https://doi.org/10.15388/vu.thesis.716
https://orcid.org/0000-0003-3456-9054

VILNIUS UNIVERSITY

Rokas Puišys

Survival with Random Effect Transformation and its Properties

Doctoral Dissertation

Natural Sciences,
Mathematics (N 001)

VILNIUS 2025

The dissertation was prepared between 2020 and 2024 at Vilnius University.

Scientific supervisor – Prof. Dr. Jonas Šiaulys (Vilnius University, Natural Sciences, Mathematics - N 001).

Academic consultant – Assoc. Prof. Dr. Andrius Grigutis (Vilnius University, Natural Sciences, Mathematics - N 001).

Dissertation Defence Panel:

Chairman – Prof. Habil. Dr. Vydas Čekanavičius (Vilnius University, Natural Sciences, Mathematics - N 001).

Members:

Prof. Dr. Paulius Drungilas (Vilnius University, Natural Sciences, Mathematics - N 001).

Prof. Dr. Jurgita Markevičiūtė (Vilnius University, Natural Sciences, Mathematics - N 001).

Assoc. Prof. Dr. Viktor Skorniakov (Vilnius University, Natural Sciences, Mathematics - N 001).

Prof. Dr. Yuliya Mishura (Taras Shevchenko National University of Kyiv, Natural Sciences, Mathematics - N 001).

The dissertation shall be defended at a public meeting of the Dissertation Defence Panel at 12 p.m. on 27 January 2025 in Room 201 of the Faculty of Mathematics and Informatics, Vilnius University. Address: Naugarduko 24, Vilnius, Lithuania. Tel. +370 52193050; e-mail: mif@mif.vu.lt.

The text of this dissertation can be accessed at the libraries of Vilnius University, as well as on the website of Vilnius University:
www.vu.lt/lt/naujienos/ivykiu-kalendorius

https://doi.org/10.15388/vu.thesis.716
https://orcid.org/0000-0003-3456-9054

VILNIAUS UNIVERSITETAS

Rokas Puišys

Išgyvenamumas su atsitiktinio efekto transformacija ir jos savybės

Daktaro disertacija

Gamtos mokslai,
Matematika (N 001)

VILNIUS 2025

Disertacija rengta 2020-2024 metais Vilniaus universitete.

Mokslinis vadovas – prof. dr. Jonas Šiaulys (Vilnius universitetas, gamtos mokslai, matematika - N 001).

Mokslinis konsultantas – doc dr. Andrius Grigutis (Vilnius universitetas, gamtos mokslai, matematika - N 001).

Gynimo taryba:

Pirmininkas – prof. habil. dr. Vydas Čekanavičius (Vilnius universitetas, gamtos mokslai, matematika - N 001).

Nariai:

prof. dr. Paulius Drungilas (Vilnius universitetas, gamtos mokslai, matematika - N 001).

prof. dr. Jurgita Markevičiūtė (Vilnius universitetas, gamtos mokslai, matematika - N 001).

doc. dr. Viktor Skorniakov (Vilnius universitetas, gamtos mokslai, matematika - N 001).

prof. dr. Yuliya Mishura (Kijevo nacionalinis Taraso Ševčenkos universitetas, gamtos mokslai, matematika - N 001).

Disertacija ginama viešame Gynimo tarybos posėdyje 2025 m. sausio 27 d. 12 val. Vilniaus universiteto Matematikos ir informatikos fakulteto 201 auditorijoje. Addresas: Naugarduko g. 24, Vilnius, Lietuva. Tel. +370 52193050; el. paštas: mif@mif.vu.lt.

Disertaciją galima peržiūrėti Vilniaus universiteto bibliotekoje ir VU interneto svetainėje adresu:
www.vu.lt/lt/naujienos/ivykiu-kalendorius

Acknowledgments

I want to express my sincere gratitude to my supervisor professor Jonas Šiaulys who guided me since my Bachelor's studies. His patience, experience and sharing of knowledge motivated and pushed me to keep working hard, especially during difficult times. I would also like to thank lecturers of Vilnius University, who always shared their time with me and answered all the questions which helped me to understand the value and importance of things they wanted to teach. I am very grateful to my study partners who very quickly became very dear friends of mine. They always showed support and believed in me. Last, but certainly not least - thank You to my family: You always wanted what's best for me and You didn't force me to choose my academic and career path and I will always cherish that.

Contents

1 Notation

$S(x)$	survival function
$F_X(x)$	distribution function of a random variable X
$\overline{F}_X(x)$	tail of distribution function
μ_x	force of mortality (hazard rate)
m_x	central mortality rate
$\widehat{S}(x) = \mathbb{E}\left(S^Z(x)\right)$	transformed survival function
$f_X(x)$	density function of a random variable X
${}_np_x$	probability of an x-aged individual to live at least n years
${}_nq_x$	probability of an x-aged individual to not live at least n years
$\mathrm{e}_{x,\overline{n}\|}$	partial remaining life expectancy of n upcoming years for an x-aged individual
H_x	cumulative hazard rate
D_x	estimated death counts at time moment x
E_x	estimated population exposed to risk at time moment x
LF	loss function
MSE	mean squared error
$\mathcal{H}$	class of heavy-tailed distribution functions
$\mathcal{L}(\gamma)$	class of exponential-type distribution functions
$\mathcal{S}$	class of subexponential distribution functions
$\mathcal{R}_\alpha$	class of regularly varying distribution functions with index α
$\mathcal{R}$	class of all regularly varying distribution functions
$\mathcal{D}$	class of dominatedly varying-tailed distribution functions
$\mathcal{L}$	class of long-tailed distribution functions
$\mathcal{OL}$	class of generalized long-tailed distribution functions
$\mathcal{C}$	class of consistently varying-tailed distribution functions
$\Gamma(k)$	standard Gamma function

2 Introduction

2.1 Relevance of the topic

Analysis of the random variables properties has long been one of the main topics in probability theory and statistics. Variety of random variables (also called as random effect in this thesis) opens the possibility to construct different distribution functions with the special properties which are widely used not only in theoretical mathematical studies but also in banking, life insurance and risk theory. The tail of distribution function has a very important meaning as its behaviour can be analyzed separately in terms of regularity classes. It turns out that the tail of distribution function can satisfy many different properties and can be assigned to its own specific regularity class. Tail of distribution function is widely used in survival analysis, therefore, it is a synonym to a survival function. In survival analysis, this function provides possibility to calculate variety of mortality characteristics such as mortality probability, force of mortality (also known as hazard rate), remaining life expectancy and similar. By knowing exact expression of survival function of a specific population, it is possible to compute mortality tables which can be used when calculating life insurance premiums, technical provisions and other financial characteristics. Additionally, some survival functions can be used for different populations by simply applying different function parameter values. The special properties of random variables and distribution functions have always been a relevant topic for mathematicians, financial industry and even medicine. It turns out that combination of different survival functions together with random variables at once opens up new possibilities to create new mortality models, which will be one of the main focus in this thesis. The following subsection will contain clear objectives of the research.

2.2 Main goals

The main objectives of the thesis will be split into two parts: theoretical and practical. Full definitions, formulas and properties will be provided in the later sections while this section will contain short descriptions and will emphasize the main goals of analysis.

