod_{i=1}^{d} \frac{1}{|\mathcal{O}_i|} = \frac{1}{|G|}.$$

5. Examples

Example 5.1. Let $X = \{1, 2, 3, 4, 5, 6\}$ and consider

$$G = \langle \alpha, \beta \rangle, \quad \alpha = (1, 2, 3, 4), \quad \beta = (1, 5, 6, 2). \tag{5.1}$$

There is a very interesting method to prove that $|G| = 120$ and the problem of computing $|G|$ is in a way unique in the theory of permutations, see [8]. However we will not focus attention on this and we will use the Schreier-Sims algorithm.

Figure 4: The generators of G given by (5.1).

We begin with an orbit of $x_1 = 1$ (we can choose x_1 because $x_1 \in Act(\alpha)$): $1 \to 1(I)$, $1 \to 2(\alpha)$, $1 \to 3(\alpha^2)$, $1 \to 4(\alpha^3)$ and $1 \to 5(\beta)$, $1 \to 6(\beta^2)$. Hence $\mathcal{O}_1 = 1^G = \{1, 2, 3, 4, 5, 6\}$ and $T_1 = \{I, \alpha, \alpha^2, \alpha^3, \beta, \beta^2\}$. Now we are ready to apply (4.2). In our case $H = G_1$, V is denoted by V_1 and $W = \{\alpha, \beta\}$. The set of Schreier generators of G_1 is the following

$$V_1 = \Big\{ \alpha(\overline{\alpha})^{-1}, \beta(\overline{\beta})^{-1}, \alpha^2(\overline{\alpha^2})^{-1}, \alpha\beta(\overline{\alpha\beta})^{-1}, \alpha^3(\overline{\alpha^3})^{-1}, \alpha^2\beta(\overline{\alpha^2\beta})^{-1},$$
$$I(\overline{I})^{-1}, \alpha^3\beta(\overline{\alpha^3\beta})^{-1}, \beta\alpha(\overline{\beta\alpha})^{-1}, \beta^2(\overline{\beta^2})^{-1}, \beta^2\alpha(\overline{\beta^2\alpha})^{-1}, \beta^3(\overline{\beta^3})^{-1} \Big\}.$$

Now we need to find the representatives: $\overline{\alpha} = \alpha$, $\overline{\beta} = \beta$, $\overline{\alpha^2} = \alpha^2$, $\overline{\alpha\beta} = I$ (because $1 \to 1(\alpha\beta)$), $\overline{\alpha^3} = \alpha^3$, $\overline{\alpha^2\beta} = \alpha^2$ (note that $\alpha^2\beta(1) = 3 = \alpha^2(1)$), $\overline{I} = I$, $\overline{\beta^2} = \beta^2$, $\overline{\beta\alpha} = \beta$ (we have $\beta\alpha(1) = 5 = \beta(1)$), $\overline{\beta^2\alpha} = \beta^2$ ($\beta^2\alpha(1) = \alpha(6) = 6 = \beta^2(1)$) and $\overline{\beta^3} = \beta^3$. For example $\alpha^2\beta(\overline{\alpha^2\beta})^{-1} = \alpha^2\beta(\alpha^2)^{-1} = \alpha^2\beta\alpha^2$, because $\alpha^{-2} = \alpha^2$. We can reduce V_1 to the following set (that we denote also by V_1)

$$V_1 = \left\{ \alpha\beta, \alpha^2\beta\alpha^2, \alpha^3\beta\alpha^{-3}, \beta\alpha\beta^{-1}, \beta^2\alpha\beta^2 \right\}.$$

We have: $\alpha\beta = (2,3,4,5,6)$, $\alpha^2\beta\alpha^2 = (3,5,6,4)$, $\alpha^3\beta\alpha^{-3} = (2,5,6,3)$, $\beta\alpha\beta^{-1} = (2,6,3,4)$ and $\beta^2\alpha\beta^2 = (3,4,6,5)$. We can finally reduce V_1 to $\{\alpha^3\beta\alpha^{-3}, \beta\alpha\beta^{-1}\}$ because $(2,6,3,4)\cdot(2,5,6,3) = (2,3,4,5,6)$, $(2,6,3,4)^2 \cdot (2,5,6,3) = (3,5,6,4)$ and $(3,4,6,5)^3 = (3,5,6,4)$. So we have

$$G_1 = \langle \alpha_1, \beta_1 \rangle, \quad \alpha_1 = (2,5,6,3), \beta_1 = (2,6,3,4),$$

with $\alpha_1 = \alpha^3\beta\alpha^{-3} = \alpha^{-1}\beta\alpha$ and $\beta_1 = \beta\alpha\beta^{-1}$. Additionaly $|G| = 6|G_1|$ by (3.1).

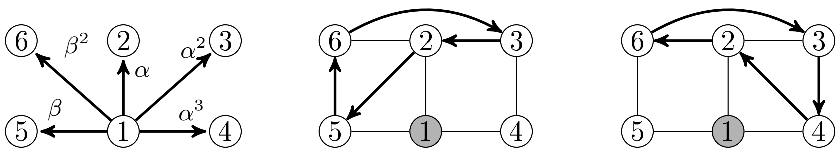


Figure 5: The transversal T_1 and generators of G_1, see Example 5.1.

Now take e.g. $x_2 = 2$. We have: $2 = I(2)$, $3 = \beta_1^2(2)$, $4 = \beta_1^3(2)$, $5 = \alpha_1(2)$, $6 = \alpha_1^2(2)$. Hence $\mathcal{O}_2 = 2^{G_1} = \{2,3,4,5,6\}$ and $T_2 = \{I, \beta_1^2, \beta_1^3, \alpha_1, \alpha_1^2\}$. The set of Schreier generators of $G_{1,2}$ is

$$V_2 = \Big\{ I, \alpha_1(\overline{\alpha_1})^{-1}, \beta_1(\overline{\beta_1})^{-1}, \alpha_1^2(\overline{\alpha_1^2})^{-1}, \alpha_1\beta_1(\overline{\alpha_1\beta_1})^{-1}, \alpha_1^3(\overline{\alpha_1^3})^{-1},$$
$$\alpha_1^2\beta_1(\overline{\alpha_1^2\beta_1})^{-1}, \beta_1^2\alpha_1(\overline{\beta_1^2\alpha_1})^{-1}, \beta_1^3(\overline{\beta_1^3})^{-1}, \beta_1^3\alpha_1(\overline{\beta_1^3\alpha_1})^{-1}, \beta_1^4(\overline{\beta_1^4})^{-1} \Big\}.$$

It is clear that $\overline{\alpha_1} = \alpha_1$, $\overline{\alpha_1^2} = \alpha_1^2$ and $\overline{\beta_1^3} = \beta_1^3$. Since $\alpha_1^3(2) = 3$, $\alpha_1\beta_1(2) = 5$, $\alpha_1^2\beta_1(2) = 3$, $\beta_1^2\alpha_1(2) =$, $\beta_1^3\alpha_1(2) =$, $\beta_1^4 = I$ therefore $\overline{\alpha_1\beta_1} = \alpha_1$, $\overline{\alpha_1^2\beta_1} = \beta_1^2$, $\overline{\beta_1^2\alpha_1} = \beta_1^2$, $\overline{\beta_1^3\alpha_1} = \beta_1^3$, $\overline{\beta_1^4} = I$. This leads to

$$V_2 = \{\alpha_1\beta_1\alpha_1^{-1}, \alpha_1^3\beta_1^{-2}, \alpha_1^2\beta_1^{-1}, \beta_1^2\alpha_1, \beta_1^3\alpha_1\beta_1^{-3}\}.$$

We have $\alpha_1\beta_1\alpha_1^{-1} = \alpha_1^2\beta_1^{-1} = \beta_1^2\alpha_1 = (3,5,6,4)$ and $\alpha_1^3\beta_1^{-2} = \beta_1^3\alpha_1\beta_1^{-3} = (3,4,6,5)$. Observe that $(3,4,6,5)^3 = (3,5,6,4)$ hence V_2 is reduced to $(3,4,6,5)$. That is

$$G_{1,2} = \langle \alpha_2 \rangle, \quad \alpha_2 = (3,4,6,5),$$

with $\alpha_2 = \alpha_1^{-1}\beta_1^{-2} = \beta_1^{-1}\alpha_1\beta_1 = \beta^2\alpha\beta^2$. We have $|G_1| = 5|G_{1,2}|$.

For $x_3 = 3$ we have $T_3 = \{I, \alpha_2, \alpha_2^2, \alpha_2^3\}$ and $\mathcal{O}_2 = 2^{G_{1,2}} = \{3,4,5,6\}$. In fact $G_{1,2} = \{I, (3,4,6,5), (3,6)(4,5), (3,5,6,4)\}$ so $G_{1,2,3} = I$ and $|G_{1,2}| = 4$. In a consequence $\mathcal{B} = \{1,2,3\}$ and $|G| = 6 \cdot 5 \cdot 4 = 120$. Now we can apply Theorem 4.4. Taking for example $t_3 = (3,6)(4,5)$, $t_2 = (2,5,6,3)$ and $t_1 = (1,2,3,4)$ we obtain

$$g = t_3 t_2 t_1 = (3,4,6)(1,2,5) = \beta^2\alpha\beta^2\alpha^{-1}\beta\alpha^2,$$

since $t_1 = \alpha$, $t_2 = \alpha_1 = \alpha^{-1}\beta\alpha$ and $t_3 = \alpha_2^2 = \beta^2\alpha\beta^2$.

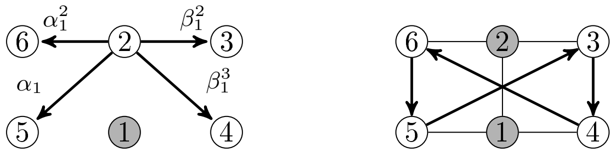


Figure 6: The transversal T_2 and the generator of $G_{1,2}$.

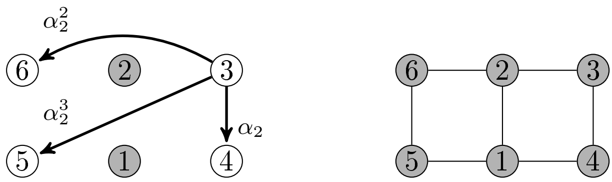


Figure 7: The transversal T_3 and $G_{1,2,3} = I$.

Table 1: The summary of the Schreier-Sims algorithm applied to the group given by (5.1).

group	*generators*	x	*orbit*	*order*
G	$\alpha = (1,2,3,4)$ $\beta = (1,5,6,2)$	1	$\mathcal{O}_1 = \{1,2,3,4,5,6\}$	$\lvert G\rvert = 6\lvert G_1\rvert$
G_1	$\alpha_1 = (2,5,6,3)$ $\beta_1 = (2,6,3,4)$	2	$\mathcal{O}_2 = \{2,3,4,5,6\}$	$\lvert G_1\rvert = 5\lvert G_{1,2}\rvert$
$G_{1,2}$	$\alpha_2 = (3,4,6,5)$	3	$\mathcal{O}_3 = \{3,4,5,6\}$	$\lvert G_{1,2}\rvert = 4$
$G_{1,2,3}$	$\alpha_3 = I$	–	–	$\lvert G_{1,2,3}\rvert = 1$

In the next example we discuss the membership problem.

Example 5.2. Let G be the group given by (5.1). Is it true that $\gamma = (1,3,4) \in G$? We will use Table 2 to answer this question. For T_1 we have $\alpha \cdot (1,3)(2,4) = (3,2,4)$. Next $(3,2,4) \cdot (2,6,3,4) = (3,6)$ for T_2 and $(3,6) \cdot (3,6)(4,5) = (4,5)$ for T_3. Therefore $\gamma \notin G$. In other words

$$\alpha \cdot \underbrace{(1,3)(2,4)}_{\in T_1^{-1}} \cdot \underbrace{(2,6,3,4)}_{\in T_2^{-1}} \cdot \underbrace{(3,6)(4,5)}_{\in T_3^{-1}} = (4,5) \neq I,$$

see Theorem 4.4 and (4.4). Let $\delta = (1,3,4)(2,6,5)$. Then $\delta \in G$ since we

Table 2: Transversals from Example 5.1. Note that $\alpha_1 = \alpha^{-1}\beta\alpha$, $\beta_1 = \beta\alpha\beta^{-1}$ and $\alpha_2 = \beta^2\alpha\beta^2$.

T_1	$=$	T_2	$=$	T_3	$=$
I	I	I	I	I	I
α	$(1,2,3,4)$	β_1^2	$(2,3)(4,6)$	α_2	$(3,4,6,5)$
α^2	$(1,3)(2,4)$	β_1^3	$(2,4,3,6)$	α_2^3	$(3,5,6,4)$
α^3	$(1,4,3,2)$	α_1	$(2,5,6,3)$	α_2^2	$(3,6)(4,5)$
β	$(1,5,6,2)$	α_1^2	$(2,6)(3,5)$		
β^2	$(1,6)(2,5)$				

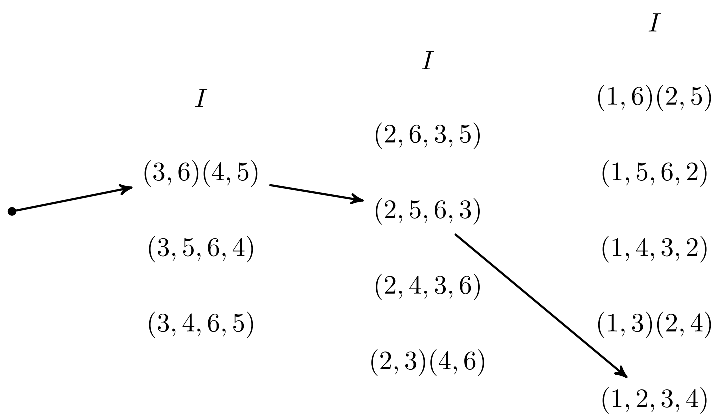


Figure 8: The procedure for generating a random permutation from G given by (5.1). Note that $|G| = 120$ and the product of $(3,6)(4,5)$, $(2,5,6,3)$ and $(1,2,3,4)$ is $(3,4,6)(1,2,5)$.

have

$$\alpha \cdot \underbrace{(1,3)(2,4)}_{\in T_1^{-1}} \cdot \underbrace{(2,6)(3,5)}_{\in T_2^{-1}} \cdot \underbrace{(3,6)(4,5)}_{\in T_3^{-1}} = I.$$

Note that $\delta' = (1,3,4)(2,5,6) \notin G$. In fact $\delta\delta' = (1,4,3) = \gamma^2$ and for this reason it is not possible.

Remark 5.3. If α and β are given by (5.1) then

$$\langle \alpha, \beta \rangle = \langle \alpha^2, \beta \rangle = \langle \alpha, \beta^2 \rangle. \tag{5.2}$$

Obviously $\langle \alpha^2, \beta \rangle \subset \langle \alpha, \beta \rangle$. One can check that $\alpha = \beta^2\alpha^2\beta^{-1}\alpha^2\beta^2$ hence $\alpha \in \langle \alpha^2, \beta \rangle$ which implies $\langle \alpha, \beta \rangle \subset \langle \alpha^2, \beta \rangle$. This proves the first equality in (5.2). The second follows from the fact that $\beta = \alpha^2\beta^2\alpha^{-1}\beta^2\alpha^2$. For a group $\langle \alpha^2, \beta^2 \rangle$ note that it is isomorphic to S_3, i.e.

$$\langle \alpha^2, \beta^2 \rangle = \{I, \alpha^2, \beta^2, \alpha^2\beta^2, \beta^2\alpha^2, \alpha^2\beta^2\alpha^2\}.$$

Example 5.4. Let $X = \{1, 2, \dots, 10\}$ and $G = \langle \alpha, \beta, \gamma, \delta \rangle$ with

$$\alpha = (1, 6)(5, 10),\ \beta = (2, 7)(5, 10),\ \gamma = (3, 8)(5, 10),\ \delta = (4, 9)(5, 10).$$

Observe that G is a subgroup of A_{10}, since all its generators are even. Clearly G does not act transitively on X. Calculations show that for $x_1 = 1$ we have $T_1 = \{I, \alpha\}$, $\mathcal{O}_1 = \{1, 6\}$ and $G_1 = \langle \beta, \gamma, \delta \rangle$. If $x_2 = 2$ then $T_2 = \{I, \beta\}$, $\mathcal{O}_2 = \{2, 7\}$ and $G_{1,2} = \langle \gamma, \delta \rangle$. Next we take $x_3 = 3$ and get $T_3 = \{I, \gamma\}$, $\mathcal{O}_3 = \{3, 8\}$. Moreover $G_{1,2,3} = \langle \delta \rangle = \{I, \delta\}$. Finally for $x_4 = 4$ we have $T_4 = \{I, \delta\}$, $\mathcal{O}_4 = \{4, 9\}$ and $G_{1,2,3,4} = I$. We conclude that $|G| = 2^4 = 16$.

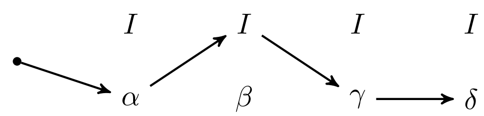


Figure 9: The procedure for generating a random permutation from the group in Example 5.4.

In the final example we consider a subgroup of a wreath product of groups.

Example 5.5. Let $X = \{1, 2, 3, 4, 5, 6, 7, 8\}$ and

$$G = \langle \alpha, \beta \rangle, \quad \alpha = (1, 6, 3, 8)(2, 7, 4, 5), \quad \beta = (1, 2)(5, 6). \tag{5.3}$$

Note that G acts transitively on X. If we take $x_1 = 1$ then $\mathcal{O}_1 = X$ with $T_1 = \{I, \beta, \alpha^2, \beta\alpha^2, \beta\alpha^3, \alpha, \beta\alpha, \alpha^3\}$, since $\beta(1) = 2$, $\alpha^2(1) = 3$, ... , and $\alpha^3(1) = 8$. Then we compute elements of V_1 (there are 16 elements in V_1 at the beginning of calculations) and reduce to $V_1 = \{(3, 4)(7, 8), (2, 4)(6, 8)\}$. Therefore the stabilizer of x_1 is as follows

$$G_1 = \langle \alpha_1, \beta_1 \rangle, \quad \alpha_1 = (3, 4)(7, 8), \quad \beta_1 = (2, 4)(6, 8), \tag{5.4}$$

Table 3: The summary of the Schreier-Sims algorithm applied to the group from Example 5.4.

group	*generators*	x	*orbit*	*order*
G	α, β, γ, δ	1	$\mathcal{O}_1 = \{1, 6\}$	$\lvert G\rvert = 2\lvert G_1\rvert$
G_1	β, γ, δ	2	$\mathcal{O}_2 = \{2, 7\}$	$\lvert G_1\rvert = 2\lvert G_{1,2}\rvert$
$G_{1,2}$	γ, δ	3	$\mathcal{O}_3 = \{3, 8\}$	$\lvert G_{1,2}\rvert = 2\lvert G_{1,2,3}\rvert$
$G_{1,2,3}$	δ	4	$\mathcal{O}_4 = \{4, 9\}$	$\lvert G_{1,2,3}\rvert = 2$
$G_{1,2,3,4}$	I	–	–	$\lvert G_{1,2,3,4}\rvert = 1$

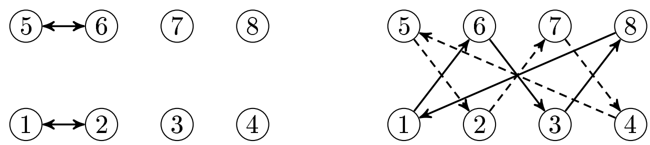


Figure 10: The generators of G given by (5.3).

with $\alpha_1 = \alpha^2\beta\alpha^2$, $\beta_1 = (\beta\alpha)\beta(\beta\alpha)^{-1}$. For $x_2 = 3$ we have $\mathcal{O}_2 = \{2, 3, 4\}$ and $T_2 = \{\alpha_1\beta_1, I, \alpha_1\}$. After computation we obtain $V_2 = \{\beta_1\}$ and in a consequence $G_{1,3} = \langle\beta_1\rangle$. Therefore if $x_3 = 4$ then $G_{1,3,4} = I$. Recall that $\mathcal{O}_3 = \{2, 4\}$ and $T_3 = \{I, \beta_1\}$. Summarizing, the base $\mathcal{B} = \{1, 3, 4\}$ and $|G| = 8 \cdot 3 \cdot 2 = 48$.

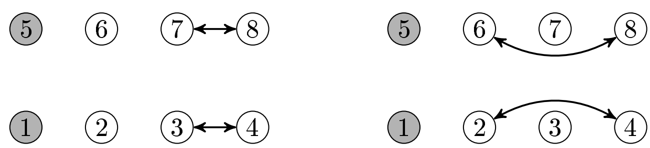


Figure 11: The generators of G_1 given by (5.4).

Remark 5.6. If we consider the group G given by (5.3) as a subgroup of S_8 then its index in S_8 is $|S_8 : G| = |S_8|/|G| = 8!/48 = 840$. However G is in fact a subgroup of $\mathbb{Z}_2 \wr S_4$, i.e. the wreath product of $\mathbb{Z}_2$ by S_4. Since $|\mathbb{Z}_2 \wr S_4| = 4! \cdot 2^4 = 384$ then $|(\mathbb{Z}_2 \wr S_4) : G| = 8$. For more details about wreath products see e.g. [3], [4].

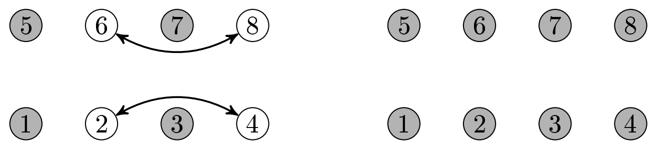


Figure 12: The generator of $G_{1,3}$ and $G_{1,3,4} = I$.

6. Appendix

To prove Lemma 2.4 and Lemma 2.7 we state the following the theorem.

Theorem 6.1 (Theorem 3.3A, [2]). *Let G be a primitive subgroup of $Sym(X)$.*

(i) *If G contains a 3-cycle, then $Alt(X) \subset G$.*

(ii) *If G contains a 2-cycle, then $G = Sym(X)$.*

The conclusions of Lemma 2.4 and Lemma 2.7 follow from Theorem 4.4 and Theorem 6.1. Namely, for S_n the base is e.g. $\mathcal{B} = \{n, n-1, \ldots, 2\}$ and $L_n, \ldots, L_2$ are transversals. For A_n we have e.g. $\mathcal{B} = \{1, 2, \ldots, n-2\}$ and $K_1, \ldots, K_{n-2}$ are appropriate transversals.

References

1. M. A. Armstrong, *Groups and symmetry*, Undergraduate texts in mathematics, Springer-Verlag, 1988.
2. J. D. Dixon, B. Mortimer, *Permutation groups*, Graduate Texts in Mathematics 163, Springer-Verlag, 1996.
3. J. A. Eidswick, *Cubelike puzzles; what are they and how do you solve the*?, "Am. Math. Month.", 1986 vol. 93(3), pp. 157-176.
4. M. Hall, *The theory of groups*, Macmillan, New York, 1959.
5. D. Joyner, *Adventures in group theory*, The Johns Hopkins Univeristy Press, 2008.
6. A. I. Kostrikin, *Wstep do algebry*, t. 1-3, Wydawnictwo naukowe PWN, Warszawa 2009.
7. A. Seress, *Permutation group algorithms*, Cambridge Tracts in Mathematics, 152, Cambridge University Press, 2003.
8. D. Singmaster, *Notes on Rubik's magic cube*, Enslow Publishers, 1981.

Examples in stochastic differential equations

Ernest Nieznaj[1]

Abstract

This paper is a short overview of Gaussian and Markov processes, especially those related to a Brownian motion and stochastic differential equations.

1. Preliminaries

We say that a real-valued random variable X has Gaussian distribution if its density is

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}, \quad x \in \mathbb{R},$$

with $\mu \in \mathbb{R}$ and $\sigma^2 > 0$. We write $X \sim N(\mu, \sigma^2)$. It is very well known that $\mathbb{E}X = \mu$ and $\mathrm{var}X = \sigma^2$. If $\sigma^2 = 0$ then we mean $\mathbb{P}(X = \mu) = 1$. Similarly a n-dimensional random vector $(X_1, \ldots, X_n)$ has Gaussian distribution if its density is given by

$$f(x_1, \ldots, x_n) = \frac{1}{\sqrt{(2\pi)^n|Q|}} \exp\left\{-\sum_{i,j=1}^{n} \frac{Q_{ij}(x_i - m_i)(x_j - m_j)}{2|Q|}\right\},$$

where $m_i = \mathbb{E}X_i$, $Q = [q_{ij}]$, $|Q| = \det Q$, $q_{ij} = \mathrm{cov}(X_i, X_j)$, and Q_{ij} is an algebraic complement of q_{ij}, $i, j = 1, \ldots, n$. Recall that $\mathrm{cov}(X_i, X_j) = \mathbb{E}(X_i - m_i)(X_j - m_j)$ and $\mathrm{var}X = \mathrm{cov}(X, X)$. In this paper we assume that $|Q| \neq 0$. Let $X \sim N(m_1, \sigma_1^2)$ and $Y \sim N(m_2, \sigma_2^2)$. If, in addition, X and Y are independent then one may prove that $X + Y \sim N(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$, see e.g. [3], Chapter 5. Given X, Y we say that they are uncorrelated if $\mathrm{cov}(X, Y) = 0$. Recall that if X, Y are Gaussian then they

[1]Department of Mathematics, Technical University of Lublin, Nadbystrzycka 38, 20-618 Lublin, Poland, *e-mail: e.nieznaj@pollub.pl*
2010 Mathematics Subject Classification. Primary 60615, 60J25; Secondary 60J60.
Keywords: Gaussian processes, Markov processes, stochastic differential equations.

are independent if and only if $\operatorname{cov}(X,Y)=0$. For a random process X_t, $t\in T$, we denote $m(t)=\mathbb{E}X_t$ and $K(t_1,t_2)=\operatorname{cov}(X_{t_1},X_{t_2})$ for all $t,t_1,t_2\in T$. A stochastic process is said to be a Gaussian process if all its finite-dimensional distributions are Gaussian. Equivalently X_t is Gaussian if for all real numbers $a_1,\dots,a_n$ the random variable $\sum_{i=1}^n a_iX_{t_i}$ is Gaussian, see [3], Chapter 5. Throughout the paper we assume that W_t, $t\geqslant 0$, is a standard Brownian motion, that is a Gaussian stochastic process with $\mathbb{E}W_t=0$, $K(t_1,t_2)=\min\{t_1,t_2\}$ and continuous trajectories.

2. Gaussian and Markov processes

Example 2.1. Let $X_t=|W_t|, t\geqslant 0$. Clearly X_t is not Gaussian. We will prove that X_t is a Markov process by showing that

$$g(x_{n+1}|x_1,\dots,x_n)=g(x_{n+1}|x_n),\tag{2.1}$$

where $g(x_{n+1}|x_1,\dots,x_n):=g(x_1,\dots,x_{n+1})/g(x_1,\dots,x_n)$ and $g(x_1,\dots,x_n)$ is the density of $(X_{t_1},\dots,X_{t_n})$ for $0<t_1<t_2<\dots<t_n$. Since W_t has independent increments then it is a Markov process, and if $B\in\mathcal{B}(\mathbb{R})$ we have $\mathbb{P}(W_{t_2}\in B|W_{t_1}=x_1)=\int_B p(t,x_1,x_2)dx_2$ where

$$p(t,x_1,x_2)=\frac{1}{\sqrt{2\pi t}}\exp\left\{-\frac{(x_2-x_1)^2}{2t}\right\},\quad x_2\in\mathbb{R},\tag{2.2}$$

and $t=t_2-t_1>0$. Hence the density of $(W_{t_1},\dots,W_{t_n})$ is given by

$$\begin{aligned}f(x_1,\dots,x_n)&=\prod_{i=1}^n p(t_i-t_{i-1},x_{i-1},x_i)\\&=\prod_{i=1}^n\frac{1}{\sqrt{2\pi(t_i-t_{i-1})}}\exp\left\{-\frac{1}{2}\sum_{i=1}^n\frac{(x_i-x_{i-1})^2}{t_i-t_{i-1}}\right\}\end{aligned}$$

with $t_0=0,x_0=0$. In particular for $t_1<t_2$ the density of (W_{t_1},W_{t_2}) is

$$\begin{aligned}f(x_1,x_2)&=p(t_1,0,x_1)p(t_2-t_1,x_1,x_2)\\&=\frac{1}{2\pi\sqrt{t_1(t_2-t_1)}}\exp\left\{-\frac{t_2x_1^2+t_1x_2^2-2t_1x_1x_2}{2t_1(t_2-t_1)}\right\},\end{aligned}$$

where $m=(0,0)$ and the covariance matrix equals

$$Q=\begin{bmatrix}\operatorname{cov}(W_{t_1},W_{t_1}) & \operatorname{cov}(W_{t_1},W_{t_2})\\ \operatorname{cov}(W_{t_1},W_{t_2}) & \operatorname{cov}(W_{t_2},W_{t_2})\end{bmatrix}=\begin{bmatrix}t_1 & t_1\\ t_1 & t_2\end{bmatrix}.$$

Let $B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R}_+)$ and $-B := \{x \in \mathbb{R} : -x \in B\}$. Then

$$\mathbb{P}\left(|W_{t_1}| \in B_1, \ldots, |W_{t_n}| \in B_n\right) = \sum_{(k_1,\ldots,k_n)} \mathbb{P}\left(W_{t_1} \in (-1)^{k_1} B_1, \ldots, W_{t_n} \in (-1)^{k_n} B_n\right),$$

where the sum runs over all 0-1 sequences $(k_1, \ldots, k_n)$, that is $k_i = 0$ or 1 for $i = 1, \ldots, n$. Note that there are 2^n such sequences. Hence for $x_1, \ldots, x_n \geqslant 0$ we have

$$\begin{aligned} g(x_1, \ldots, x_n) &= \sum_{(k_1,\ldots,k_n)} f((-1)^{k_1} x_1, \ldots, (-1)^{k_n} x_n) \\ &= \sum_{(k_1,\ldots,k_n)} \prod_{i=1}^{n} p(t_i - t_{i-1}, (-1)^{k_{i-1}} x_{i-1}, (-1)^{k_i} x_i). \end{aligned}$$

We will show that $g(x_1, \ldots, x_{n+1}) = g(x_{n+1}|x_n) g(x_1, \ldots, x_n)$, where

$$g(x_{n+1}|x_n) = p(t_{n+1} - t_n, x_n, x_{n+1}) + p(t_{n+1} - t_n, x_n, -x_{n+1}). \tag{2.3}$$

In other words for any $t_1 < t_2$ we have

$$\mathbb{P}(X_{t_2} \in B | X_{t_1} = x_1) = \frac{1}{\sqrt{2\pi(t_2 - t_1)}} \int_B \left[e^{-\frac{(x_2 - x_1)^2}{2(t_2 - t_1)}} + e^{-\frac{(x_2 + x_1)^2}{2(t_2 - t_1)}} \right] dx_2.$$

Suppose that $n \geqslant 1$ and let $\mathbf{x}_n = (x_1, \ldots, x_n)$, $\mathbf{k}_n = (k_1, \ldots, k_n)$. Then

$$\begin{aligned} g(\mathbf{x}_{n+1}) &= \sum_{\mathbf{k}_{n+1}} \prod_{i=1}^{n+1} p(t_i - t_{i-1}, (-1)^{k_{i-1}} x_{i-1}, (-1)^{k_i} x_i) \\ &= \sum_{\mathbf{k}_n} [p(t_{n+1} - t_n, (-1)^{k_n} x_n, x_{n+1}) + p(t_{n+1} - t_n, (-1)^{k_n} x_n, -x_{n+1})] P_n \end{aligned}$$

with $P_n := \prod_{i=1}^n p(t_i - t_{i-1}, (-1)^{k_{i-1}} x_{i-1}, (-1)^{k_i} x_i)$. Clearly P_n does depend on k_i-s and x_i-s but we don't need to write that explicitly. Hence $g(\mathbf{x}_{n+1}) = S_1 + S_2$ where

$$S_1 = [p(t_{n+1} - t_n, x_n, x_{n+1}) + p(t_{n+1} - t_n, x_n, -x_{n+1})] \sum_{(\mathbf{k}_{n-1}, 0)} P_n$$

and

$$S_2 = [p(t_{n+1} - t_n, -x_n, x_{n+1}) + p(t_{n+1} - t_n, -x_n, -x_{n+1})] \sum_{(\mathbf{k}_{n-1}, 1)} P_n.$$

The key observation is that

$$p(t, x_n, x_{n+1}) + p(t, x_n, -x_{n+1}) = p(t, -x_n, x_{n+1}) + p(t, -x_n, -x_{n+1}),$$

see (2.2). Therefore we have

$$g(\mathbf{x}_{n+1}) = [p(t_{n+1} - t_n, x_n, x_{n+1}) + p(t_{n+1} - t_n, x_n, -x_{n+1})]g(\mathbf{x}_n),$$

because

$$\sum_{(\mathbf{k}_{n-1},0)} P_n + \sum_{(\mathbf{k}_{n-1},1)} P_n = \sum_{\mathbf{k}_n} P_n = g(\mathbf{x}_n).$$

We have just proved (2.1).