2.2.1 Theoretical

Assume that we have a probability space $(\mathbf{W}, \mathcal{F}, \mathbf{P})$ with a random variable X defined on it. Let us define the distribution function (d.f.) of a random variable X:

$$F_X(x) := \mathbf{P}(X \leqslant x) = \mathbf{P}(w \in \mathbf{W} : X(w) \leqslant x),\ x \in \mathbb{R}.$$

The tail of such distribution function is equal to

$$\overline{F}_X(x) = 1 - F_X(x) = \mathbf{P}(X > x),$$

In survival analysis, the tail of distribution function is called a survival function and has the expression

$$S(x) = \mathbf{P}(T > x),$$

where $T : \mathbf{W} \to \mathbb{R}$ denotes random variable of the remaining life expectancy. We note here that in survival analysis it is considered that the remaining life expectancy T is an absolutely continuous non-negative random variable. In such a case, the survival function S is related to the random variable density f_T by the usual equality

$$S(x) = \int_x^\infty f_T(u)\mathrm{d}u,\ x \geqslant 0.$$

Let's assume that we have another positive random variable Z, defined on the another probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Now let's denote the transformation of survival function as

$$\widehat{S}(x) = \mathbb{E}\left(S^Z(x)\right) = \int\limits_{(0,\infty)} \left(S(x)\right)^u \mathrm{d}\,\mathbb{P}(Z \leqslant u).$$

Our first theoretical objective will be to prove that $\widehat{S}(x)$ is a new survival function, meaning that it maintains properties of a survival function. A theory will be constructed based on this assumption in 4.1 and the proof will be provided in Section 5.

Since survival function is equal to the tail of the distribution function, and d.f. tails can be assigned to different regularity classes in terms of their behaviour, the same can be said about survival functions. Therefore, our second theoretical objective will be to prove that if tail function $\overline{F}(x) = \overline{F}_X(x)$ of a random variable X belongs to a specific regularity class, then newly transformed tail function

$$\mathbb{E}\Big(\big(\overline{F}(x)\big)^Z\Big) = \int\limits_{(0,\infty)} \big(\overline{F}(x)\big)^u \mathrm{d}\,\mathbb{P}(Z \leqslant u)$$

with a positive random variable Z will also belong to the same regularity class. The theorem of this assumption will be provided in 4.2 and the proof will be shown in Section 5.

2.2.2 Practical

In the practical part, we will show that based on theorems constructed in 4.1 and 4.2, set of newly constructed transformed survival functions can be used in survival analysis and can precisely describe specific population's mortality behaviour based on real statistics. Two different practical applications are shortly described in paragraphs below.

Baltic population mortality. Mortality data of Lithuanian, Latvian and Estonian populations will be chosen. Variety of different (transformed) survival functions will be applied to the dataset in order to calculate the best estimate parameters of the survival functions based on the mean squared error metric. Calculations will be done in R programming language, using the *MortalityLaws* library. Estimated functions will be used to calculate survival probabilities of each population at each age. The main goal of practical part will be firstly to select which of the survival functions can actually be used in mortality analysis. Later on, model fit for each survival function will be evaluated through graphics as well as a mean squared error metric to determine which of the analyzed models describe specific population's mortality the best. Detailed descriptions, calculations and conclusions are provided in Section 7.

Lithuanian prostate cancer screening data. Two sets of Lithuanian population data will be analyzed. One dataset set will represent male groups who were a part of the prostate screening program and the other datasets will consist of people who were not a part of the program. We will compare the prostate cancer survival probabilities as well as general survival probabilities between the two datasets by applying Weibull mortality model (survival function) to determine if the prostate cancer screening program in Lithuania has been effective. Weibull mortality model's parameters will be calculated based on binomial loss function in R programming language, using the *MortalityLaws* library. Detailed descriptions, calculations and conclusions are provided in Section 8.

2.3 Publications

The main results of this thesis and other related results were published in these scientific papers:

- Šiaulys, J.; Puišys, R. Survival with random effect. *Mathematics* **2022**, *10*, 1097.
- Puišys, R.; Lewkiewicz, S.; Šiaulys, J. Properties of the random effect transformation. *Lith. Math. J.* **2024**, *64*, 177-189.
- Skučaitė, A.; Puvačiauskienė, A.; Puišys, R.; Šiaulys, J. Actuarial analysis of survival among breast cancer patients in Lithuania. *Healthcare* **2021**, *9*, 383.
- Levickytė, J.; Skučaitė, A.; Šiaulys, J.; Puišys, R.; Vincerževskienė, I. Actuarial analysis of survival after breast cancer diagnosis among Lithuanian females. *Healthcare* **2024**, *12*, 746.

2.4 Conferences

The results obtained during the preparation of the thesis were presented in following conferences and seminars:

- R. Puišys. *Kelios išgyvenamumo funkcijos su atsitiktiniu efektu*, The Conference of Lithuanian Mathematical Society, June 16-17, 2021.
- R. Puišys, J. Šiaulys, S.Lewkiewicz. *Kelios atsitiktinio efekto transformacijos savybės*. The Conference of Lithuanian Mathematical Society, June 21-22, 2023, Vilnius, Lithuania.
- R. Puišys, J. Šiaulys. *Survival with random effect*. Data Analysis Methods for Software Systems, November 30 - December 2, 2023, Druskinininkai, Lithuania.
- R. Puišys, J. Šiaulys. *Survival with random effect*. 27th International Congress on Insurance: Mathematics and Economics, July 8 - July 11, 2024. Chicago, Illinois, United States.

3 The main concepts

3.1 Distribution, tail and survival functions

Let X be a random variable, defined on a probability space $(\mathbf{W}, \mathcal{F}, \mathbf{P})$.

Definition 1. *The function, describing behaviour of random variable* $X : \mathbf{W} \to \mathbb{R}$ *is called a distribution function (d.f.) and is defined as*

$$F_X(x) := \mathbf{P}(X \leqslant x) = \mathbf{P}(w \in \mathbf{W} : X(w) \leqslant x),\ x \in \mathbb{R}. \tag{1}$$

The properties of probability space imply the following properties of a d.f. F:

(i) $0 \leqslant F(x) \leqslant 1, \quad x \in \mathbb{R},$ (2)

(ii) $\lim\limits_{x \to -\infty} F(x) = 0, \quad \lim\limits_{x \to +\infty} F(x) = 1,$ (3)

(iii) F is non-decreasing on $\mathbb{R}$, (4)

(iv) F is right continuous and has a finite limit on the left at each $a \in \mathbb{R}$, i.e. $\lim\limits_{x \to a+} F(x) = F(a)$, and $\lim\limits_{x \to a-} F(x)$ exists for $a \in \mathbb{R}$. (5)

It is well known that pure random variables can be either discrete, continuous or singular [27]. In the case of a discrete random variable, distribution function has the expression:

$$F_X(x) = \sum_{x_i \leqslant x} \mathbf{P}(X = x_i), \tag{6}$$

where $x_i \in \mathcal{N}$ and $\mathcal{N} = \{x_i\}$ is a finite or countable set.

If X is absolutely continuous, then distribution function of X has the expression

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\mathrm{d}t,\ x \in \mathbb{R}, \tag{7}$$

where f_X is a density function, i.e. measurable non-negative function such that

$$\int_{-\infty}^{\infty} f_X(t)\mathrm{d}t = 1.$$

We remark only that the above equality (7) implies that $f_X(t) = F^{'}(t)$ almost surely on $\mathbb{R}$.