Example 2.2. Let $X_t = W_t^2$, $t \geqslant 0$. As in the previous example X_t is not a Gaussian process. However observe that for $B_1, \dots, B_n \in \mathcal{B}(\mathbb{R}_+)$ we have

$$\mathbb{P}(W_{t_1}^2 \in B_1, \dots, W_{t_n}^2 \in B_n) = \mathbb{P}(|W_{t_1}| \in \sqrt{B_1}, \dots, |W_{t_n}| \in \sqrt{B_n}),$$

where $\sqrt{B} = \{x \in \mathbb{R} : x^2 \in B\}$ for $B \in \mathcal{B}(\mathbb{R}_+)$. Therefore X_t is a Markov process because $|W_t|$ is, see Example 2.1. Note that $EW_t^2 = t$ and

$$\begin{aligned}\mathbb{E}(W_{t_1}^2 W_{t_2}^2) &= \mathbb{E}(W_{t_1} W_{t_1})\mathbb{E}(W_{t_2} W_{t_2}) + 2\mathbb{E}(W_{t_1} W_{t_2})\mathbb{E}(W_{t_1} W_{t_2}) \\ &= t_1 t_2 + 2(\min\{t_1, t_2\})^2,\end{aligned}$$

hence $K(t_1, t_2) = 2(\min\{t_1, t_2\})^2$.

For Gaussian processes the condition (2.1) is equivalent to

$$K(t_1, t_3)K(t_2, t_2) = K(t_1, t_2)K(t_2, t_3) \tag{2.4}$$

for all $t_1 < t_2 < t_3$, see e.g. [5]. For Gaussian Markov processes the covariance function can be characterize in more explicit form. First we cite an auxiliary lemma. The following is Lemma 5.1.8. in [3].

Lemma 2.3. *Let $p(t)$ and $q(t)$ be positive functions on $T \subset \mathbb{R}$ with $p(t)/q(t)$ strictly increasing. Define*

$$K(t_1, t_2) = \begin{cases} p(t_1)q(t_2), & t_1 \leqslant t_2 \\ p(t_2)q(t_1), & t_2 < t_1 \end{cases} \tag{2.5}$$

and suppose that p and q are such that $K(t_1, t_2) > 0$ for all $t_1, t_2 \in T$. Then $K(t_1, t_2)$ is a strictly positive definite function on $T \times T$.

Recall that a function $K(t_1,t_2)$ defined on $T\times T$ is a positive definite (or non-negative definite) function, if for every $n\geqslant 1$ and all and $t_1,\dots,t_n\in T$

$$\sum_{i,j=1}^{n} a_i a_j K(t_i,t_j)\geqslant 0 \tag{2.6}$$

for all real numbers $a_1,\dots,a_n$. We say that a positive definite function $K(t_1,t_2)$ is strictly positive definite if equality in (2.6) implies that $a_1=\dots=a_n=0$. The next is Lemma 5.1.9 in [3].

Lemma 2.4. *Let $T\subset\mathbb{R}$ be an open or closed interval and let X_t, $t\in T$ be a mean zero Gaussian process with continuous strictly positive definite covariance $K(t_1,t_2)$. Then X_t is a Gaussian Markov process if and only if $K(t_1,t_2)$ can be expressed as in* (2.5).

Example 2.5. Let $X_t=at+bW_t$, $t\geqslant 0$, where $a,b\in\mathbb{R}$ and $b\neq 0$. Observe that $X_t\sim N(at,b^2)$ and $\sum_{i=1}^n a_iX_{t_i}\sim N(m,\sigma^2)$ where $m=a\sum_{i=1}^n a_it_i$ and $\sigma^2=b^2(\sum_{i=1}^n a_i)^2$. The covariance

$$K(t_1,t_2)=b^2\mathbb{E}(W_{t_1}W_{t_2})=b^2\min\{t_1,t_2\},$$

clearly satisfies (2.4), so it is a Markov process. In fact increments of X_t are independent. Namely, if $t_1<t_2<t_3$ then

$$\operatorname{cov}(X_{t_2}-X_{t_1},X_{t_3}-X_{t_2})=b^2\mathbb{E}(W_{t_2}-W_{t_1})(W_{t_3}-W_{t_2})=0,$$

because increments of W_t are independent. For $b=1$ from (2.2) we obtain

$$\mathbb{P}(X_{t_2}\in B|X_{t_1}=x_1)=\frac{1}{\sqrt{2\pi t}}\int_B\exp\left\{-\frac{(x_2-x_1-at)^2}{2t}\right\}dx_2.$$

Example 2.6. Let $T>0$ and define

$$X_t=a+\frac{t}{T}(b-a-W_T)+W_t,\quad t\in\langle 0,T\rangle,$$

called the Brownian bridge. Observe that $x_0=a$, $x_T=b$ and this is a Gaussian process with $m(t)=a+t(b-a)/T$. The covariance function given by

$$K(t_1,t_2)=\min\{t_1,t_2\}-\frac{t_1t_2}{T},\quad 0\leqslant t_1,t_2\leqslant T, \tag{2.7}$$

satisfies (2.4) therefore X_t is also a Markov process. Indeed, for $t_1<t_2<t_3$ from $(0,T)$ we have

$$\begin{aligned}K(t_1,t_3)K(t_2,t_2)&=\left(t_1-\frac{t_1t_3}{T}\right)\left(t_2-\frac{t_2^2}{T}\right)=t_1t_2\left(1-\frac{t_3}{T}\right)\left(1-\frac{t_2}{T}\right)\\&=\left(t_1-\frac{t_1t_2}{T}\right)\left(t_2-\frac{t_2t_3}{T}\right)=K(t_1,t_2)K(t_2,t_3).\end{aligned}$$

Increments of X_t are not independent but they are stationary, i.e.

$$\operatorname{cov}(X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}) = \frac{(t_2 - t_1)(t_2 - t_3)}{T} < 0,$$

for $t_1 < t_2 < t_3$. As for Lemma 2.3 we have

$$K(t_1, t_2) = \begin{cases} t_1 \left(1 - \frac{t_2}{T}\right), & t_1 \leqslant t_2, \\ t_2 \left(1 - \frac{t_1}{T}\right), & t_2 < t_1, \end{cases}$$

hence $p(t) = t$, $q(t) = 1 - \frac{t}{T}$, $t \in (0, T)$ in (2.5). The function

$$\frac{p(t)}{q(t)} = \frac{tT}{T - t} = \frac{T^2}{T - t} - T, \quad t \in (0, T),$$

is strictly increasing.

3. The Ornstein-Uhlenbeck process and fractional Brownian motion

The stochastic process X_t, $t \in T$, is self-similar with index $H > 0$, if for every $a > 0$ the processes X_{at} and $a^H X_t$ have the same finite dimensional distributions. That means that for any $n \geqslant 1$, $t_1, \ldots, t_n \in T$ and $a > 0$ the distribution of $(X_{at_1}, \ldots, X_{at_n})$ is the same as $(a^H X_{t_1}, \ldots, a^H X_{t_n})$. Since we multiply $t \in T$ by any $a > 0$ it makes sense to take $T = \mathbb{R}$, $T = (0, +\infty)$ or $T = \langle 0, +\infty)$.

Lemma 3.1 (Proposition 7.1.4, [6]). *If X_t, $t > 0$, is self similar with index H, then*

$$Y_t = e^{-tH} X_{e^t}, \quad t \in \mathbb{R},$$

is stationary. Conversely, if Y_t, $t \in \mathbb{R}$ is stationary, then

$$X_t = t^H Y_{\ln t}, \quad t > 0,$$

is self similar with index H.

Observe that if $a > 0$ then $\mathbb{E}(W_{at_1} W_{at_2}) = \min\{at_1, at_2\} = a \min\{t_1, t_2\}$ and $\mathbb{E}(\sqrt{a} W_{t_1} \sqrt{a} W_{t_2}) = a \min\{t_1, t_2\}$. Hence, because Brownian motion is a Gaussian process it is self similar with $H = 1/2$.

Example 3.2. According to Lemma 3.1 the stochastic process

$$Y_t = e^{-t/2} W_{e^t}, \quad t \in \mathbb{R}, \tag{3.1}$$

is stationary with $m(t) = 0$ and

$$K(t_1, t_2) = e^{-\frac{1}{2}|t_1 - t_2|}, \quad t_1, t_2 \in \mathbb{R}. \tag{3.2}$$

It is also a Markov process since Y_t is Gaussian and

$$K(t_1, t_3)K(t_2, t_2) = e^{\frac{1}{2}(t_3 - t_1)} = K(t_1, t_2)K(t_2, t_3) = e^{\frac{1}{2}(t_2 - t_1)}e^{\frac{1}{2}(t_3 - t_2)},$$

where $t_1 < t_2 < t_3$. One can also investigate Y_t only for $t \geqslant 0$. We will construct the Ornstein-Uhlenbeck process in a different way. Since $Y_0 \sim N(0, 1)$ define

$$\widetilde{Y}_t = e^{-\frac{1}{2}t}X_0 + e^{-\frac{1}{2}t}\int_0^t e^{\frac{1}{2}s}dW_s, \quad t \geqslant 0, \tag{3.3}$$

where $X_0 \sim N(0, 1)$ is independent of the Brownian motion. Therefore $\widetilde{Y}_t$ is a Gaussian process with zero mean and the covariance function (3.2). Using the integration by parts formula, see (3.6) below, we can write $\widetilde{Y}_t$ as follows

$$\widetilde{Y}_t = e^{-\frac{1}{2}t}X_0 + W_t - \tfrac{1}{2}e^{-\frac{1}{2}t}\int_0^t e^{\frac{1}{2}s}W_s ds, \quad t \geqslant 0.$$

From the above we obtain

$$\begin{aligned}
d\widetilde{Y}_t &= -\tfrac{1}{2}e^{-\frac{1}{2}t}X_0 dt + dW_t - \tfrac{1}{2}\left[-\tfrac{1}{2}e^{-\frac{1}{2}t}\int_0^t e^{\frac{1}{2}s}W_s ds + W_t\right] dt \\
&= -\tfrac{1}{2}\left[e^{-\frac{1}{2}t}X_0 + W_t - \tfrac{1}{2}e^{-\frac{1}{2}t}\int_0^t e^{\frac{1}{2}s}W_s ds\right] dt + dW_t \\
&= -\tfrac{1}{2}\widetilde{Y}_t dt + dW_t.
\end{aligned}$$

Hence $\widetilde{Y}_t$ is the solution of $d\widetilde{Y}_t = -\frac{1}{2}\widetilde{Y}_t dt + dW_t$ with $\widetilde{Y}_0 = X_0$.

Example 3.3. One can prove, see Lemma 2.10.8 in [6], that the function

$$K_H(t_1, t_2) = \frac{1}{2}\left(|t_1|^{2H} + |t_2|^{2H} - |t_1 - t_2|^{2H}\right), \quad t_1, t_2 \in \mathbb{R},$$

satisfies (2.6) for $H \in (0, 1)$. Therefore there exists a Gaussian process X_t, $t \geqslant 0$, with zero mean and the covariance function $K_H(t_1, t_2)$. This process is called the fractional Brownian motion. For $H = 1/2$ it is in fact a Brownian motion, since

$$K_{\frac{1}{2}}(t_1, t_2) = \frac{t_1 + t_2 - |t_1 - t_2|}{2} = \min\{t_1, t_2\}.$$

For $H \neq \frac{1}{2}$ the increments of X_t are not independent. Indeed, we have

$$\operatorname{cov}(X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}) = \tfrac{1}{2}\left[(t_3 - t_1)^{2H} - (t_3 - t_2)^{2H} - (t_2 - t_1)^{2H}\right],$$

for $0 < t_1 < t_2 < t_3$. Define for fixed $t_1 < t_2$ the function

$$f_H(t) = (t - t_1)^{2H} - (t - t_2)^{2H} - (t_2 - t_1)^2, \quad t \geqslant t_2.$$

Note that $f(t_2) = 0$ and

$$f'_H(t) = 2H\left[(t - t_1)^{2H-1} - (t - t_2)^{2H-1}\right], \quad t > t_2.$$

If $H > 1/2$ then $f'_H(t) > 0$ and f_H is strictly increasing. If $H < 1/2$ we have $f'_H(t) < 0$ and f_H is decreasing. From the above we conclude that

$$\operatorname{cov}(X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}) : \quad \begin{cases} > 0, & H \in (\frac{1}{2}, 1) \\ = 0, & H = \frac{1}{2} \\ < 0, & H \in (0, \frac{1}{2}). \end{cases}$$

Hence X_t is not a Markov process for $H \neq \frac{1}{2}$.

Example 3.4. Let

$$X_t = A\cos(t + \varphi), \quad t \in \mathbb{R}, \tag{3.4}$$

where A is a random variable with density $f(x) = xe^{-x^2/2}$, $x \geqslant 0$, φ is uniformly distributed on $\langle 0, 2\pi)$ and independent of A. We have $\mathbb{E}X_t = \mathbb{E}A \cdot \mathbb{E}\cos(t + \varphi) = 0$ since

$$\mathbb{E}\cos(t + \varphi) = \frac{1}{2\pi}\int_0^{2\pi} \cos(t + x)dx = 0, \quad t \in \mathbb{R}.$$

The covariance $K(t_1, t_2) = \cos(t_2 - t_1)$. Indeed, we have

$$\begin{aligned} K(t_1, t_2) &= \mathbb{E}X_sX_t = \mathbb{E}(A^2)\mathbb{E}[\cos(t_1 + \varphi)\cos(t_2 + \varphi)] \\ &= \mathbb{E}[\cos(t_2 - t_1) + \cos(t_1 + t_2 + 2\varphi)] = \cos(t_2 - t_1). \end{aligned}$$

The crucial observation is that X_t can be written in the form

$$X_t = X\cos t - Y\sin t, \quad t \in \mathbb{R},$$

where $X = A\cos\varphi$, $Y = A\sin\varphi$. First we will prove that both X and Y has a $N(0,1)$ distribution. Furthermore from $\mathbb{E}(XY) = \mathbb{E}(A^2)\mathbb{E}(\sin\varphi\cos\varphi) = 0$ we conclude that X, Y are independent. Hence for all $a_1, \ldots, a_n \in \mathbb{R}$ the random variable

$$\sum_{i=1}^n a_iX_{t_i} = X\sum_{i=1}^n a_i\cos(t_i) - Y\sum_{i=1}^n a_i\sin(t_i)$$

has a normal distribution $N(0, \sigma^2)$, where

$$\sigma^2 = (\sum_{i=1}^{n} a_i \cos t_i)^2 + (\sum_{i=1}^{n} a_i \sin t_i)^2.$$

That means that X_t is a Gaussian process and since

$$K(t_1 + h, t_2 + h) = K(t_1, t_2), \quad \forall\, h \in \mathbb{R}$$

it is also stationary. In particular $(X_{t_1}, \ldots, X_{t_n})$ has a normal distribution with zero mean and the covariance function $K(t_i, t_j) = \cos(t_i - t_j)$, $i, j = 1, \ldots, n$. What remains to show is that $X \sim N(0, 1)$. Observe that the density of $\cos\varphi$ is

$$f(x) = \frac{1}{\pi} \frac{1}{\sqrt{1 - x^2}}, \quad x \in (-1, 1). \tag{3.5}$$

Indeed, for $x \in (-1, 1)$ we have

$$F(x) = \mathbb{P}(\cos\varphi \leqslant x) = \frac{(2\pi - 2\arccos x)}{2\pi} = 1 - \frac{\arccos x}{\pi},$$

which proves (3.5). Since X is the product of two independent random variables (i.e. A and $\cos\varphi$) its density equals

$$\begin{aligned} g(z) &= \frac{1}{\pi} \int_{|z|}^{+\infty} \frac{x}{\sqrt{x^2 - z^2}} e^{-\frac{1}{2}x^2} dx = \frac{1}{\pi} \int_{|z|}^{+\infty} x\sqrt{x^2 - z^2} e^{-\frac{1}{2}x^2} dx \\ &= \frac{1}{2\pi} e^{-\frac{1}{2}z^2} \int_0^{+\infty} \sqrt{t} e^{-\frac{1}{2}t} dt = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}, \quad z \in \mathbb{R}. \end{aligned}$$

Hence $X \sim N(0, 1)$. Similarly we show that $Y \sim N(0, 1)$.

Remark 3.5. Let $Z = X \cdot Y$ and $X \neq 0$. If X and Y are independent then the density of Z is given by

$$g(z) = \int_{\mathbb{R}} f_1(x) f_2\left(\frac{z}{x}\right) \frac{1}{|x|} dx,$$

where f_1 is the density of X and f_2 the density of Y, see e.g. [5].

Example 3.6. Consider the following Gaussian processes

$$X_t = \int_0^t W_s ds, \quad Y_t = tW_t, \quad Z_t = \int_0^t s dW_s, \quad t \geqslant 0.$$

All of them are Gaussian since each is a linear transformation of W_t. For X_t we have $m_X(t) = \int_0^t \mathbb{E}W_s ds = 0$ and $K_X(t_1,t_2) = \frac{1}{3}t_1^3 + \frac{1}{2}t_1^2(t_2 - t_1)$, $t_1 < t_2$. Indeed from $K_X(t_1,t_2) = \int_0^{t_1}\int_0^{t_2} \mathbb{E}(W_{s_1}W_{s_2})ds_1 ds_2$ we get

$$\begin{aligned} K_X(t_1,t_2) &= \int_0^{t_1}\int_0^{s_1} s_2 ds_2 ds_1 + \int_0^{t_1}\int_0^{s_2} s_1 ds_1 ds_2 + \int_0^{t_1}\int_{t_1}^{t_2} s_1 ds_2 ds_1 \\ &= \tfrac{1}{6}t_1^3 + \tfrac{1}{6}t_1^3 + \tfrac{1}{2}t_1^2(t_2 - t_1). \end{aligned}$$

The function $K_X(t_1,t_2)$ does not satisfy (2.4), so X_t is not a Markov process. For example $K_X(1,3)K_X(2,2) = \frac{32}{9}$ and $K_X(1,2)K_X(2,3) = \frac{35}{9}$. Increments of X_t are not stationary because

$$\operatorname{cov}(X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}) = \tfrac{1}{2}(t_2 - t_1)(t_2 + t_1)(t_3 - t_2),$$

where $0 \leqslant t_1 < t_2 < t_3$. For Y_t we have $m_Y(t) = 0$ and the covariance function equals

$$K_Y(t_1,t_2) = t_1 t_2 \mathbb{E}(W_{t_1}W_{t_2}) = t_1 t_2 \min\{t_1,t_2\}.$$

For $t_1 < t_2 < t_3$ we have $K_Y(t_1,t_3)K_Y(t_2,t_2) = K_Y(t_1,t_2)K_Y(t_2,t_3)$ and both those terms are equal to $t_1^2 t_2^3 t_3$. Therefore Y_t is a Markov process. However the increments of Y_t are not independent. If $t_1 < t_2$ then $\mathbb{E}(Y_{t_2} - Y_{t_1})(Y_{t_1}) = t_1^2(t_2 - t_1)$. For Z_t we have $m_Z(t) = 0$ and $K_Z(t_1,t_2) = \frac{1}{3}\min\{t_1^3, t_2^3\}$. Note that Z_t has independent increments, i.e.

$$\mathbb{E}(Z_{t_3} - Z_{t_2})(Z_{t_2} - Z_{t_1}) = \tfrac{1}{3}(t_2^3 - t_1^3 - t_2^3 + t_1^3) = 0.$$

If $t_1 < t_2$ then $\mathbb{P}(Z_{t_2} \in B | Z_{t_1} = x_1) = \int_B p(t_1,x_1,t_2,x_2)dx_2$, where

$$p(t_1,x_1,t_2,x_2) = \left(\tfrac{2\pi}{3}(t_2^3 - t_1^3)\right)^{-1/2} \exp\left\{-\frac{3(x_2-x_1)^2}{2(t_2^3 - t_1^3)}\right\}, \quad x_2 \in \mathbb{R}.$$

Therefore Z_t is a non-homogeneous Markov process. In fact $Z_t = Y_t - X_t$. Using (3.6) we have

$$\int_0^t s dW_s = tW_t - \int_0^t W_s ds, \quad t \geqslant 0.$$

Remark 3.7. For stochastic integrals we have (see e.g. Theorem 4.5, [4]): if a deterministic function f is continuous and of bounded variation on $\langle 0,t\rangle$ then

$$\int_0^t f(s)dW_s = f(t)W_t - \int_0^t W_s df(s). \tag{3.6}$$

Table 1

process	W_t	$\|W_t\|$	W_t^2	$\int_0^t W_s ds$	tW_t	$\int_0^t s dW_s$	$e^{-t/2}W_{e^t}$
$m(t)$	0	$\sqrt{\frac{2t}{\pi}}$	t	0	0	0	0
$\mathrm{var} X_t$	t	$(1-\frac{2}{\pi})t$	$2t^2$	$\frac{1}{3}t^3$	t^3	$\frac{1}{3}t^3$	1
Gaussian	*yes*	*no*	*no*	*yes*	*yes*	*yes*	*yes*
Markov	*yes*	*yes*	*yes*	*no*	*yes*	*yes*	*yes*
stationary	*no*	*no*	*no*	*no*	*no*	*no*	*yes*

4. Applications of Ito formula

Suppose that X_t has the stochastic differential $dX_t = a(t)dt + b(t)dW_t$ and $f(t,x) \in C^2([0,+\infty) \times \mathbb{R})$. Then (see e.g. Theorem 4.2, [4])

$$\begin{aligned} df(t,X_t) = {} & \left[f_t'(t,X_t) + a(t)f_x'(t,X_t) + \tfrac{1}{2}b^2(t)f_{xx}''(t,X_t)\right] dt \\ & + f_x'(t,X_t)b(t)dW_t. \end{aligned}$$

We will apply the above formula for computing expectations.

Example 4.1. Let $dX_t = a(t)dt + b(t)dW_t$, where $a(t)$ and $b(t)$ are deterministic functions, i.e.

$$X_t = \int_0^t a(s)ds + \int_0^t b(s)dW_s, \quad t \geqslant 0.$$

We have $m(t) = \mathbb{E}X_t = \int_0^t a(s)ds$ since $\mathbb{E}\int_0^t b(s)dW_s = 0$. In a consequence

$$\mathbb{E}(X_t - m(t))^2 = \mathbb{E}\left(\int_0^t b(s)dW_s\right)\left(\int_0^t b(s')dW_{s'}\right) = \int_0^t b^2(s)ds,$$

where we used $dW_s dW_{s'} = \delta(s-s')ds$. Now take $f(t,x) = (x-m(t))^n$, $n \geqslant 2$. Then by Ito formula we have

$$\begin{aligned} df(t,X_t) = {} & \left[a(t)n(X_t - m(t))^{n-1} + \tfrac{1}{2}b(t)^2 n(n-1)(X_t-m(t))^{n-2}\right] dt \\ & + b(t)n(X_t - m(t))^{n-1}dW_t. \end{aligned}$$

Denote $Y_t = X_t - m(t)$, and by the fact that the expectation of a stochastic integral is zero we obtain

$$\mathbb{E}Y_t^n = n\int_0^t a(t)\mathbb{E}Y_s^{n-1}ds + \tfrac{1}{2}n(n-1)\int_0^t b^2(s)\mathbb{E}Y_s^{n-2}ds, \quad n \geqslant 2,$$

with $EY_t = 0$. For a Brownian motion we have

$$dW_t^n = \tfrac{1}{2}n(n-1)W_t^{n-2}dt + nW_t^{n-1}dW_t, \quad n \geqslant 2,$$

and in consequence we get $\mathbb{E}W_t^n = \frac{1}{2}n(n-1)\int_0^t \mathbb{E}W_s^{n-2}ds$, $n \geqslant 2$. This leads to

$$\mathbb{E}W_t^n = \begin{cases} 1 \cdot 3 \cdot 5 \cdot \ldots \cdot (2k-1)t, & n = 2k \\ 0, \quad n = 2k+1. \end{cases}$$

Example 4.2. Let $f(x) = \sin x$. Then by Ito formula we have

$$\sin W_t = -\tfrac{1}{2}\int_0^t \sin(W_s)ds + \int_0^t \cos(W_s)dW_s.$$

Denote $m(t) = \mathbb{E}(\sin W_t)$, $t \geqslant 0$. Then $m'(t) = -\frac{1}{2}m(t)$ with $m(0) = 0$. That gives $m(t) = 0$, $t \geqslant 0$. In a similar way we get $\mathbb{E}(\cos W_t) = \exp(-t/2)$, $t \geqslant 0$.

Now take $f(x) = \sin^2 x$. Then $f'(x) = \sin(2x)$, $f''(x) = 2\cos(2x)$ and

$$\sin^2 W_t = \tfrac{1}{2}(1 - \cos(2W_t)) = \int_0^t \cos(2W_s)ds + \int_0^t \sin(2W_s)dW_s.$$

Denote $n(t) = \mathbb{E}\cos(2W_t)$, $t \geqslant 0$. The function $n(t)$ satisfies $n'(t) = -2n(t)$ for $t > 0$ with $n(0) = 1$. That implies $n(t) = \exp(-2t)$, $t \geqslant 0$. Finally

$$\mathbb{E}\sin^2 W_t = \tfrac{1}{2}(1 + e^{-2t}), \quad \mathbb{E}\cos^2 W_t = \tfrac{1}{2}(1 - e^{-2t}), \quad t \geqslant 0.$$

5. Stochastic Equations

In this section we use the notation $X_t = (X_t^1, X_t^2)$ for 2-dimensional stochastic processes. The following is the Exercise 5.8 from [4].

Example 5.1. Consider the equation

$$\begin{cases} dX_t^1 = X_t^2 dt + \alpha dW_t^1 \\ dX_t^2 = X_t^1 dt + \beta dW_t^2 \end{cases} \tag{5.1}$$

where (W_t^1, W_t^2) is 2-dimensional Brownian motion. This means that W_t^1 and W_t^2 are independent standard Brownian motions. We can write (5.1) in the form

$$\begin{bmatrix} dX_t^1 \\ dX_t^2 \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}}_{=A} \begin{bmatrix} X_t^1 \\ X_t^2 \end{bmatrix} dt + \underbrace{\begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix}}_{=M} \begin{bmatrix} dW_t^1 \\ dW_t^2 \end{bmatrix}, \quad t \geqslant 0,$$

with given X_0^1 and X_0^2. We assume that $\alpha, \beta \in \mathbb{R}$. According to e.g. [2], [4] or [7] the solution of $dX_t = Adt + MdW_t$ is given by

$$X_t = e^{tA} X_0 + e^{tA} \int_0^t e^{-sA} M dW_s, \quad t \geqslant 0. \tag{5.2}$$

Using (6.5) we have

$$e^{tA} = \sinh(t) \cdot A + \cosh(t) \cdot I = \begin{bmatrix} \cosh(t) & \sinh(t) \\ \sinh(t) & \cosh(t) \end{bmatrix}, \quad t \in \mathbb{R}.$$

Taking $X_0^1 = 0$, $X_0^2 = 0$ we get

$$X_t^1 = \alpha \int_0^t \cosh(t-s) dW_s^1 + \beta \int_0^t \sinh(t-s) dW_s^2$$

and

$$X_t^2 = \alpha \int_0^t \sinh(t-s) dW_s^1 + \beta \int_0^t \cosh(t-s) dW_s^2.$$

With this initial condition X_t is not stationary because e.g.

$$\begin{aligned} \mathrm{var}(X_t^1) = \mathbb{E}(X_t^1)^2 &= \alpha^2 \int_0^t \cosh^2(t-s) ds + \beta^2 \int_0^t \sinh^2(t-s) ds \\ &= \tfrac{\alpha^2+\beta^2}{2} \sinh(t)\cosh(t) + \tfrac{1}{2}(\alpha^2 - \beta^2) t. \end{aligned}$$

In addition

$$\mathbb{E}(X_t^1 X_t^2) = (\alpha^2 + \beta^2) \int_0^t \cosh(t-s)\sinh(t-s) ds = \tfrac{a^2+b^2}{4}(\cosh(2t) - 1).$$

6. Appendix

We will find the formula for e^{At} where A is a matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}. \tag{6.1}$$

We assume that entries of A are real. First we will compute A^n, $n \geqslant 2$. Let $w(\lambda)$ be the characteristic polynomial of A, that is

$$w(\lambda) = \det \begin{bmatrix} a-\lambda & b \\ c & d-\lambda \end{bmatrix} = \lambda^2 - (a+d)\lambda + ad - bc. \tag{6.2}$$

The roots of $w(\lambda)$ are

$$\lambda_1 = r - p, \quad \lambda_2 = r + p,$$

where $r = (a+d)/2$, $p = \sqrt{\Delta}/2$ and $\Delta = (a+d)^2 - 4D$ with $D = ad - bc$. Assume first that $\Delta > 0$ and take $n \geqslant 3$. Then

$$\lambda^n = Q_n(\lambda)w(\lambda) + x_n\lambda + y_n, \tag{6.3}$$

where $Q_n(\lambda)$ is a polynomial and $x_n, y_n \in \mathbb{R}$. Then by the Cayley-Hamilton theorem $w(A) = \Theta$, where Θ is a zero matrix. We conclude that $A^n = x_nA + y_nI$. Solving $\lambda_1^n = x_n\lambda_1 + y_n$, $\lambda_2^n = x_n\lambda_2 + y_n$ we find $x_n = (\lambda_1^n - \lambda_2^n)/(\lambda_1 - \lambda_2)$ and $y_n = (\lambda_1^n\lambda_2 - \lambda_2^n\lambda_1)/(\lambda_2 - \lambda_1)$. In consequence

$$A^n = \frac{\lambda_2^n - \lambda_1^n}{2p} \cdot A + \frac{\lambda_1^n\lambda_2 - \lambda_2^n\lambda_1}{2p} \cdot I, \quad n \geqslant 0. \tag{6.4}$$

Hence we have

$$e^{At} = A\sum_{n=0}^{+\infty} \frac{\lambda_2^n - \lambda_1^n}{2p} \cdot \frac{t^n}{n!} + I\sum_{n=0}^{+\infty} \frac{\lambda_1^n\lambda_2 - \lambda_2^n\lambda_1}{2p} \cdot \frac{t^n}{n!}.$$

Note that

$$\sum_{n=0}^{+\infty}(\lambda_2^n - \lambda_1^n) \cdot \frac{t^n}{n!} = e^{\lambda_2 t} - e^{\lambda_1 t} = 2e^{rt}\sinh(pt)$$

and similarly

$$\sum_{n=0}^{+\infty}(\lambda_1^n\lambda_2 - \lambda_2^n\lambda_1) \cdot \frac{t^n}{n!} = 2e^{rt}\left[p\cosh(pt) - r\sinh(pt)\right].$$

Finally we have

$$e^{At} = e^{rt}\frac{\sinh(pt)}{p} \cdot A + e^{rt}\left[\cosh(pt) - \frac{r}{p}\sinh(pt)\right] \cdot I, \quad \Delta > 0. \tag{6.5}$$

Now consider the case when $\Delta = 0$. Then $w(\lambda) = (\lambda - r)^2$ and in order to compute x_n we differentiate both sides of (6.3)

$$n\lambda^{n-1} = (\lambda - r)^2Q_n'(\lambda) + 2(\lambda - r)Q_n(\lambda) + x_n.$$

Putting $\lambda = r$ to the above equation we find $x_n = \lambda r^{n-1}$ and from (6.3) we get $y_n = (1-n)r^n$. It is worth to mention that $A^n = nr^{n-1}A + (1-n)r^nI$, $n \geqslant 1$. We have

$$\begin{aligned} e^{At} &= I + At\sum_{n=1}^{\infty}\frac{(rt)^{n-1}}{(n-1)!} + I\left[\sum_{n=1}^{\infty}\frac{(rt)^n}{n!} - rt\sum_{n=1}^{\infty}\frac{(rt)^{n-1}}{(n-1)!}\right] \\ &= I + Ate^{rt} + I[(e^{rt} - 1) - rte^{rt}] \end{aligned}$$

and finally

$$e^{At} = te^{rt} \cdot A + (1 - rt)e^{rt} \cdot I, \quad \Delta = 0. \tag{6.6}$$

References

1. R.J. Adler, *An introduction to continuity, extrema, and related topics for general Gaussian processes*, Vol. 12, 1990.
2. I. Karatzas, S.E. Shreve, *Brownian motion and stochastic calculus*, Springer, 1991.
3. M.B. Marcus, J. Rosen, *Markov Processes, Gaussian Processes and Local Times*, Cambridge, 2006
4. B. Oksendal, *Stochastic Differential Equations*, Springer, 1995.
5. A. Pluci"nska, E. Pluci"nski, *Rachunek prawdopodobie"nstwa, Statystyka matematyczna, Procesy stochastyczne*, WNT, Warszawa, 2000.
6. G. Samorodnitsky, M.S. Taqqu, *Stable non-Gaussian Random Processes: Stochastic Models with Infinite Variance*, Chapman and Hall, 1994.
7. K. Sobczyk, *Stochastyczne r"ownania r"o"rniczkowe*, WNT, Warszawa, 1996.
8. Wentzell, *Wstep do teorii proces"ow stochastycznych*, PWN, Warszawa, 1980.