Definition 2. *Let $F = F_X$ be a distribution function of a random variable X. Then*

$$\overline{F}_X(x) = 1 - F_X(x) = \mathbf{P}(X > x),\ x \in \mathbb{R}, \tag{8}$$

is called a tail function of F.

3.2 Concepts of survival analysis

Let $T \geqslant 0$ be a non-negative absolutely continuous random variable, describing the life expectancy of a newborn. The main function describing mortality of a population is the survival function (see [17, 18, 19]):

$$S(x) = S_T(x) = \mathbf{P}(T > x).$$

Since $S(x) = S_T(x) = 1 - F_T(x) = 1 - \mathbf{P}(T \leqslant x)$ and $T \geqslant 0$, the survival functions satisfy the following properties:

(i) $S(x)$ is non-increasing for $x \in [0, \infty)$; (9)

(ii) $S(0) = 1$; (10)

(iii) $S(\infty) := \lim_{x \to \infty} S(x) = 0$; (11)

(iv) $S(x)$ is absolutely continuous on the interval $[0, \infty)$. (12)

According to the definition of absolute continuity, for an arbitrary $\varepsilon > 0$ there is $\delta > 0$, such that for every finite collection of pairwise disjoint intervals $\{(a_1, b_1), \ldots, (a_n, b_n)\}$ with

$$\sum_{k=1}^{n} (b_k - a_k) < \delta$$

it holds that

$$\sum_{k=1}^{n} \left(S(a_k) - S(b_k)\right) < \varepsilon.$$

By the known results (see for instance [20]), function $S(x)$ is absolutely continuous if and only if:

$$S(x) = \int_x^\infty f_T(u)\mathrm{d}u,\ x \geqslant 0, \tag{13}$$

where f_T is a density function, i.e. non-negative and integrable function such that

$$\int_0^\infty f_T(u)\mathrm{d}u = 1.$$

Additional population characteristics can be calculated with the usage of survival function. For instance, probability of an x-aged individual to live at least n years

$$ {}_np_x = \frac{S(x+n)}{S(x)}, \tag{14} $$

probability that an x-aged individual will not live at least n years

$$ {}_nq_x = 1 - \frac{S(x+n)}{S(x)}, \tag{15} $$

partial remaining life expectancy of n upcoming years for an x-aged individual

$$ \mathring{e}_{x,\overline{n}|} = \frac{1}{S(x)} \int_x^{x+n} S(u)\mathrm{d}u. \tag{16} $$

Another characteristic used in mortality analysis is called the force of mortality, also known as the hazard rate, and is denoted as μ_x. Hazard rate is calculated using following formula:

$$ \mu_x = \frac{f(x)}{S(x)} = -\frac{S'(x)}{S(x)}, \tag{17} $$

where $x \in [0, \infty)$, such that $S(x) > 0$.

Based on the last formula, survival function at each point $x \in [0, \infty)$ can be expressed by equality

$$ S(x) = \exp\left\{ -\int_0^x \mu_u \mathrm{d}u \right\}. \tag{18} $$

The behaviour of hazard rate function in interval $[x, x+1)$ can be described by the central mortality rate:

$$ m_x = \frac{\int_0^1 S(x+u)\mu_{x+u}\mathrm{d}u}{\int_0^1 S(x+u)\mathrm{d}u} = \frac{S(x) - S(x+1)}{\int_0^1 S(x+u)\mathrm{d}u}. \tag{19} $$

If it is assumed that the force of mortality in $[x, x+1)$ is constant, i.e.

$\mu_{x+t} = \mu_x$ for any $x \in \mathbb{N}_0 = \{0, 1, ...\}$ and $t \in [0, 1)$,

then the central mortality rate can be compared with the hazard rate:

$$m_x = \mu_x,\ x \in \mathbb{N}_0.$$

Additonally, the cumulative hazard rate function has the following expression:

$$H_x = -\ln S(x) = \int_0^x \mu_u \mathrm{d}u. \tag{20}$$

3.3 Regularity classes

There is a variety of many different survival functions (tails of distribution functions). Its behaviour at infinity can be split into various categories. Those categories are called the regularity classes. Below we define a list of some well-known regularity classes. A more extensive list can be found in [64], for example.

Definition 3. *A d.f. F is said to be heavy-tailed, $F \in \mathcal{H}$, if for any fixed $\delta > 0$*

$$\limsup_{x\to\infty} \mathrm{e}^{\delta x}\, \overline{F}(x) = \infty. \tag{21}$$

Definition 4. *We say that d.f. F belongs to the exponential-type distribution class $\mathcal{L}(\gamma)$ with $\gamma > 0$ if for any $y > 0$*

$$\lim_{x\to\infty} \frac{\overline{F}(x+y)}{\overline{F}(x)} = \mathrm{e}^{-\gamma y}. \tag{22}$$

Definition 5. *A d.f. F is said to be regularly varying with index $\alpha \geqslant 0$, denoted $F \in \mathcal{R}_\alpha$, if for any $y > 0$ it holds that*

$$\lim_{x\to\infty} \frac{\overline{F}(xy)}{\overline{F}(x)} = y^{-\alpha}. \tag{23}$$

Set of all regularly varying functions is denoted as

$$\mathcal{R} := \bigcup_{\alpha\geqslant 0} \mathcal{R}_\alpha.$$

Definition 6. *A d.f. F is said to be dominatedly varying, $F \in \mathcal{D}$, if*

$$\limsup_{x\to\infty} \frac{\overline{F}(xy)}{\overline{F}(x)} < \infty \tag{24}$$

for any (equivalently for some) $y \in (0, 1)$.

Definition 7. *A d.f. F is said to be long-tailed, $F \in \mathcal{L}$, if*

$$\overline{F}(x-y) \underset{x\to\infty}{\sim} \overline{F}(x) \tag{25}$$

for each (*equivalently for some*) $y > 0$.

Definition 8. *We say that d.f. F belongs to the class $\mathcal{OL}$ if for any* (*or, equivalently, for some*) $y > 0$

$$\limsup_{x\to\infty} \frac{\overline{F}(x-y)}{\overline{F}(x)} < \infty. \tag{26}$$

Definition 9. *A d.f. $F = F_X$ of a non-negative random variable X is said to be subexponential, $F \in \mathcal{S}$, if*

$$\lim_{x\to\infty} \frac{\overline{F*F}(x)}{\overline{F}(x)} = 2. \tag{27}$$

Definition 10. *We say that d.f. F belongs to the class of consistently varying-tailed distribution functions $\mathcal{C}$ if*

$$\lim_{y\uparrow 1} \limsup_{x\to\infty} \frac{\overline{F}(xy)}{\overline{F}(x)} = 1. \tag{28}$$

In this and in the next sections we will cover some basic examples behind formulas and definitions described above.

3.4 Distribution functions assigned to specific regularity classes

List of distribution functions, used in this subsection, can be found in the Appendix.