Problem of multistage, strictly positional games with delayed information

Witold Rzymowski[1], **Tomasz Warowny**[2]

Abstract

A differential game can be viewed as a certain limit of the appropriate sequence of discrete dynamic games. Proofs of the existence of saddle points in differential games usually involve analogous facts concerning discrete dynamical games [1], [2]. In majority of the articles about differential games, authors assume that each player has complete and up-to-date information about his own position and the position of his opponent. In this paper we deal with discrete positional, dynamic, games with a fixed duration and final pay-off functional. We assumed that each player receives information about the position of his opponent with a certain delay. Consequently, both players have to apply mixed, positional strategies. A problem concerning the existence of saddle points, if a special case of mixed, positional strategies is involved, is presented in the document. Pursuit games with the delay of information received by one player were investigated by L.A Petrosian in *Differential Games of Pursuit* [3]. Other differential games, with incomplete information, were considered by E. Dockner and R. Isaacs [4], [5]. The special case of a pursuit game with almost absolutely incomplete information was solved by M.I. Zelikin [6].

1. Pure Strategies

Game Description

For every set Z, a symbol 2^Z denotes a family of all subsets of the set Z, and, if $Z \neq \varnothing$, $\mathfrak{Fin}\, Z$ denotes a family of all non-empty and finite subsets of

[1]Lublin University of Technology, Fundamental of Technology Faculty, Department of Applied Mathematics, Nadbystrzycka 38D, 20-618 Lublin, Poland, *e-mail: w.rzymowski@pollub.pl*

[2]Lublin University of Technology, Management Faculty, Department of Quantitative Methods in Management, Nadbystrzycka 38D, 20-618 Lublin, Poland, *e-mail: t.warowny@pollub.pl*

Keywords: Game theory, Multistage games, Local and global strategies.

this set. For any non-empty sets A and B, A^B means a set of all functions

$$f : A \to B. \tag{1.1}$$

Symbols $\mathbb{N}$ and $\mathbb{Z}$ stand for sets of all natural and integer number respectively. We assume

$$\mathbb{N}_0 = \{0\} \cup \mathbb{N} = \{0, 1, 2, ...\}, \quad \mathbb{R}^n_+ = \prod_{k=1}^{n} [0, \infty) \tag{1.2}$$

and for all $t \in \mathbb{N}_0$

$$\mathbb{Z}_{t]} = \{s \in \mathbb{Z} : s \leqslant t\} \tag{1.3}$$

Trajectories and sets of availability

Two players participate in a game: **E** – maximizer and **P** – minimizer. Player **E** moves in a set $X \neq \varnothing$ and player **P** moves in a set $Y \neq \varnothing$. Game dynamics is generated with a multifunction pair

$$F : X \to 2^{\mathfrak{Fin}\, X}, \quad G : Y \to 2^{\mathfrak{Fin}\, Y} \tag{1.4}$$

At an instant $t \in \mathbb{N}_0$ the players **E** and **P** take positions $x \in X$ and $y \in Y$ respectively. At the following instant $t + 1$ the player **E** can move to any element of the set $F(x)$ and the player **P** can move to any element of the set $G(y)$.

The initial position is set to $(x_0, y_0) \in X \times Y$. A symbol $\mathcal{X}(x_0)$ denotes a set of all functions (sequences) $\xi \in X^{\mathbb{N}_0}$ meeting the criterion:

$$\xi(t) \in F(\xi(t-1)), \quad t \in \mathbb{N}. \tag{1.5}$$

Similarly, a symbol $\mathcal{Y}(y_0)$ denotes a set of all functions (sequences) $\eta \in Y^{\mathbb{N}_0}$ meeting the criterion:

$$\eta(t) \in G(\eta(t-1)), \quad t \in \mathbb{N}. \tag{1.6}$$

The set $\mathcal{X}(x_0)$ is called the set of admissible trajectories of the player **E** and the set $\mathcal{Y}(y_0)$ is called the set of admissible trajectories of the player **P**. For any $(\xi, \eta) \in \mathcal{X}(x_0) \times \mathcal{Y}(y_0)$ and all $\tau \in \mathbb{Z}$ we define

$$\xi_{\tau]}(t) = \xi(t), \eta_{\tau]}(t) = \eta(t), t \in (Z)_{\tau]} \tag{1.7}$$

and assume

$$\mathcal{X}_{\tau]}(x_0) = \{\xi_{\tau]} : \xi \in \mathcal{X}(x_0)\},\ \mathcal{Y}_{\tau]}(y_0) = \{\eta_{\tau]} : \eta \in \mathcal{Y}(y_0)\}. \tag{1.8}$$

We establish a random initial position $(x_0, y_0) \in X \times Y$. For all $t \in \mathbb{Z}$ we define:

$$X_t(x_0) = \{x_0\}, \quad Y_t(y_0) = \{y_0\}, \text{ if } t \leqslant 0, \tag{1.9}$$

$$X_t(x_0) = F(X_{t-1}(x_0)), \quad T_t(y_0) = G(Y_{t-1}(y_0)), \text{ if } t > 0, \tag{1.10}$$

$$\mathbf{X}_{t\rceil}(x_0) = \bigcup_{s \leqslant t} X_t(x_0), \quad \mathbf{Y}_{t\rceil}(y_0) = \bigcup_{s \leqslant y} Y_t(y_0). \tag{1.11}$$

Delay of information

We presume that, at every instant $t \in \mathbb{N}_0$, each of the players **E** and **P** knows the initial position (x_0, y_0) and their current position $x_t \in X_t(x_0)$ and $y_t \in Y_t(y_0)$ consecutively. Additionally, at the instant t, player **E** receives the information about the position of the player **P** with a delay α, hence they know the position $y_{t-a} \in Y_{t-\alpha}(y_0)$ taken by **P** at $t-\alpha$. Likewise, at the instant t, player **P** receives the information about the position of the player **E** with the delay β, hence they know the position $x_{t-\beta} \in X_{t-\beta}$ taken by **E** at $t - \beta$. On the basis of these pieces of information both players choose their next position simultaneously.

$$\begin{aligned} x_{t+1} &\in F(x_1) \text{ (player } \mathbf{E}), \\ y_{t+1} &\in G(y_t) \text{ (player } \mathbf{P}). \end{aligned} \tag{1.12}$$

Strictly positional, pure strategies

Let us specify a random initial position $(x_0, y_0) \in X \times Y$, delays $\alpha, \beta \in \mathbb{N}_0$ and game duration $N \in \mathbb{N}$. Every function $f : \mathbf{X}_{N-1\rceil}(x_0) \times \mathbf{Y}_{N-1\rceil}(y_0) \to X$ fulfilling a condition:

$$(x, y) \in \mathbf{X}_{N-1\rceil}(x_0) \times \mathbf{Y}_{N-1\rceil}(y_0) \implies f(x, y) \in F(x) \tag{1.13}$$

is called a strictly positional, pure strategy of the player **E**. The set of all such strategies is denoted $S^{spp}_{E-\alpha}(x_0, y_0)$. Every function $g : \mathbf{X}_{N-1\rceil}(x_0) \times \mathbf{Y}_{N-1\rceil}(y_0) \to Y$ fulfilling a condition:

$$(x, y) \in \mathbf{X}_{N-1\rceil}(x_0) \times \mathbf{Y}_{N-1\rceil}(y_0) \implies g(x, y) \in G(y) \tag{1.14}$$

is called a strictly positional, pure strategy of the player **P**. The set of all such strategies is denoted $S^{spp}_{P-\beta}(x_0, y_0)$

Note 1.1. For any $(f, g) \in S^{spp}_{E-\alpha}(x_0, y_0) \times S^{spp}_{P-\beta}(x_0, y_0)$, there exist exactly one pair

$$(\xi^{(f,g)}, \eta^{(f,g)}) \in \mathcal{X}_{N\rceil}(x_0) \times \mathcal{Y}_{N\rceil}(y_0) \tag{1.15}$$

meeting a condition

$$(\xi^{(f,g)}, \eta^{(f,g)}) = (f(\eta^{(f,g)}), g(\xi^{(f,g)})) \tag{1.16}$$

We say that the pair of trajectories $(\xi^{(f,g)}, \eta^{(f,g)})$ is determined (generated) by the pair of strategies (f, g).

Game

The following conditions are given: $N \in \mathbb{N}$, the initial position (x_0, y_0) and the function $w : X \times Y \to \mathbb{R}$. For any pair of strategies $(f, g) \in S^{spp}_{E-\alpha}(x_0, y_0) \times S^{spp}_{P-\beta}(x_0, y_0)$ we maintain

$$w(\varphi, \upsilon) = w(\xi^{(f,g)}(N), \eta^{(f,g))}(N)). \tag{1.17}$$

The functional defined as $w : S^{spp}_{E-\alpha}(x_0, y_0) \times S^{spp}_{P-\beta}(x_0, y_0) \to \mathbb{R}$ is called a final pay-off functional. The arrangement

$$G^{spp}_{\alpha,\beta}(x_0, y_0, N, w) = (S^{spp}_{E-\alpha}(x_0, y_0) \times S^{spp}_{P-\beta}(x_0, y_0), N, w) \tag{1.18}$$

is called a strictly positional, multistage game (N-stage game) with the strictly positional, pure strategies, final pay-off functional w, initial position (x_0, y_0) and the delay of information – α for the player **E** and β for the player **P**.

Note 1.2. Every game $G^{spp}_{\alpha,\beta}(x_0, y_0, 1, w)$ is a matrix game.

Upper and lower value of a game

A strategy $f \in S^{spp}_{E-\alpha}(x_0, y_0)$ in the game $G^{spp}_{\alpha,\beta}(x_0, y_0, N, w)$ provides a result u to the player **E** if

$$w_{\mathbf{E}}(f) \stackrel{\text{def}}{=} \min_{g \in S^{spp}_{P-\beta}(x_0,y_0)} w(f, g) \geqslant u. \tag{1.19}$$

In such a case, we can write $f \succeq u$. A number

$$\begin{aligned} V^- &= V^-(S^{spp}_{E-\alpha}(x_0, y_0), S^{spp}_{P-\beta}(x_0, y_0), N, w) \\ &= \max_{f \in S^{spp}_{E-\alpha}(x_0,y_0)} w_E(f) \\ &= \max_{f \in S^{spp}_{E-\alpha}(x_0,y_0)} \min_{g \in S^{spp}_{P-\beta}(x_0,y_0)} w(f, g) \end{aligned} \tag{1.20}$$

is called the lower value of the game $G^{spp}_{\alpha,\beta}(x_0, y_0, N, w)$. The lower value V^- is the maximum result which may be obtained by the player **E** in a considered game. Since $S^{spp}_{E-\alpha}(x_0, y_0)$ is a finite set, then there exists such $f^* \in S^{spp}_{E-\alpha}(x_0, y_0)$ meeting $f^* \succeq V^-$.

A strategy $g \in S^{spp}_{P-\beta}(x_0, y_0)$ provides the player **P** the result v in the game $G^{spp}_{\alpha,\beta}(x_0, y_0, N, w)$ if

$$w_P(g) \overset{\text{def}}{=} \max_{f \in S^{spp}_{E-\alpha}(x_0,y_0)} w(f,g) \leqslant v. \tag{1.21}$$

Then, we can write $g \preceq v$. A number

$$\begin{aligned} V^+ &= V^+(S^{spp}_{E-\alpha}(x_0, y_0), S^{spp}_{P-\beta}(x_0, y_0), N, w) \\ &= \min_{g \in SP-\beta^{spp}(x_0,y_0)} w_P(g) \\ &= \min_{g \in S^{spp}_{P-\beta}(x_0,y_0)} \max_{f \in S^{spp}_{E-\alpha}(x_0,y_0)} w(f,g) \end{aligned} \tag{1.22}$$

is called the upper value of the game $G^{spp}_{\alpha,\beta}(x_0, y_0, N, w)$. The upper value V^+ is the minimum result which may be obtained by the player **P** in a considered game. Since $S^{spp}_{P-\beta}(x_0, y_0)$ is a finite set, then there exist such $g^* \in S^{spp}_{P-\beta}(x_0, y_0)$ meeting $g^* \preceq Y^+$. Obviously, $V^- \leqslant V^+$.

Value of the game

If $V^- = V^+$, then the number

$$V = V(S^{spp}_{E-\alpha}(x_0, y_0), S^{spp}_{P-\beta}(x_0, y_0), N, w) = V^- = V^+ \tag{1.23}$$

is called a value of the game $G^{spp}_{\alpha,\beta}(x_0, y_0, N, w)$. As both strategy sets are finite

$$V^- = V^+ \Leftrightarrow \exists (f^*, g^*) \in S^{spp}_{E-\alpha}(x_0, y_0) \times S^{spp}_{P-\beta}(x_0, y_0) \left(w_E(f^*) = w_P(g^*)\right). \tag{1.24}$$

The saddle point (solution of the game)

Every pair $(f^*, g^*) \in S^{spp}_{E-\alpha}(x_0, y_0) \times S^{spp}_{P-\beta}(x_0, y_0)$ meeting a condition

$$w_E(f^*) = w_P(g^*) \tag{1.25}$$

is called an equilibrium point or a solution of the game $G^{spp}_{\alpha,\beta}(x_0, y_0, N, w)$.

Note 1.3. The game $G^{spp}_{\alpha,\beta}(x_0, y_0, N, w)$ has, in most cases, no solution.

Example 1.4. Let $X = Y = \{0, 1, 2\}, N = 2$ and $F = G$, where $F : X \to 2^X$, be defined by the equation

x	0	1	2
$F(x)$	$\{0,1\}$	$\{0,2\}$	$\{1,2\}$

The pay-off functional $w : X \times Y \to \mathbb{R}$ is defined using a table:

				y
x	$w(x,y)$	**0**	**1**	**2**
	0	4	1	2
	1	1	1	3
	2	0	3	0

For the initial position $(x_0, y_0) = (0, 0)$ we get

$$\begin{aligned} X_0(x_0) &= Y_0(y_0) = \{0\}, \\ X_1(x_0) &= Y_1(y_0) = \{0, 1\}, \\ X_2(x_0) &= Y_2(y_0) = \{0, 1, 2\}, \\ \mathbf{X}_{N-1|}(x_0) &= \mathbf{Y}_{N-1|}(x_0) = \{0, 1\}, \\ \mathcal{X}_2(x_0) &= \mathcal{Y}_2(y_0) = \{\zeta_1, \zeta_2, \zeta_3, \zeta_4\} \end{aligned} \tag{1.26}$$

where

$$\zeta_1 = (0, 0, 0), \zeta_2 = (0, 0, 1), \zeta_3 = (0, 1, 0), \zeta_4 = (0, 1, 2). \tag{1.27}$$

Therefore, both players can move along four trajectories $\zeta_i, i = 1, 2, 3, 4$ only.