Example 1. Let's consider the Log-logistic distribution function with shape parameter $\beta = 2$ and scale parameter $\alpha = 3$. Then the survival function (or tail of a d.f.) would be

$$\overline{F}(x) = 1 - \frac{x^2}{9+x^2},\ x \geqslant 0. \tag{29}$$

We will prove that the Log-logistic d.f. belongs to the heavy-tailed distribution class $\mathcal{H}$. This will be proven if for any fixed $\delta > 0$ we will get that

$$\limsup_{x\to\infty} \mathrm{e}^{\delta x}\overline{F}(x) = \infty.$$

We can calculate such limit by applying L'Hospital's rule two times:

$$\limsup_{x\to\infty} \mathrm{e}^{\delta x}\overline{F}(x) = \lim_{x\to\infty}\left(1 - \frac{x^2}{9+x^2}\right)\mathrm{e}^{\delta x} = \lim_{x\to\infty}\frac{9\,\mathrm{e}^{\delta x}}{9+x^2}$$
$$= \lim_{x\to\infty}\frac{9\,\delta^2\,\mathrm{e}^{\delta x}}{2} = \infty.$$

Example 2. Let's consider the Weibull distribution function with parameters $\alpha = 0$ and $a = 1$. Then the survival function (or tail of d.f.) would be

$$\overline{F}(x) = \mathrm{e}^{-x},\ x \geqslant 0. \tag{30}$$

We will prove that this particular case of the Weibull d.f. belongs to the exponential-type distribution class $\mathcal{L}(1)$. This will be proven if for any $y > 0$ we will get

$$\lim_{x\to\infty}\frac{\overline{F}(x+y)}{\overline{F}(x)} = \mathrm{e}^{-y}.$$

Let's take $y > 0$. Then

$$\lim_{x\to\infty}\frac{\overline{F}(x+y)}{\overline{F}(x)} = \lim_{x\to\infty}\frac{\mathrm{e}^{-(x+y)}}{\mathrm{e}^{-x}} = e^{-y}.$$

Example 3. Suppose we have the Peter and Paul d.f. with parameters $a = 1$ and $b = 3$. The tail of such d.f. would be

$$\overline{F}(x) = (3^{-1})^{\lfloor \log_3 x\rfloor} = \left(\frac{1}{3}\right)^{\lfloor \log_3 x\rfloor},\ x \geqslant 1, \tag{31}$$

where symbol $\lfloor a \rfloor$ denotes the integer part of a real number a.

Peter and Paul d.f. has a dominatedly varying tail ($F \in \mathcal{D}$), because it satisfies the following property:

$$\limsup_{x\to\infty}\frac{\overline{F}(xy)}{\overline{F}(x)} < \infty \tag{32}$$

with fixed $y \in (0,1)$. Namely, for $y \in (0,1)$ we obtain

$$\begin{aligned}\limsup_{x\to\infty} \frac{\overline{F}(xy)}{\overline{F}(x)} &= \limsup_{x\to\infty} \left(\frac{1}{3}\right)^{\lfloor \log_3(xy)\rfloor - \lfloor \log_3 x\rfloor} \\ &= \limsup_{x\to\infty} \left(\frac{1}{3}\right)^{\lfloor \log_3 x + \log_3 y\rfloor - \lfloor \log_3 x\rfloor} \\ &= \limsup_{x\to\infty} \left(\frac{1}{3}\right)^{\langle \log_3 x\rangle + \log_3 y - \langle \log_3 x + \log_3 y\rangle} \\ &\leqslant y^{-1} < \infty,\end{aligned}$$

where symbol $\langle a\rangle = a - \lfloor a\rfloor$ denotes the fractional part of a real number a.

We remark here that the class of dominatedly varying functions was introduced by Feller [72] as a generalization of regularly varying distributions and later was considered by many other authors, see [73, 62, 74, 88, 87, 75, 76, 80, 81, 61, 82], for instance.

Example 4. Let's take the Burr distribution with parameters $c = 1$ and $k = 2$. Then the tail of such d.f. is

$$\overline{F}(x) = \frac{1}{(1+x)^2}, \quad x \geqslant 0. \tag{33}$$

Tail of the Burr d.f. belongs to the class of long-tailed distributions $\mathcal{L}$, because

$$\lim_{x\to\infty} \frac{\overline{F}(x+y)}{\overline{F}(x)} = 1$$

for every y. Namely, for the fixed real y we have

$$\begin{aligned}\lim_{x\to\infty} \frac{\overline{F}(x+y)}{\overline{F}(x)} &= \lim_{x\to\infty} \frac{(1+x)^2}{(1+x+y)^2} \\ &= \lim_{x\to\infty} \frac{\left(1+\frac{1}{x}\right)^2}{\left(1+\frac{1}{x}+\frac{y}{x}\right)^2} = 1.\end{aligned}$$

We remark that the class of long-tailed d.f.s was indroduced by Chistyakov [83] in the context of branching processes. Class $\mathcal{L}$, either alone or in intersection with the class $\mathcal{D}$, was considered in [84, 86, 85, 65, 89, 77, 78, 79]. A detailed analysis of distributions from the class $\mathcal{L}$ is presented in Chapter 2 of [65]. For instance, in Lemma 2.23 in [65], closure under mixture and closure under maximum are established. Closure under mixture means that when d.f. $F_1 \in \mathcal{L}$ and either $F_2 \in \mathcal{L}$ or $\overline{F}_2(x) = o(\overline{F}_1(x))$, then $pF_1 + (1-p)F_2 \in \mathcal{L}$ for any $p \in (0,1)$. Closure under maximum means that d.f. $F_{X \vee Y} \in \mathcal{L}$ if $F_X \in \mathcal{L}$, $F_Y \in \mathcal{L}$ for two independent random variables X, Y.

Example 5. Let's consider the Weibull distribution function with parameters $\alpha = -1/2$ and $a = 1/2$. Then the survival function (or tail of d.f.) would be

$$\overline{F}(x) = \mathrm{e}^{-\sqrt{x}},\ x \geqslant 0. \tag{34}$$

Weibull d.f. has the generalized long-tailed distribution ($F \in \mathcal{OL}$) because it meets the requirement of

$$\lim_{x\to\infty} \frac{\overline{F}(x-y)}{\overline{F}(x)} < \infty$$

for any $y > 0$. Namely, for fixed $y \in \mathbb{R}$ we have

$$\begin{aligned}\lim_{x\to\infty} \frac{\overline{F}(x-y)}{\overline{F}(x)} &= \lim_{x\to\infty} \mathrm{e}^{\sqrt{x}-\sqrt{x-y}} \\ &= \lim_{x\to\infty} \mathrm{e}^{\frac{y}{\sqrt{x}+\sqrt{x-y}}} = e^0 = 1 < \infty.\end{aligned}$$

Class $\mathcal{OL}$ was proposed by Shimura and Watanabe [79] as generalization of the exponential-like-tailed distributions. The full range of the properties of the class $\mathcal{OL}$ are presented in [57, 59, 24, 58, 60].

Example 6. Let's analyze the Burr distribution function with parameters $c = 1$ and $k = 2$ as in Example 4. The tail of such d.f. would be

$$\overline{F}(x) = 1 - F(x) = \left(\frac{1}{1+x}\right)^2,\ x \geqslant 0. \tag{35}$$

In Example 4, it was shown that Burr d.f. belongs to class $\mathcal{L}$. However, it can also be proven that Burr d.f. also has a regularly varying tail ($F \in \mathcal{R}_2$) because for any fixed $y > 0$ we have

$$\lim_{x\to\infty} \frac{\overline{F}(xy)}{\overline{F}(x)} = \lim_{x\to\infty} \left(\frac{1+x}{1+xy}\right)^2 = \lim_{x\to\infty} \left(\frac{\frac{1}{x}+1}{\frac{1}{x}+y}\right)^2 = y^{-2}.$$

Other popular distributions, which belong to class $\mathcal{R}$, are Pareto, loggamma, Cauchy. Historical notes, properties and different application of regularly varying functions are described in [63, 65, 69, 64, 66, 67, 68, 26]. Moreover, regular distributions contain variety of closure properties, for example, according to Proposition 3.3 of [64] (see also Proposition on page 278 of [71] or Lemma 1.3.1 of [70]), the class of regularly varying d.f.s $\mathcal{R}$ is closed under strong tail equivalence and convolution. The closure under strong tail equivalence means that

$$F_1 \in \mathcal{R}, \text{ and } \overline{F}_2(x) \underset{x\to\infty}{\sim} c\overline{F}_1(x),\ c > 0 \ \Rightarrow\ F_2 \in \mathcal{R}.$$

Whereas the closure under convolution means that

$$F_1, F_2 \in \mathcal{R} \ \Rightarrow\ F_1 * F_2 \in \mathcal{R},$$

where $F_1 * F_2$ denotes the standard convolution of F_1 and F_2, i.e.

$$F_1 * F_2(x) = \int_{-\infty}^{\infty} F_1(x-y)\mathrm{d}F_2(y),\ x \in \mathbb{R}.$$

Example 7. Let's consider the Pareto distribution function with shape parameter $\alpha = 4$ and scale parameter $x_m = 1$. Then the survival function (or tail of d.f.) would be

$$\overline{F}(x) = \left(\frac{1}{1+x}\right)^4,\ x \geqslant 0. \tag{36}$$

We will prove that Pareto d.f. belongs to the subexponential distribution class $\mathcal{S}$. This will be proven if in interval $[0, \infty)$

$$\lim_{x\to\infty} \frac{\overline{F * F}(x)}{\overline{F}(x)} = 2.$$

Assume we have a random variable X, which has Pareto distribution with d.f. F. It is evident that d.f. $F^{*2}(x) = F * F(x)$ corresponds to the sum of $X_1 + X_2$, where X_1 and X_2 are independent copies of r.v. X. For any positive x we have

$$\begin{aligned}\overline{F^{*2}}(x) &= \mathbb{P}(X_1 + X_2 > x) \geqslant \mathbb{P}(\{X_1 > x\} \cup \{X_2 > x\})\\ &= \mathbb{P}(X_1 > x) + \mathbb{P}(X_2 > x) - \mathbb{P}(X_1 > x)\mathbb{P}(X_2 > x)\\ &= 2\overline{F}(x) - (\overline{F}(x))^2.\end{aligned}$$

Based on upper inequality we get

$$\liminf_{x\to\infty} \frac{\overline{F*F}(x)}{\overline{F}(x)} \geqslant 2.$$

On the other hand, for any $0 < \delta < \frac{1}{2}$ and any $x > 0$

$$\begin{aligned}\overline{F^{*2}}(x) &= \mathbb{P}(X_1 + X_2 > x) \\ &\leqslant \mathbb{P}(\{X_1 > (1-\delta)x\} \cup \{X_2 > (1-\delta)x\} \cup \{X_1 > \delta x, X_2 > \delta x\}) \\ &\leqslant 2\overline{F}((1-\delta)x) + (\overline{F}(\delta x))^2.\end{aligned}$$

We have

$$\begin{aligned}\limsup_{x\to\infty} \frac{\overline{F^{*2}}(x)}{\overline{F}(x)} &\leqslant \limsup_{x\to\infty} \left(\frac{2\overline{F}((1-\delta)x)}{\overline{F}(x)} + \frac{(\overline{F}(\delta x))^2}{\overline{F}(x)} \right) \\ &= 2 \limsup_{x\to\infty} \left(\frac{1+x}{1+(1-\delta)x} \right)^4 + \limsup_{x\to\infty} \frac{(1+x)^4}{(1+\delta x)^8} \\ &= 2(1-\delta)^{-4}.\end{aligned}$$

Since δ is arbitrary from the interval $\left(0, \frac{1}{2}\right)$, we have

$$\limsup_{x\to\infty} \frac{\overline{F^{*2}}(x)}{\overline{F}(x)} \leqslant 2.$$

From the inequalities above we conclude that

$$\lim_{x\to\infty} \frac{\overline{F*F}(x)}{\overline{F}(x)} = 2.$$

Example 8. Let's consider the Pareto distribution function with shape parameter $\alpha = 3$ and scale parameter $x_m = 1$. Then the survival function (or tail of d.f.) would be

$$\overline{F}(x) = \left(\frac{1}{1+x} \right)^3, \; x \geqslant 0. \tag{37}$$

We will prove that Pareto d.f. belongs to consistently varying-tailed distribution class $\mathcal{C}$. This will be proven if we get that

$$\lim_{y\uparrow 1} \limsup_{x\to\infty} \frac{\overline{F}(xy)}{\overline{F}(x)} = 1.$$

Based on L'Hospital's rule, we have

$$\lim_{y\uparrow 1}\limsup_{x\to\infty}\frac{\overline{F}(xy)}{\overline{F}(x)}=\lim_{y\uparrow 1}\limsup_{x\to\infty}\left(\frac{1+x}{1+xy}\right)^3=\lim_{y\uparrow 1}\frac{1}{y^3}=1.$$

Regularity classes, mentioned in both definitions and examples, satisfy the following inclusions (see [70, 83], for instance):

$$\mathcal{R}\subset\mathcal{C}\subset\mathcal{L}\cap\mathcal{D}\subset\mathcal{S}\subset\mathcal{L}\subset\mathcal{H},\qquad \mathcal{D}\subset\mathcal{H},\qquad \mathcal{D}\subset\mathcal{OL}.$$

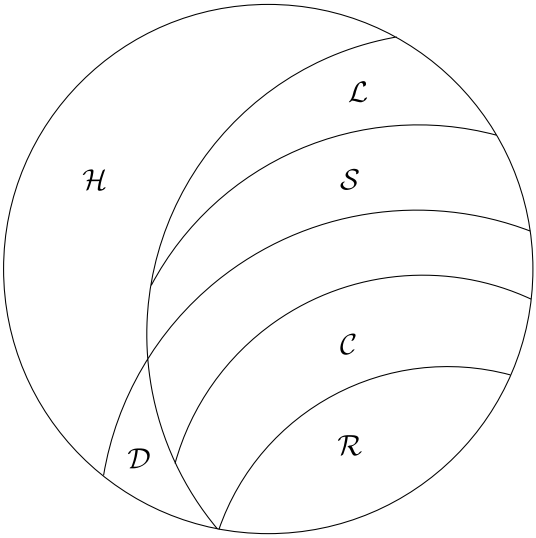

Figure 1: Regularity classes of heavy-tailed distributions

Additionally, it is evident from the definitions that

$$\mathcal{L}\subset\mathcal{OL},\qquad \bigcup_{\gamma>0}\mathcal{L}(\gamma)\subset\mathcal{OL}.$$

It is important to emphasize that tail functions in general do not require absolute continuity, while in survival analysis, the survival function (or tail function) must be absolutely continuous.