We are going to determine sets of strategies S^{spp}_{E-1}, S^{spp}_{P-0} for delays $\alpha = 1$, $\beta = 0$. Every strategy $f \in S^{spp}_{E-1}$ only depends on the first variable because $\alpha = N - 1$. As a result

$$S^{spp}_{E-1} = \{f_1, f_2, f_3, f_4\} \tag{1.28}$$

where

	$(0, y)$	$(1, y)$
f_1	0	0
f_2	0	2
f_3	1	0
f_4	1	2

The set S^{plp}_{P-0} comprises 16 strategies:

	(0,0)	(1,0)	(0,1)	(1,1)
g_1	0	0	0	0
g_2	0	0	0	2
g_3	0	0	2	0
g_4	0	0	2	2
g_5	0	1	0	0
g_6	0	1	0	2
g_7	0	1	2	0
g_8	0	1	2	2
	(0,0)	(1,0)	(0,1)	(1,1)
g_9	1	0	0	0
g_{10}	1	0	0	2
g_{11}	1	0	2	0
g_{12}	1	0	2	2
g_{13}	1	1	0	0
g_{14}	1	1	0	2
g_{15}	1	1	2	0
g_{16}	1	1	2	2

Notice that $g_{12} \preceq 2 \preceq f_1$, thus the pair (f_1, g_{12}) is a saddle point of a considered game $G^{spp}_{1,0}(x_0, y_0, 2, w)$. Indeed,

$$\xi^{(f_1,g_j)} = \zeta_1 \,, j = 1, 2, 3, ..., 16 \tag{1.29}$$

and

$$\eta^{(f_k,g_{12})} = \zeta_4 \,, k = 1, 2, 3, 4. \tag{1.30}$$

2. Mixed global strategies

Consider a game

$$G^{spp}_{\alpha,\beta}(x_0, y_0, N, w) = (S^{spp}_{E-\alpha}(x_0, y_0), S^{spp}_{P-\beta}(x_0, y_0), N, w) \tag{2.1}$$

and assume that

$$S_E = S^{spp}_{E-\alpha}(x_0, y_0) = \{f_1, f_2, ..., f_m\}, \tag{2.2}$$

$$S_P = S^{spp}_{P-\beta}(x_0, y_0) = \{g_1, g_2, ..., g_n\}. \tag{2.3}$$

The game $G^{spp}_{\alpha,\beta}(x_0, y_0, N, w)$ is equivalent to a matrix game $\mathbf{G}(S_E, S_P, M)$ with sets of pure strategies S_E, S_P and a pay-off matrix

$$M = [w(f_j, g_k)]_{\substack{j=1,2,...,m \\ k=1,2,...,n}} \tag{2.4}$$

Further, we assume:

$$P = \left\{ p = (p_1, p_2, ..., p_m) \in \mathbb{R}^m_+ : \sum_{j=1}^{m} p_j = 1 \right\}, \tag{2.5}$$

$$Q = \left\{ q = (q_1, q_2, ..., q_n) \in \mathbb{R}^n_+ : \sum_{k=1}^{n} q_k = 1 \right\} \tag{2.6}$$

and, for any $(p, q) \in P \times Q$ we define:

$$\mathbf{w}(p, q) = \sum_{j-1}^{m} \sum_{k=1}^{n} p_j q_k w(f_i, g_k). \tag{2.7}$$

The matrix game $\mathbf{G}(P, Q, \mathbf{w})$ with mixed strategy sets P, Q and pay-off functional $\mathbf{w}$ has a saddle point (p^*, q^*) and a value $\mathbf{V}^* = \mathbf{w}(p^*, q^*)$ [7].

As far as the game $G^{spp}_{\alpha,\beta}(x_0, y_0, N, w)$ is concerned, we presume:

$$S^{spgm}_{E-\alpha}(x_0.y_0) = P, \tag{2.8}$$

$$S^{spgm}_{P-\beta}(x_0, y_0) = Q, \tag{2.9}$$

$$\mathbf{w}(p, q) = \sum_{j=1}^{m} \sum_{k=1}^{n} p_j q_k w(\xi^{(f_i, g_k)}, \eta^{f_i, g_k)}). \tag{2.10}$$

Therefore, we obtain the game

$$G^{spgm}_{\alpha,\beta}(x_0, y_0, N, \mathbf{w}) = (S^{spgm}_{E-\alpha}(x_0, y_0) \times S^{spgm}_{P-\beta}(x_0.y_0), N, \mathbf{w}) \tag{2.11}$$

with a saddle point (p^*, q^*) and a value $\mathbf{V}^* = \mathbf{w}(p^*, q^*)$.

In case the initial game $G^{spp}_{\alpha,\beta}(x_0.y_0.N, w)$ is played multiple times, strategies $p \in S^{spmg}_{E-\alpha}(x_0, y_0)$ and $q \in S^{spmg}_{P-\beta}(x_0, y_0)$ can be interpreted as the frequency of use of the pure strategies.

Example 2.1. A pair (f_1, g_{12}) is a saddle point of a game $G^{spp}_{1,0}(x_0, y_0, 2, w)$ from the example 1.4. Thus, the pair (p^*, q^*) where

$$p^* = (1, 0, 0, 0), q^* = (0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0) \tag{2.12}$$

is a saddle point of the game $G^{spp}_{1,0}(x_0, y_0, 2, w)$. Clearly, the values of the games G^{spp} and G^{spmg} are the same and equal to 2.

3. Mixed local strategies

There exist another, more natural, way to derive mixed strategies. Once again, consider the game $G^{spp}_{\alpha,\beta}(x_0, y_0, N, w)$. Each pair

$$(x, y) \in \mathbf{X}_{N-1]}(x_0) \times \mathbf{Y}_{N-1]}(y_0) \tag{3.1}$$

is assigned a probability distribution $\mathfrak{D}_E(x, y)$ and $\mathfrak{D}_P(x, y)$. Assume that

$$F(x) = \{x_1, x_2, ..., x_m\}, \quad G(y) = \{y_1, y_2, ..., y_n\}, \tag{3.2}$$

$$\mathcal{D}_E(x, y) = \left\{ p = (p_1, p_2, ..., p_m) \in \mathbb{R}^m_+ : \sum_{j=1}^{m} p_j = 1 \right\}, \tag{3.3}$$

$$\mathcal{D}_P(x, y) = \left\{ q = (q_1, q_2, ..., q_n) \in \mathbb{R}^n_+ : \sum_{k=1}^{n} q_k = 1 \right\}. \tag{3.4}$$

The set $S^{spml}_{E-\alpha}(x_0, y_0)$ comprising all functions:

$$f : \mathbf{X}_{N-1]}(x_0) \times \mathbf{Y}_{N-1]}(y_0) \to \bigcup_{(x,y)\in \mathbf{X}_{N-1]}(x_0)\times \mathbf{Y}_{N-1]}(y_0)} \mathfrak{D}_E(x, y) \tag{3.5}$$

meeting the criterion:

$$(x, y) \in \mathbf{X}_{N-1]}(x_0) \times \mathbf{Y}_{N-1]}(y_0) \implies f(x, y) \in \mathfrak{D}_E(x, y) \tag{3.6}$$

is the set of the strictly positional, mixed, local strategies of the player **E**.

The set $S^{spml}_{P-\beta}(x_0, y_0)$ comprising all functions:

$$g : \mathbf{X}_{N-1]}(x_0) \times \mathbf{Y}_{N-1]}(y_0) \to \bigcup_{(x,y)\in \mathbf{X}_{N-1]}(x_0)\times \mathbf{Y}_{N-1]}(y_0)} \mathfrak{D}_P(x, y) \tag{3.7}$$

meeting the criterion

$$(x, y) \in \mathbf{X}_{N-1]}(x_0) \times \mathbf{Y}_{N-1]}(y_0) \implies g(x, y) \in \mathfrak{D}_P(x, y) \tag{3.8}$$

is the set of the strictly positional, mixed, local strategies of the player **P**.

The interpretation of the operation of the strategy $f \in S^{spml}_{E-\alpha}(x_0, y_0)$ is as follows: Suppose that, at the instant $t \in \{0, 1, ..., N - 1\}$, the player **E** is located at the point x, and, at the instant $t - \alpha$, the player **P** was located at the point y. Assume that

$$F(x) = \{x_1, x_2, ..., x_m\}, f(x, y) = (p_1, p_2, ..., p_m). \tag{3.9}$$

At the instant $t+1$ the player **E** will be located at the point x_j with a probability p_j. The interpretation of how the strategy $g \in S^{spml}_{P-\beta}(x_0, y_0)$ works is similar.

Each pair of strategies $(f, g) \in S^{spml}_{E-\alpha}(x_0, y_0) \times S^{spml}_{P-\beta}(x_0, y_0)$ generated a distribution

$$f \otimes g : \mathcal{X}_{N\rceil}(x_0) \times \mathcal{Y}_{N\rceil}(y_0) \to [0,1], \qquad \sum_{(\xi,\eta)\in \mathcal{X}_{N\rceil}(x_0)\times \mathcal{Y}_{N\rceil}(y_0)} (f \otimes g)_{(\xi,\eta)} = 1 \quad (3.10)$$

in a set $\mathcal{X}_{N\rceil}(x_0) \times \mathcal{Y}_{N\rceil}(y_0)$. Each such a pair of strategies is assigned an expected pay-off amount

$$w(f,g) = \sum_{(\xi,\eta)\in \mathcal{X}_{N\rceil}(x_0)\times \mathcal{Y}_{N\rceil}(y_0)} (f \otimes g)_{(\xi,\eta)} w(\xi(N), \eta(N)). \qquad (3.11)$$

leading to the game $G^{spml}_{\alpha,\beta}(x_0, y_0, N, w) = (S^{spml}_{E-\alpha}(x_0, y_0) \times S^{spml}_{P-\beta}(x_0, y_0), N, w)$.

Note 3.1. Games

$$G^{spml}_{\alpha,\beta}(x_0, y_0, N, w), \quad G^{spml}_{\alpha,\beta}(x_0, y_0, N, w) \qquad (3.12)$$

are not equivalent. The existence of a saddle point in the game

$$G^{spml}_{\alpha,\beta}(x_0, y_0, N, w)$$

is an open problem. The following example is a sign of this statement.

Example 3.2. In a game $G^{spp}_{1,0}(x_0, y_0, N, w)$ from the example 1.4, we use a strictly positional, mixed, local strategy. In the other words, we consider a game $G^{spml}_{1,0}(x_0, y_0, 2, w)$.

Local distribution of the player **E** is defined as follows: we choose random

$$p = (p_1, p_2) \in [0,1] \times [0,1] \qquad (3.13)$$

and assume:

$$f_p(0,0) = f_p(0,1) : \begin{array}{|c|c|}\hline 0 & 1 \\ \hline p_1 & 1-p_1 \\ \hline\end{array}, \quad f_p(1,0) = f_p(1,1) : \begin{array}{|c|c|}\hline 0 & 2 \\ \hline p_2 & 1-p_2 \\ \hline\end{array} \qquad (3.14)$$

Local distribution of the player **P** is defined in a following manner: we choose random

$$q = (q_1, q_2, q_3, q_4) \in [0,1] \times [0,1] \times [0,1] \times [0,1] \qquad (3.15)$$

and assume:

$$
\begin{aligned}
g_q(0,0) &: \begin{array}{|c|c|}\hline 0 & 1 \\ \hline q_1 & 1-q_1 \\ \hline \end{array}, \quad g_q(1,0) : \begin{array}{|c|c|}\hline 0 & 1 \\ \hline q_2 & 1-q_2 \\ \hline \end{array}, \\
g_q(0,1) &: \begin{array}{|c|c|}\hline 0 & 2 \\ \hline q_3 & 1-q_3 \\ \hline \end{array}, \quad g_q(1,1) : \begin{array}{|c|c|}\hline 0 & 2 \\ \hline q_4 & 1-q_4 \\ \hline \end{array}.
\end{aligned}
\tag{3.16}
$$

A distribution $f_p \otimes g_q$ inside the set $\{\xi_1, \xi_2, \xi_3, \xi_4\} \times \{\eta_1, \eta_2, \eta_3, \eta_4\}$ is given by

$f_p \otimes g_q$	η_1	η_2
ξ_1	$p_1^2 q_1^2$	$p_1^2 q_1(1-q_1)$
ξ_2	$p_1(1-p_1)q_1^2$	$p_1(1-p_1)q_1(1-q_1)$
ξ_3	$(1-p_1)p_2 q_1 q_2$	$(1-p_1)p_2 q_1(1-q_2)$
ξ_4	$(1-p_1)(1-p_2)q_1 q_2$	$(1-p_1)(1-p_2)q_1(1-q_2)$

$f_p \otimes g_q$	η_3	η_4
ξ_1	$p_1^2(1-q_1)q_3$	$p_q^2(1-q_1)(1-q_3)$
ξ_2	$p_1(1-p_1)(1-q_1)q_3$	$p_1(1-p_1)(1-q_1)(1-q_3)$
ξ_3	$(1-p_1)p_2(1-q_1)q_4$	$(1-p_1)p_2(1-q_1)(1-q_4)$
ξ_4	$(1-p_1)(1-p_2)(1-q_1)q_4$	$(1-p_1)(1-p_2)(1-q_1)(1-q_4)$

A pay-off functional is expressed by the formula:

$$
\begin{aligned}
w(p,q) &= w(p_1, p_2, q_1, q_2, q_3) = w(f_p, g_q) \\
&= p_1^2(4q_1^2 + q_1(1-q_1) + 4(1-q_1)q_3 + 2(1-q_1)(1-q_3)) \\
&\quad + p_1(1-p_1)(q_1^2 + q_1(1-q_1) + (1-q_1)q_3 + 3(1-q_1)(1-q_3)) \\
&\quad + (1-p_1)p_2(4q_1q_2 + q_1(1-q_2) + 4(1-q_1)q_4 + 2(1-q_1)(1-q_4)) \\
&\quad + (1-p_1)(1-p_2)(0 + 3q_1(1-q_2) + 0 + 0).
\end{aligned}
\tag{3.17}
$$

Regarding a saddle point, it is easy to prove that a function $w(\cdot, q)$ does not have to be concave and a function $w(p, \cdot)$ does not have to be convex. Therefore, the existence of a saddle point in the considered game is not a result of a standard minimax theorems. Assume that

$$
p = (p_1, 1), \quad q = (q_1, 0, 0, 0), \tag{3.18}
$$

then

$$
\begin{aligned}
w(p,q) &= p_1^2(3q_1^2 - q_1 + 2) + p_1(1-p_1)(3-2q_1) + (1-p_1)(2-q_1) \\
&= p_1^2(3q_1^2 + q_1 - 1) + p(1-q_1) - q_1 + 2.
\end{aligned}
\tag{3.19}
$$

Note that:

$$3q_1^2 + q_1 - 1 < 0 \Leftrightarrow q_1 \in \left[0, \frac{\sqrt{13}-1}{6}\right) \tag{3.20}$$

We set $q_0 = \frac{\sqrt{13}-1}{6} \approx 0.434258$.