3.5 Survival function examples used in mortality analysis

Let's consider the following hazard rate function:

$$\mu_x = ax^{\alpha}, \quad x \geqslant 0, \tag{38}$$

where $a > 0, \ \alpha > 0$. Survival function of such force of mortality would be

$$S(x) = \mathrm{e}^{-\frac{a}{\alpha+1}x^{\alpha+1}}, \ x \geqslant 0. \tag{39}$$

These functions are called the Weibull hazard rate and the Weibull survival functions respectively. These functions were first introduced by Swedish mathematican Waloddi Weibull back in 1939 [90]. A graph containing Weibull survival function with different pairs of a and α values is provided below in Figure 2.

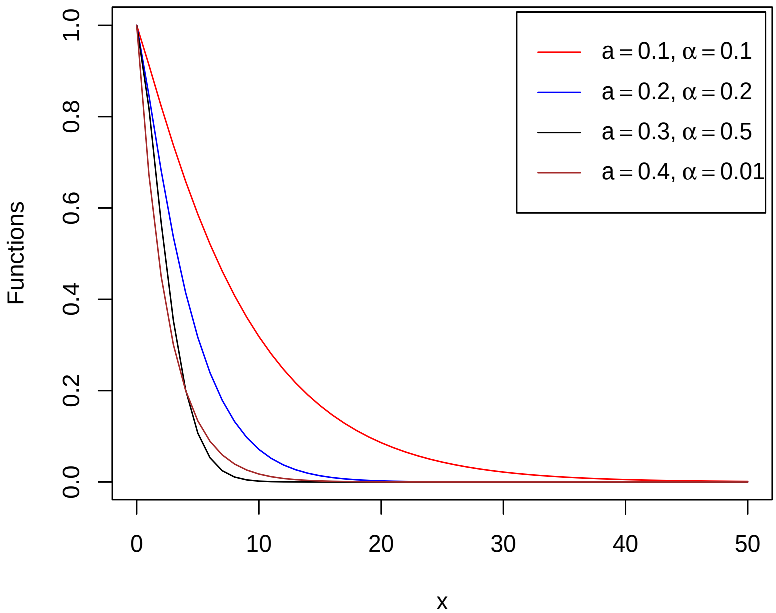

Figure 2: Weibull survival function

Another example of a hazard rate function would be the Gompertz force of mortality:

$$\mu_x = B\mathrm{e}^{\alpha x}, \quad x \geqslant 0, \tag{40}$$

where $B > 0,\ \alpha > 0$. Gompertz survival function has the following expression:

$$S(x) = \mathrm{e}^{-\frac{B(\mathrm{e}^{\alpha x}-1)}{\alpha}},\ x \geqslant 0. \tag{41}$$

These functions were proposed by British self-taught mathematician Benjamin Gompertz in 1825 [53]. A graph containing Gompertz survival function with different pairs of B and α values is provided below in Figure 3.

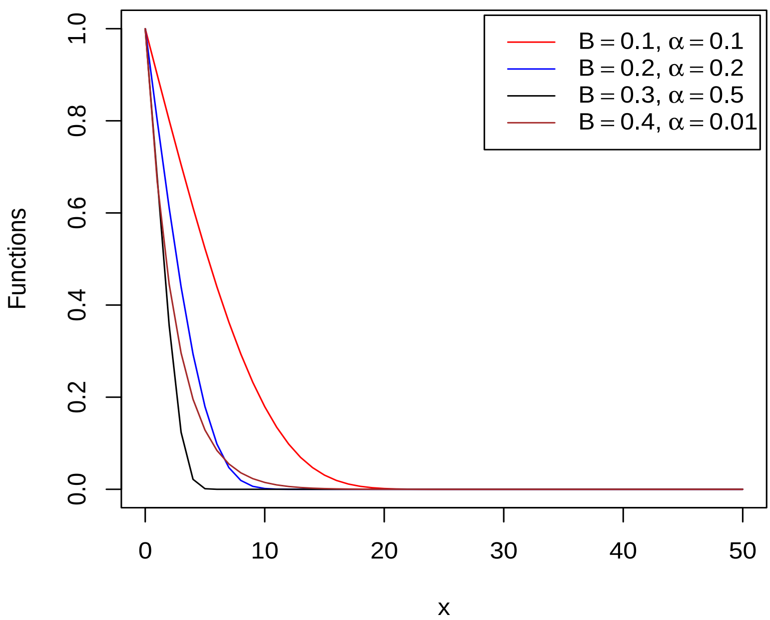

Figure 3: Gompertz survival function

In 1860, English actuary and mathematician William Makeham proposed a slightly modified version of the Gompertz hazard rate [53]:

$$\mu_x = A + B\mathrm{e}^{\alpha x}, \quad x \geqslant 0, \tag{42}$$

where $A \geqslant -B$ is age-independent term, while B and α are the same parameters as in Gompertz functions. The Makeham (or Makeham - Gompertz) survival function is equal to

$$S(x) = \mathrm{e}^{-Ax - \frac{B(\mathrm{e}^{\alpha x}-1)}{\alpha}},\ x \geqslant 0. \tag{43}$$

It is evident that Gompertz functions can also be called the Makeham - Gompertz functions with age-independent term equal to 0. A graph containing Makeham survival function with different pairs of B and α values is provided below in Figure 4.

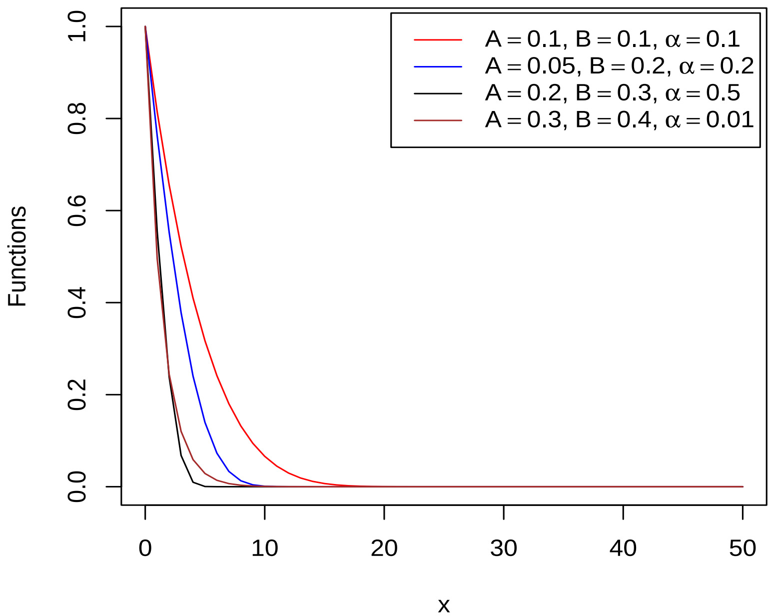

Figure 4: Makeham survival function

4 Main results

4.1 The historical origin of the random effect concept

In 1979, James Vaupel, Kenneth Manton and Eric Stallard [21, 3] presented the idea of modifying mortality force by applying random variable to it:

$$\mu_{x|Z} = Z\mu_x. \tag{44}$$

In such models Z is also called random effect or frailty parameter [22, 23]. Models, followed by presented equation, are called frailty models or random effect models. In some instances, definition of Cox models may also be used, see Lai [4] and Wienke [5], for instance. In frailty models, random effect Z is defined on probability space $(\Omega, \mathcal{A}, \mathbb{P})$, while random variable T, which denotes the remaining life expectancy of a newborn and is used to calculate function μ_x, is defined on another probability space $(\mathbf{W}, \mathcal{F}, \mathbf{P})$. The purpose of random effect models was to obtain more precise results compared to standard mortality models. Frailty models and their applications to specific data sets have been widely used and analyzed by Manton and Vaupel, who collaborated with Stallard, Yashin, Iachine, Begun [3, 6, 7, 8], as well as by Butt and Haberman [9], Moger and Allen [10], Hougaard [11], Finkelstein [12] and Pitacco [13, 14]. The idea of frailty models was expanded as mixed hazard models suggested the hazard rate to have a polynomial expression, see Spreeuw et al. [15], while Assabil [16] analyzed frailty models with time-dependable random effects.

However, in order to be sure that random effect models indeed can be analyzed in survival analysis, all the necessary properties need to be satisfied. If transformed mortality force has (44) expression, then function

$$\widehat{S}(x) = \mathbb{E}\left(S^Z(x)\right) = \mathbb{E}\left(\exp\Big\{-Z\int_0^x \mu_u \mathrm{du}\Big\}\right) \tag{45}$$

should be a new survival function, and

$$\widehat{\mu}(x) = -\frac{\widehat{S}'(x)}{\widehat{S}(x)} = -\ln\left(\mathbb{E}\left(S^Z(x)\right)\right)' \tag{46}$$

should be a new force of mortality. Previously mentioned statements are true if requirements (9)-(12) are fulfilled, or equivalently, when function $\hat{S}(x)$ satisfies equality of type (13). Below we present a theorem which states that function $\hat{S}(x)$, defined by formula (45), is a survival function in the case where $S(x)$ is such.

Theorem 1. *Let $S = S(x)$ be a survival function with a force of mortality $\mu = \mu_x$, and let Z be a positive random variable. Then function $\widehat{S} = \widehat{S}(x)$, defined by equality* (45)*, is a new survival function.*

4.2 Tail function with random effect

From (45) equality we see that random effect transformation can be applied not only to a survival function, which is absolutely continuous, but also to any tail function regardless, if that function is absolutely continuous or not. Based on (45) formula, for any tail function $\overline{F}$ we can define

$$\overline{F^{(Z)}}(x) := \mathbb{E}\Big(\overline{F}^Z(x)\Big) = \mathbb{E}\Big(\big(\overline{F}(x)\big)^Z\Big),\ x \in \mathbb{R} \tag{47}$$

and the d.f. for this transformed t.f. is

$$F^{(Z)}(x) = 1 - \overline{F^{(Z)}}(x) = 1 - \mathbb{E}\Big(\overline{F}^Z(x)\Big),\ x \in \mathbb{R}. \tag{48}$$

The following theorem shows that random effect transformation maintains the main properties of a distribution function.

Theorem 2. *Let $F = F(x)$ be a distribution function, and let Z be a positive random variable. Then function $F^{(Z)}(x)$, defined by equality* (48)*, is a new distribution function.*

Additionally, it turns out that in some cases application of random effect does not affect d.f. properties in terms of specific regularity classes. Below we present a theorem behind this statement.

Theorem 3. *Let F be a distribution function, Z be a positive random variable, and let $F^{(Z)}$ be d.f. with random effect Z defined by equality* (48)*. Then the following statements hold:*

(i) $F \in \mathcal{R}_\alpha \ \Rightarrow\ F^{(Z)} \in \mathcal{R}_{\alpha a}$, *where a is the greatest lower bound of r.v. Z, i.e:*

$$\mathbb{P}(Z \geqslant a) = 1 \textit{ and } \mathbb{P}(Z \geqslant a + \delta) < 1 \textit{ for any } \delta > 0;$$

(ii) $F \in \mathcal{L} \ \Rightarrow\ F^{(Z)} \in \mathcal{L}$;

(iii) $F \in \mathcal{OL} \ \Rightarrow\ F^{(Z)} \in \mathcal{OL}$;

(iv) $F \in \mathcal{D} \ \Rightarrow\ F^{(Z)} \in \mathcal{D}$.

5 Proofs of theorems

5.1 Proof of Theorem 1

The proof can be found in article [55].

Proof. Suppose that survival function S is absolutely continuous with the hazard rate μ_x. We will prove that function $\widehat{S}$ satisfies equality of type (13). According to definition (45), we get:

$$\widehat{S}(x) = \mathbb{E}\left(S^Z(x)\right) = \int_{[0,\infty)} S^z(x)\mathrm{d}F_Z(z)$$
$$= \int_{[0,\infty)} \exp\Big\{ - z\int_0^x \mu_u \mathrm{d}u\Big\}\mathrm{d}F_Z(z)$$

for each $x \geqslant 0$, where F_Z is a distribution function of the random effect Z.

Let us define two real numbers

$$a = \sup\{x \geqslant 0 : S(x) = 1\},$$
$$b = \inf\{x \geqslant 0 : S(x) = 0\}.$$

It is obvious that $\widehat{S}(x) = 1$ if $x \in [0, a]$ and $\widehat{S}(x) = 0$ if $x \in [b, \infty)$ in the case of finite b.

It remains to consider $x \in (a, b)$. For these x we have that $0 < S(x) < 1$. Therefore, for $z \geqslant 0$ and $h > 0$ we have

$$\begin{aligned}
\left|\frac{S^z(x+h) - S^z(x)}{h}\right| &= \frac{S^z(x)}{h}\left|\frac{S^z(x+h)}{S^z(x)} - 1\right| \\
&= \frac{S^z(x)}{h}\left|\exp\left\{-z\int_x^{x+h} \mu_u \mathrm{d}u\right\} - 1\right| \\
&\leqslant z\, S^z(x)\frac{1}{h}\int_x^{x+h} \mu_u \mathrm{d}u \\
&= z\, S^z(x)\frac{1}{h}\ln\frac{S(x)}{S(x+h)}.
\end{aligned}$$

Since S is absolutely continuous, derivative $S'(x)$ exists almost everywhere on interval (a, b). If $S'(x)$ exists, then

$$\ln S(x) - \ln S(x+h) \leqslant \left(1 - \frac{S'(x)}{S(x)}\right) h$$

for sufficiently small $h > 0$, which implies that

$$\left|\frac{S^z(x+h) - S^z(x)}{h}\right| \leqslant zS^z(x)\left(1 - \frac{S'(x)}{S(x)}\right).$$

For $z \geqslant 0$ and sufficiently small $h < 0$ the same estimate can be derived analogously. In addition, it is obvious that integral

$$\int_{[0,\infty)} zS^z(x)\mathrm{d}F_Z(z)$$

is finite if $x \in (a, b)$.

Therefore, due to the Lebesgue's dominated convergence theorem we have that

$$\begin{aligned}
\widehat{S}'(x) &= \int_{[0,\infty)} \lim_{h\to 0} \frac{S^z(x+h) - S^z(x)}{h}\mathrm{d}F_Z(z) \\
&= \int_{[0,\infty)} (S^z(x))'\,\mathrm{d}F_Z(z) \\
&= -\mu_x \int_{[0,\infty)} zS^z(x)\mathrm{d}F_Z(z)
\end{aligned}$$

for almost all $x \in (a, b)$.