If $q_1 \in [q_0, 1]$, then

$$\begin{aligned}\max_{p_1\in[0,1]} w(p_1, q_1) &= \max_{p_1\in\{0,1\}} w(p_1, q_1) = \max\{2 - q_1, 3q_1^2 - q_1 + 2\}\\ &= 3q_1^2 - q_1 + 2\end{aligned} \tag{3.21}$$

and

$$\begin{aligned}\min_{q_1\in[q_0,1]} w(p_1, q_1) &= \min_{q_1\in[q_0,1]} (3q_1^2 - q_1 + 2) = 3q_0^2 - q_0 + 2\\ &= \frac{10-\sqrt{13}}{3} \approx 2.131483.\end{aligned} \tag{3.22}$$

Note that $\min_{q_1\in[0,1]} (3q_q^2 - q_1 + 2) = 2$ for $q_1 = \frac{1}{3}$

If $q_1 \in [0, \frac{1}{3})$ then $0 < \frac{q_1-1}{2(3q_1^2-q_1-1} < 1$, hence:

$$\begin{aligned}\max_{p_1\in[0,1]} w(p_1, q_1) &= w\left(\frac{q_1-1}{2(3q_1^2+q_1-1)}, q_1\right)\\ &= \frac{12q_1^3 - 19q_1^2 - 14q_1 + 9}{4(1 - q_1 - 3q_1^2)} = h(q_1)\end{aligned} \tag{3.23}$$

We have

$$\frac{\mathrm{d}}{\mathrm{d}q_1} h(q_1) = \frac{-36q_1^4 - 24q_1^3 + 13q_1^2 + 16q_1 - 5}{4(1 - q_1 - 3q_1^2)^2}. \tag{3.24}$$

The only root of the polynomial $-36q_1^4 - 24q_1^3 + 13q_1^2 + 16q_1 - 5$ within the range $[0, \frac{1}{3})$ is $q_1^* \approx 0.297\,747\,3$. Therefore,

$$\min_{q_1\in[0,\frac{1}{3})} w(p_1, q_1) = h(q_1^*) \approx 1.984\,838 \tag{3.25}$$

If $q_1 \in \left[\frac{1}{3}, q_0\right)$ then:

$$\frac{q_1-1}{2(3q_1^2+q_1-1)} \geqslant 1, \tag{3.26}$$

thus

$$\max_{p_1\in[0,1]} w(p_1, q_1) = w(1, q_1) = 3q_1^2 - q_1 + 2 \tag{3.27}$$

and

$$\min_{q_1 \in [\frac{1}{3}, q_0]} w(p_1, q_1) = 2 \text{ , for } \quad q_1 = \frac{1}{3} \tag{3.28}$$

Finally, the upper value of the game is given by

$$V^+(q_1^*) \approx 1.984\,838 \tag{3.29}$$

and

$$q_1^* \approx 0.297\,747\,3. \tag{3.30}$$

The lower value of the game can be determined in a following way. Suppose that

$$w(p_1, q_1) = 3p_1^2 q_1^2 + (p_1^2 - p_1 - 1)q_1 - p_1^2 + p_1 + 2 \tag{3.31}$$

We may be notice that

$$p_1^2 - p_1 - 1 < 0 \text{ when } p_1 \in [0, 1] \tag{3.32}$$

and

$$0 < \frac{1 - p_1 - p_1^2}{6p_1^2} < 1 \Leftrightarrow p_1 \in \left(\frac{\sqrt{29}+1}{14}, 1\right]. \tag{3.33}$$

Suppose $p_0 = \frac{\sqrt{29}+1}{14} \approx 0.456\,083\,1$. If $p_1 \in [0, p_0]$, then

$$\min_{q_1 \in [0,1]} w(p_1, q_1) = w(p_1, 1) = 3p_1^2 + 1. \tag{3.34}$$

and

$$\max_{p_1 \in [0, p_0]} (3p_1^2 + 1) = 3p_0^2 + 1 \approx 1.624\,035\,66 \tag{3.35}$$

If $p_1 \in (p_0, 1]$, then:

$$\begin{aligned} \min_{q_1 \in [0,1]} w(p_1, q_1) &= w\left(p_1, \frac{1 + p_1 - p_1^2}{6p_1^2}\right) \\ &= \frac{-13p_1^4 + 14p_1^3 + 25p_1^2 - 2p_1 - 1}{12p_1^2} = r(p_1). \end{aligned} \tag{3.36}$$

We have

$$\frac{\mathrm{d}}{\mathrm{d}p_1} r(p_1) = \frac{1 + p_1 + 7p_1^3 - 13p_1^4}{6p_1^3} \tag{3.37}$$

The only root of the polynomial $1 + p_1 + 7p_1^3 - 13p_1^4$ within the range $(p_0, 1]$ is $p_1^* \approx 0.804\,796$ and $g(p_1^*) \approx 1.984\,838$. Finally, the lower value of the game is given by:

$$V^-(p_1^*) \approx 1.984\,838 \tag{3.38}$$

and

$$p_1^* \approx 0.804\,796 \tag{3.39}$$

Notice (example 2.1) that this time games

$$G_{1,0}^{spp}(x_0, y_0, 2, w)$$

and

$$G_{1,0}^{spml}(x_0, y_0, 2, w)$$

have distinct saddle points and distinct values.

References

1. A. Friedman, *Differential Games*, Wiley, New Yourt, 1971.
2. N.N. Krasovskii, A.I. Subbotin, *Game-Theoretical Control Problems*, Springer-Verlag, Berlin 1988.
3. L.A. Petrosian, *Differential Games of Pursuit*, Izd. LGU, Leningrad 1977, pp. 172-186.
4. E. Dockner, S. Jørgensen, Ngo Van Long, G. Sorger, *Differential Games in Economics and Management Science*, Cambridge University Press, New York 2000. Chapter 8.
5. R. Isaacs *Differential games*, John Wiley and and Sons Inc., New York 1965, Chapter 12.
6. M.I.Zelikin, *On a certain differential game with incomplete information*, Game Theory, Izd. AN ASSR, Erevan 1973, pp 168-171.
7. J. von Neumann, O. Morgenstern, *The Theory of Games and Economics Behaviour*, Princeton University Press, Princeton 1944.

Identification of periodic components

Witold Rzymowski[1], **Agnieszka Surowiec**[2]

Abstract

The econometric model of life expectancy for men in Poland in the age from 1 to 100 is considered. The purpose of the investigation is identification of the periodic components in the sequences of the residuals of this model. In the process of identifications the following methods were used: the periodogram, the classical harmonic analysis and the modified classical harmonic analysis. The results of these methods are compared with the results of the deterministic models with the predetermined periodic components.

1. Introduction

Life expectancy is the average number of years that a group of individuals of age t is likely to live. In this work men are chosen for analysis.

Knowledge of this statistical measure of how long a person at a given age may live is very important. Thanks to the knowledge of this measure you can get a synthetic picture of the developments in the process of extinction of the population of a certain age. This knowledge may be used for example by policy makers. They usually consider life expectancy while deciding the retirement plan, e.g. retirement age, minimal number of years spent working, minimal pension etc. Therefore, life expectancy is one of the factors in determining the standard of living. Knowledge of life expectancy is necessary for insurance companies to determine the value of a life insurance policy. Disparities in life expectancy are often cited as demonstrating the need for better medical care or increased social support [9].

The present work is an attempt to find the econometric model with periodic components of life expectancy P_t for men in Poland in the age from $t = 1, ..., 100$ and it presents the main problems connected with the

[1]Department of Applied Mathematics, Faculty of Fundamentals of Technology, Lublin University of Technology, Nadbystrzycka 38D, 20-618 Lublin, Poland, *e-mail: w.rzymowski@pollub.pl*

[2]Department of Quantitative Methods in Management, Faculty of Management,Lublin University of Technology, Nadbystrzycka 38D, 20-618 Lublin, Poland, *e-mail: a.surowiec@pollub.pl*

Keywords: Life expectancy, harmonic analysis, Fibonacci numbers, periodogram.

identification of these components. The analyzed time interval includes the years 2000-2013. The data came from [8]. The methodology of calculation these data are presented in [1], [5]. Life expectancy at age zero ($t = 0$) is not considered in the paper. The model parameters are estimated according to the Least Square Method. (LSM).

Because the correctness of the econometric model is usually verified by the means of relative uncertainty of prognoses the aim of the present work was to find the models with the least relative errors. Models with periodic components usually can have small relative errors. Within the scope of investigation in the identification of periodic components the following times series methods are used: periodogram [2], [7], [4],, classical harmonic analysis [6] and modified classical harmonic analysis. Also, to create the periodic components the numbers that are given in the work [6] and Fibonacci numbers are used.

The starting point for the identification of periodic component in the model of life expectancy for men in Poland P_t in the age $t = 1, ..., 100$ in years $2000 - 2013$ is the sequence of the residuals in the following form:

$$R_t = P_t - f(t), \tag{1.1}$$

where $f(t)$, $t = 1, ..., 100$ is polynomial of degree p.

Identification of the periodic components in the model P_t is the determination of such angles ω_j that maximal relative error:

$$\delta_{t_{\max}} = \max_t |\delta_t|, \tag{1.2}$$

of life expectancy model

$$P_t = f(t) + \sum_{j=1}^{q} (\alpha_j \cos \omega_j t + \beta_j \sin \omega_j t) + \varepsilon_t. \tag{1.3}$$

was the least. Relative error δ_t in the equation is given by the following formula:

$$\delta_t = \frac{P_t - \widehat{P}_t}{P_t} 100\% \tag{1.4}$$

where

$$\widehat{P}_t = f(t) + \sum_{j=1}^{q} (\alpha_j \cos \omega_j t + \beta_j \sin \omega_j t)\,. \tag{1.5}$$

The determination of degree p of polynomial $f(t)$ is the first problem. Unfortunately, the relative errors for all analyzed years for the polynomial model of degree $p = 1$ established on the basis of time series methods [2] are very big and the relative errors of parameters estimation in these models are

Table 1: Maximal relative errors and maximal relative error parameters estimation of models (1), (2) and (3) of life expectancy for men for years 2000-2013

	$model(1)$		$model(2)$		$model(3)$	
$years$	a	b	a	b	a	b
2000	$454,527$	$1,885$	$25,798$	$23,924$	$19,191$	$0,439$
2001	$431,139$	$1,866$	$24,963$	$23,020$	$18,293$	$0,452$
2002	$450,624$	$1,831$	$23,486$	$15,827$	$14,409$	$0,444$
2003	$484,189$	$1,823$	$22,548$	$12,280$	$11,754$	$0,420$
2004	$463,461$	$1,806$	$19,648$	$18,047$	$12,758$	$0,363$
2005	$459,893$	$1,792$	$19,003$	$20,957$	$13,044$	$0,355$
2006	$443,540$	$1,784$	$14,861$	$20,799$	$9,532$	$0,329$
2007	$438,416$	$1,780$	$14,348$	$21,663$	$9,306$	$0,326$
2008	$443,061$	$1,752$	$12,853$	$15,700$	$6,369$	$0,327$
2009	$460,842$	$1,741$	$13,677$	$11,613$	$4,005$	$0,389$
2010	$430,176$	$1,722$	$13,396$	$11,108$	$3,724$	$0,407$
2011	$422,406$	$1,691$	$11,682$	$9,921$	$2,867$	$0,411$
2012	$419,002$	$1,689$	$12,403$	$8,457$	$3,502$	$0,463$
2013	$413,610$	$1,678$	$13,825$	$7,387$	$4,142$	$0,534$

greater than 10%. The polynomial of degree $p = 2$ is not suitable either. The following polynomial model, see [6]:

$$f(t) = \gamma_0 + \gamma_1 t + \gamma_3 t^3 \tag{1.6}$$

comes in handy.

Only the results of testing the polynomials models for $p = 1$ (1), $p = 3$ (2) and the results of testing model (1.6) (3) of the life expectancy for years 2000-2013 are shown in Table 1.

In Table 1 we present maximal relative errors $\delta_{t_{\max}}$ (kolumn a) for models (1), (2) and (3) and maximal relative errors rel$S_{\gamma_{\max}}$(kolumn b) for these models. We define $\mathrm{rel}S_{\gamma_{\max}}$ by the following formula:

$$\mathrm{rel}S_{\gamma_{\max}} = \max_i \mathrm{rel}S_{\gamma_i}, \tag{1.7}$$

where $relS_{\gamma_i}$ is the relative estimation error of parameter i.

In model 1 the parameter γ_1 has the biggest relative estimation error, in model (2) γ_2 and in model (3) γ_3.

2. Periodogram

Lets construct the periodogram [2] for the times series R_t. For even number of observation $N = 100$ ($N = 2q$, $q = 50$), the following function:

$$I(f_j) = \begin{cases} \frac{N}{2}(a_j^2 + b_j^2) & \text{for } j = 1, 2, ..., q-1, \\ Na_j^2 & \text{for } j = q. \end{cases} \tag{2.1}$$

is the periodogram, where

$$a_j = \frac{2}{N}\sum_{t=1}^{N} R_t \cos\omega_j t, j = 1, 2, \ldots, q-1, \tag{2.2}$$

$$b_j = \frac{2}{N}\sum_{t=1}^{N} R_t \sin\omega_j t, j = 1, 2, \ldots, q-1, \tag{2.3}$$

$$a_q = \frac{1}{N}\sum_{t=1}^{N} (-1)^t R_t, \tag{2.4}$$

We determine the maxima of periodogram. These maxima correspond to such i, that the angles ω_i are conected with determined i by the following equation:

$$\omega_i = 2\pi\frac{i}{N}. \tag{2.5}$$

3. Harmonic analysis and modified harmonic analysis

In harmonic analysis for times series of residuals R_t we determine the mean value $\overline{R}$:

$$\overline{R} = \frac{1}{N}\sum_{t=1}^{N} R_t \tag{3.1}$$

and standard deviation $\sigma(R_t)$:

$$\sigma(R_t) = \sqrt{\frac{1}{N-1}\sum_{t=1}^{N}(R_t - \overline{R_t})^2}. \tag{3.2}$$

We set the following critical value connected with the standard deviation:

$$\kappa^* = 3\sqrt{\frac{N}{2}\sigma(R_t)}. \tag{3.3}$$

for the following coefficients

$$\kappa_c(k) = \sum_{t=1}^{N} R_t \cos(\frac{2k\pi}{n}t). \tag{3.4}$$

$$\kappa_c(k) = \sum_{t=1}^{N} R_t \sin(\frac{2k\pi}{N}t). \tag{3.5}$$

where $k \in [0, \frac{N}{2}]$.

In classical harmonic analysis we determine the number q of local extrema of functions κ_c and κ_s. The absolute value of these extrema must be greater than critical value κ^*. We look for such k_i, where $i = 1, 2, ..., q$ that correspond to these extrema.

Similarly, we determine the number q of local extrema of functions κ_c and κ_s that in the same way are connected with critical value κ^* in modified harmonic analysis but in this method we look for such k_i, where $i = 1, 2, ..., q$, that correspond to zeroes of a function κ_s for extrema of function κ_c and that correspond to zeroes of a function κ_c for extrema of function κ_s. The angles ω_i are connected with determined k_i by the following equation:

$$\omega_i = \frac{2k_i\pi}{N}. \tag{3.6}$$

4. Models with predetermined periodic components

4.1. Polynomial model with 4 angles

In this section we test the model of life expectancy P_t for men in the age from $t = 1, ..., 100$ in Poland introduced in [6] with only 4 angles in the following form:

$$\begin{aligned} P_t = \gamma_0 + \gamma_1 t + \gamma_3 t^3 + \\ +\alpha_1 \cos(\frac{\pi}{42}t) + \beta_1 \sin(\frac{\pi}{42}t) + \\ +\alpha_2 \cos(\frac{\pi}{35}t) + \beta_2 \sin(\frac{\pi}{35}t) + \\ +\beta_3 \sin(\frac{\pi}{28}t) + \\ +\alpha_4 \cos(\frac{\pi}{21}t) + \beta_4 \sin(\frac{\pi}{21}t) + \varepsilon_t. \end{aligned} \tag{4.1}$$

where $f(t)$ is chosen polynomial given by equation (1.6).

4.2. Models with Fibonacci numbers

We can obtain the model with small relative errors by using only the periodic components (without the polynomial).

In this section we define the model of life expectancy for men in Poland by Fibonacci numbers. Lets $\omega_j = \frac{2\pi}{F_{j+7}}$ for $j = 1, 2, ...7$ where F_{j+7} is $(j+7)^{th}$ Fibonacci number. Therefore we take the following Fibonacci numbers: $F_8 = 21$, $F_9 = 34$, $F_{10} = 55$, $F_{11} = 89$, $F_{12} = 144$, $F_{13} = 233$, $F_{14} = 377$.

We consider only 7 angles because of simplicity of econometric model.

Therefore the analyzed model has the following form:

$$P_t = \sum_{j=1}^{7} (\alpha_j \cos \omega_j t + \beta_j \sin \omega_j t) + \varepsilon_t. \tag{4.2}$$

5. Results

The results presented in this work have been obtained by means of these methods described in chapters above. These methods are to determine which angles (periods) should enter the final models. To estimate the parameters of constructed models the LSM is used. The effectiveness of these models is tested by the relative error.

5.1. Results from the periodogram

The periodogram for all analyzed years 2000-2013 has the same shape. Maximum of all periodogram corresponds to $i = 2$. It changes from $0,77$ in year 2008 to 2.10 in year 2001. A sample periodogram is presented in Figure 1 for year 2013, for $i = 1, 2, ..., 10$.

Because of the following models:
model (1):

$$P_t = \gamma_0 + \gamma_1 t + \gamma_3 t^3 + \alpha_2 \cos(\frac{2\pi}{50}t) + \beta_2 \sin(\frac{2\pi}{50}t) + \varepsilon_t,$$

model ($1a$):

$$P_t = \gamma_0 + \gamma_1 t + \gamma_3 t^3 + \alpha_2 \cos(\frac{2\pi}{50}t) + \varepsilon_t,$$

model ($1b$):

$$P_t = \gamma_0 + \gamma_1 t + \gamma_3 t^3 + \beta_2 \sin(\frac{2\pi}{50}t) + \varepsilon_t$$

are not very good (see Figure 2 and 3) we analyze the following models as well:
model (2):

$$\begin{aligned} P_t = \gamma_0 + \gamma_1 t + \gamma_3 t^3 + \\ +\alpha_1 \cos(\frac{\pi}{50}t) + \beta_1 \sin(\frac{\pi}{50}t) + \\ +\alpha_2 \cos(\frac{2\pi}{50}t) + \beta_2 \sin(\frac{2\pi}{50}t) + \\ +\alpha_3 \cos(\frac{3\pi}{50}t) + \beta_3 \sin(\frac{3\pi}{50}t) + \varepsilon_t, \end{aligned}$$

model (3):

$$\begin{aligned} P_t = \gamma_0 + \gamma_1 t + \gamma_3 t^3 + \\ +\alpha_1 \cos(\frac{\pi}{50}t) + \beta_1 \sin(\frac{\pi}{50}t) + \\ +\alpha_3 \cos(\frac{3\pi}{50}t) + \beta_3 \sin(\frac{3\pi}{50}t) + \varepsilon_t, \end{aligned}$$

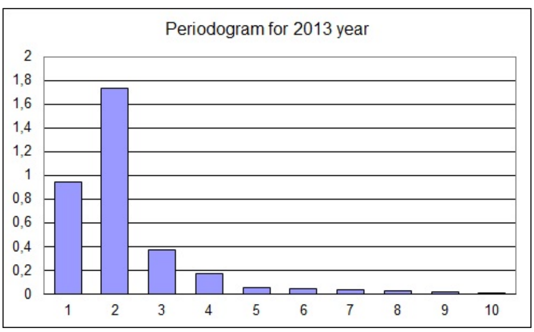


Figure 1: Periodogram of life expectancy for year 2013 for $i = 1, 2, ..., 10$.

model (4):

$$P_t = \gamma_0 + \gamma_1 t + \gamma_3 t^3 + \alpha_1 \cos(\frac{\pi}{50} t) + \beta_1 \sin(\frac{\pi}{50} t) + \varepsilon_t.$$

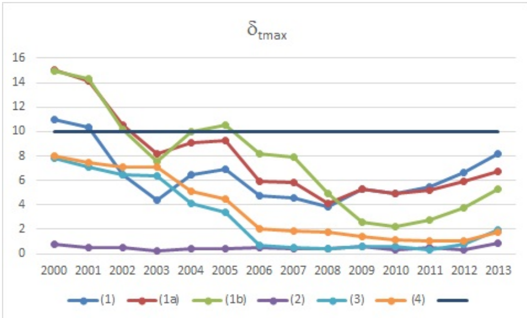


Figure 2: Maximal relative errors for models (1), (1a), (1b), (2), (3) and (4) in years 2000-2013.

	model (1)		model (1a)	model (1b)	model (2)				model (3)	
years	2s	2c	2c	2s	3s	3c	2s	2c	3s	3c
2000			11	13		13		33	134	48
2001			11	14				25	84	33
2002			13	13				20	42	22
2003		11	15	13				11	29	16
2004	12		10	18				14	14	
2005	12			20				23	12	
2006	22			36			324	86		
2007	21			34			30	194		
2008	23			32			90	169		
2009	16	10	12	22			30	52		
2010	14		12	19			11	25		
2011	15	11	13	19	12		26	34	15	
2012	14	13	16	17	11		15	40	18	
2013	12	15	19	15	19			35	198	

Figure 3: The map of the relative errors of parameters estimation for models (1), (1a), (1b), (2) and (3) in years 2000-2013. Black square with integer value of error is for the parameters errors greater than 10%.

The maximal relative errors and all relative errors of parameters estimation of model (4) of P_t are less than 10%. The relative errors of parameters estimation γ_i, α_1 and β_1 in all analyzed models of P_t are less than 10% too. To this end we do not present them in Figure 3.

5.2. Results from harmonic analysis and from the modified harmonic analysis

Firstly we analyze the graphs of functions given by equation (3.4) and (3.5). The sample of these graphs are presented in Figure 4. For years 2000 – 2005 and 2012 we can have 4 angles (see Figure 4 on left), for years 2006 – 2011 and 2013 we can have only 3 angles (see Figure 4 on right). Taking into account the results presented in Figure 4 we can obtain different values of k_i in methods: the classical harmonic analysis and modified harmonic analysis. These values of k_i in equation (3.6) for models of life expectancy for men in years 2000-2013 are presented in Table 2.

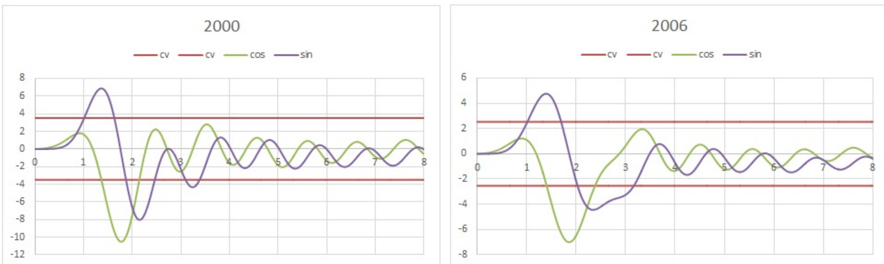


Figure 4: Graph of functions Cosinus and Sinus for 2000 and 2006 year

Therefore in this section we test 4 following models:

- $classical4$ is the model with 4 angles obtained in classical harmonic analysis.
- $classical3$ is the model with 3 angles obtained in classical harmonic analysis.
- $modified4$ is the model with 4 angles obtained in modified harmonic analysis.
- $modified3$ is the model with 3 angles obtained in modified harmonic analysis.

The results of maximal relative errors for models with 3 and 4 angles in classical and modified analysis in years 2000-2013 are presented in Figure 5). The results for 4 angles in years 2000 – 2005 are nearly the same. In modified harmonic analysis the maximal relative errors are less than in classical harmonic analysis in all years.

The maximal relative errors for the models with 4 angles are less then the maximal relative errors for the corresponding models with 3 angles see Figure

Table 2: Values of k_i in equation (3.6) for the classical harmonic analysis and for the modified harmonic analysis for models of life expectancy for men in years 2000-2013

years	classical method				modified method			
i	1	2	3	4	1	2	3	4
2000	1.774	1.365	2.151	3.241	1.756	1.186	2.281	3.247
2001	1.777	1.365	2.157	3.228	1.761	1.184	2.301	3.232
2002	1.731	1.299	2.131	3.169	1.715	1.102	2.324	3.159
2003	1.691	1.252	2.102	3.1	1.681	1.055	2.399	3.102
2004	1.777	1.34	2.188	3.091	1.771	1.144	2.542	3.06
2005	1.805	1.372	2.213	3.098	1.798	1.181	2.57	3.048
2006	1.86	1.396	2.344	–	1.86	1.195	2.873	–
2007	1.862	1.401	2.338	–	1.859	1.202	2.853	–
2008	1.811	1.313	2.612	–	1.81	1.103	2.794	–
2009	1.719	1.238	2.366	–	1.721	1.038	2.727	–
2010	1.708	1.235	2.252	–	1.708	1.035	2.731	–
2011	1.675	1.206	2.305	–	1.68	1.011	2.698	–
2012	1.63	1.179	2.139	2.593	1.638	0.99	2.735	–
2013	1.6	1.161	2.058	–	1.603	0.975	2.733	–

5. Nevertheless the relative errors of almost all parameters estimation for the model with 4 angles are significantly greater than 10% (see Figures 6). Both models (obtained in classical and modified harmonic analysis) with 3 angles work worse in years $2000-2005$ than in years $2006-2013$ (see Figure 5 and 7).

5.3. Results for polynomial model with 4 predetermined angles

In this section the results for deterministic model with predetermined angles (see equation (4.1) are presented. The Figure 8 presents maximal relative errors for this model. The values of this model in all analyzed years does

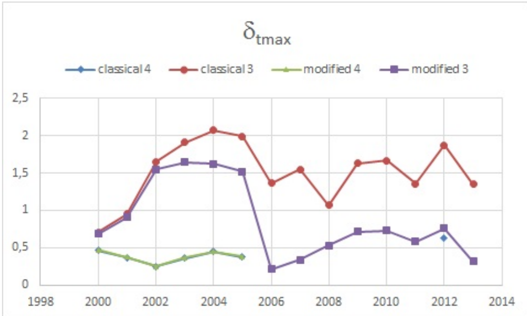


Figure 5: Maximal relative errors for models with 3 and 4 angles in classical and modified analysis in years 2000-2013.

years	classical method for 4 angles								modified method for 4 angles							
i	4	4	3	3	2	2	1	1	4	4	3	3	2	2	1	1
2000		1117								513						
2001		52						10		58		11				
2002		20		12	11	10		11		26	10	26	11	12		
2003	14	24	22	23	20	21	16	18	16	35	23	145	18	27	13	14
2004	14	20	24	28	25	22	16	27	14	43	30	28	23	26	13	15
2005	11	22	18	24	25	17	13	52	10	80	26	18	24	20	11	13
2012	20	112	200	62	255	3979	99	372	-	-	-	-	-	-	-	-

Figure 6: The map of the relative errors of parameters estimation for models with 4 angles in years 2000-2013. Black square with integer value of error is for the parameters errors greater than 10%.

years	classical method for 3 angles						modified method for 3 angles					
i	3	3	2	2	1	1	3	3	2	2	1	1
2000	1082	10		26	11		36			54	13	
2001	65			17			19			26		
2002	88			18			15			32	11	
2003	61			13						20		
2004	16							11		11		
2005	15							15		12		
2006												
2007												
2008		418										
2009										12		
2010										13		
2011										11		
2012										18		
2013										78		

Figure 7: The map of the relative errors of parameters estimation for models with 3 angles in years 2000-2013. Black square with integer value of error is for the parameters errors greater than 10%.

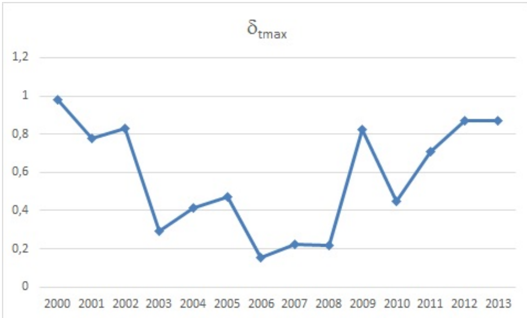


Figure 8: Relative errors for models of the life expectancy with 4 predetermined angles in years 2000-2013.

not exceed 1% but the relative errors of the parameters estimation for this model do not exceed 10% for only year 2002 (see Figure 9).

years	β4	α4	β3	β2	α2	β1	α1
2000	11	10	13	17	15	17	18
2001			11	15	12	11	16
2002							
2003				15			13
2004				14			13
2005				10			
2006				15			12
2007				12			10
2008				17			13
2009		12	22	108	25		165
2010			12	75	13		31
2011		15	30	46	33		235
2012		14	26	62	29		850
2013		30	148	24	297		38

Figure 9: The map of the relative errors of parameters estimation for models with 4 predetermined angles in years 2000-2013. Black square with integer value of error is for the parameters errors greater than 10%.

5.4. Results for models with Fibonacci numbers

In this section 3 models with Fibonacci numbers are analyzed. In the model (1) the following Fibonacci numbers are used: $F_9 - F_{13}$, in the model (2) the same Fibonacci numbers are used like in the model 1 but the component $\cos(\frac{2\pi}{F_9})$ is absent. In the model (3) the following Fibonacci numbers are used: $F_8 - F_{14}$. The maximal relative errors of these models are presented in Figure 10 and the map of relative errors of parameter estimation are presented in Figure 11. The maximal relative errors of the model with 7 Fibonacci numbers are less than 0.5% see (Figure 10). This model is very stable in all analyzed period $2000 - 2013$ even though the relative errors of almost all parameters estimation for this model are significantly greater than 10% (see Figures 11. The model (1) are best in the period $2000 - 2005$ than the model (2) and the model (2) are best in the period $2006 - 2013$ than the model (1).

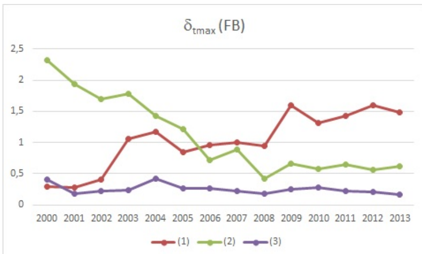


Figure 10: Relative errors of models of the life expectancy with Fibonacci numbers in years 2000-2013.

model with Fibonacci numbers(1)										model with Fibonacci numbers (2)									model with Fibonacci numbers (3)													
cos	sin	cos	sin	cos	sin	cos	sin	cos	sin	cos	sin	cos	sin	cos	sin	cos	sin	sin	cos	sin	cos	sin	cos	sin	cos	sin	cos	sin	cos	sin	cos	sin
233	233	144	144	89	89	55	55	34	34	233	233	144	144	89	89	55	55	34	377	377	233	233	144	144	89	89	55	55	34	34	21	21
															11		22	18	19	20	20	18	19	23	34	18	12	21	12	11	31	11
							11										18	14	29	33	32	36	31	35	45	30	20	31	54	14	17	13
							12										18		19	21	21	24	20	21	23	20	16	19	81		30	14
								23									11	10	23	26	25	37	26	24	24	25	26	22	15	13	878	
								66											80	137	104	398	131	83	68	120	203	71	13	23	26	28
							10	42									10		22	24	23	33	24	22	22	23	23	21	14	11	252	11
								43											34	41	39	55	39	41	53	38	28	44	11	14	266	15
								93											90	136	124	231	111	167	3988	101	45	195	15	19	108	24
								16											45	55	52	75	51	56	78	49	35	59	11	16	7908	13
					10			14	18						11			20	68	97	87	195	88	97	191	86	53	144		30	32	
								16	12									13	33	40	38	56	38	39	48	38	29	43		18	34	11
								13	31						11			32	39	50	46	84	48	46	54	47	40	54		27	24	
					10			13	27						11			28	24	30	27	52	29	26	26	29	30	28		17	36	
					13			14	##						13			##	30	39	35	73	38	35	40	38	35	43		31	28	

Figure 11: The map of the relative errors of parameters estimation in models with Fibonacci numbers in years 2000-2013. Black square with integer value of error is for the parameters errors greater than 10%.

Conclusion

Because of *"Econometric model building is an art, as an art is to design a building"* [3] we tested many different methods of identification of the periodic components in the sequences of the residuals of the life expectancy model for men in years 2000-2013. The methods described in section (2) and (3): the periodogram, the classical harmonic analysis and the modified harmonic analysis are based on times series theory [2], [7]. These models are probabilistic models. The models described in section (4) are the deterministic models that are not substantiated any theory. Nevertheless, the maximal relative errors of these deterministic models are significantly less than the maximal relative errors of probabilistic models. The models obtained by means the periodogram are the worst (see Figure 2). The most accurate model is the model with 7 Fibonacci numbers even though the relative errors of almost all parameters estimation for this model are significantly greater than 10% (see Figures 11). The maximal relative errors for model with 7 Fibonacci number are less than 0.5%. Furthermore, this model is stable in all analyzed years (see Figure 10). The maximal relative errors for the probabilistic models obtained in classical and modified harmonic analysis are less than 0.5% as well, but these models exist only in $2000 - 2005$ years (see Figure 5). The relative errors of almost all parameters estimation for this model are significantly greater than 10% (see Figures 6) as well.

References

1. L. Bolesławski, *Trwanie życia w 2000 r., (Life tables of Poland 2000), „Informacje i Opracowania Statystyczne"*, GUS, Warszawa 2001.
2. G. E. P. Box, G. M. Jenkins, *Analiza szeregów czasowych, Prognozowanie i sterowanie*, PWN, Warszawa, 1983.
3. G. C. Chow, *Econometrics*, Mc-Graw-Hill Book company, New York, 1983.
4. M. Osińska, *Ekonometria finansowa*, PWE, Warszawa 2006.
5. L. Rutkowska, *Trwanie życia w 2013 r., (Life expectancy tables of Poland 2013), „Informacje i Opracowania Statystyczne"*, GUS, Warszawa 2014.
6. W. Rzymowski, *Three models with periodic variables*, in Process control in management, Ed. T. Banek, Toruń, 2009.
7. J. W. Wiśniewski, Z. Zieliński, *Elementy ekonometrii*, Wyd. Uniwersytetu Mikołaja Kopernika, Toruń 2004.
8. http://stat.gov.pl/obszary-tematyczne/ludnosc/trwanie-zycia/trwanie-zycia,1,1.html
9. *"International Human Development Indicators – UNDP"*. Hdrstats.undp.org. Retrieved November 4, 2010.