Function μ_x is non-negative, bounded and integrable on interval $[a, B]$ with an arbitrary $a < B < b$, while function

$$\int_{[0,\infty)} zS^z(x)\,\mathrm{d}F_Z(z)$$

is continuous on interval (a, b).

Consequently, derivative $\widehat{S}'(x)$ is integrable on the interval $[a, B]$ with $a < B < b$. For $x \in [0, \infty)$, let us define

$$\widehat{f}(x) = -\widehat{S}'(x)\,\mathbb{1}_{(a,b)}(x).$$

Using the Tonelli's theorem we get

$$\int\limits_0^\infty \widehat{f}(x)\mathrm{d}x = \int\limits_a^b (-\widehat{S}'(x))\mathrm{d}x = \int\limits_a^b \Bigg(\int\limits_{[0,\infty)} z\mu_x \mathrm{e}^{-z\int\limits_0^x \mu_u \mathrm{d}u} \mathrm{d}F_Z(z) \Bigg) \mathrm{d}x$$

$$= \int\limits_{[0,\infty)} \Bigg(\int\limits_a^b \mathrm{d}\Big(-\mathrm{e}^{-z\int\limits_0^x \mu_u \mathrm{d}u} \Big) \Bigg) \mathrm{d}F_Z(z)$$

$$= \int\limits_{[0,\infty)} \left(S^z(a) - S^z(b)\right) \mathrm{d}F_Z(z) = 1. \tag{49}$$

Similarly, for $x \in (a, b)$

$$\int\limits_x^\infty \widehat{f}(y)\mathrm{d}y = \int\limits_x^b (-\widehat{S}'(y))\mathrm{d}y = \int\limits_x^b \Bigg(\int\limits_{[0,\infty)} z\mu_y \mathrm{e}^{-z\int\limits_0^y \mu_u \mathrm{d}u} \mathrm{d}F_Z(z) \Bigg) \mathrm{d}y$$

$$= \int\limits_{[0,\infty)} \Bigg(\int\limits_x^b \mathrm{d}\Big(-\mathrm{e}^{-z\int\limits_0^y \mu_u \mathrm{d}u} \Big) \Bigg) \mathrm{d}F_Z(z)$$

$$= \int\limits_{[0,\infty)} \left(S^z(x) - S^z(b)\right) \mathrm{d}F_Z(z) = \widehat{S}(x).$$

It follows from this that

$$\widehat{S}(x) = \int_x^\infty \widehat{f}(y)\mathrm{d}y,\ x \in [0, \infty)$$

for an integrable non-negative function with property (49).

The last equality has the form (13). Consequently, the function is absolutely continuous and so $\widehat{S}$ is a new survival function in the sense of survival analysis. Theorem is proved. □

5.2 Proof of Theorem 2

The proof can be found in article [56].

Proof. Our considerations are based on equation (47). We use known tail function properties to prove that requirements for transformed distribution function are also fulfilled.

• Since $0 \leqslant \overline{F}(x) \leqslant 1$, for all $x \in \mathbb{R}$ we have:

$$F^{(Z)}(x) = 1 - \mathbb{E}\Big(\big(\overline{F}(x)\big)^Z \Big) \geqslant 0,$$
$$F^{(Z)}(x) = 1 - \mathbb{E}\Big(\big(\overline{F}(x)\big)^Z \Big) \leqslant 1.$$

• According to the dominated convergence theorem

$$\begin{aligned}\lim_{x\to-\infty} F^{(Z)}(x) &= 1 - \int_{(0,\infty)} \Big(\lim_{x\to-\infty} \overline{F}(x)\Big)^u \mathrm{d}\mathbb{P}(Z \leqslant u)\\ &= 1 - \mathbb{P}(Z > 0) = 0,\end{aligned}$$

and, similarly, $\lim_{x\to\infty} F^{(Z)}(x) = 1$.

• If $x_1 < x_2$, then $\overline{F}(x_1) \geqslant \overline{F}(x_2)$ imply that

$$\begin{aligned}F^{(Z)}(x_1) &= 1 - \mathbb{E}\Big(\big(\overline{F}(x_1)\big)^Z \Big)\\ &\leqslant 1 - \mathbb{E}\Big(\big(\overline{F}(x_2)\big)^Z \Big) = F^{(Z)}(x_2).\end{aligned}$$

• By using the dominated convergence theorem again, for each $a \in \mathbb{R}$ we get that

$$\begin{aligned}\lim_{x\to a+} F^{(Z)}(x) &= 1 - \int_{(0,\infty)} \Big(\lim_{x\to a+} \overline{F}(x)\Big)^u \mathrm{d}\mathbb{P}(Z \leqslant u)\\ &= 1 - \int_{(0,\infty)} \big(\overline{F}(a)\big)^u \mathrm{d}\mathbb{P}(Z \leqslant u) = F^{(Z)}(a),\end{aligned}$$

and, similarly,

$$\lim_{x\to a-} F^{(Z)}(x) = 1 - \int_{(0,\infty)} \big(\overline{F}(a-)\big)^u \mathrm{d}\mathbb{P}(Z \leqslant u)$$

exists for each $a \in \mathbb{R}$. We can see from the derived relations that function $F^{(Z)}$ satisfies the standard properties (i) - (iv) of d.f., see (2)-(5).

□

5.3 Proof of Theorem 3

The proof can also be found in article [56], written by thesis author, J. Šiaulys and S. Lewkiewicz. We present proofs of all parts of Theorem 3.

Proof of part **(i)** If F is an arbitrary d.f. and $\alpha \geqslant 0$, then

$$F \in \mathcal{R}_\alpha \Leftrightarrow \lim_{x\to\infty} \frac{\overline{F}(xy)}{\overline{F}(x)} = y^{-\alpha},$$

for any fixed $y > 1$.

Since a is the greatest lower bound of Z we can express tail of d.f. with random effect $F^{(Z)}$ by the following integral

$$\overline{F^{(Z)}}(x) = \int\limits_{[a,\infty)} \big(\overline{F}(x)\big)^u \mathrm{d}F_Z(u).$$

Therefore, for an arbitrary fixed $y > 1$ and positive x we get

$$\begin{aligned}
\sup_{v\geqslant x} \frac{\overline{F^{(Z)}}(vy)}{\overline{F^{(Z)}}(v)} &= \sup_{v\geqslant x} \frac{\int\limits_{[a,\infty)} \left(\frac{\overline{F}(vy)}{\overline{F}(v)}\right)^u \big(\overline{F}(v)\big)^u \mathrm{d}F_Z(u)}{\int\limits_{[a,\infty)} \big(\overline{F}(v)\big)^u \mathrm{d}F_Z(u)} \\
&\leqslant \sup_{v\geqslant x} \frac{\int\limits_{[a,\infty)} \left(\frac{\overline{F}(vy)}{\overline{F}(v)}\right)^a \big(\overline{F}(v)\big)^u \mathrm{d}F_Z(u)}{\int\limits_{[a,\infty)} \big(\overline{F}(v)\big)^u \mathrm{d}F_Z(u)} \\
&\leqslant \sup_{v\geqslant x} \left(\frac{\overline{F}(vy)}{\overline{F}(v)}\right)^a.
\end{aligned}$$

This estimate implies that

$$\limsup_{x\to\infty} \frac{\overline{F^{(Z)}}(xy)}{\overline{F^{(Z)}}(x)} \leqslant \limsup_{x\to\infty} \left(\frac{\overline{F}(xy)}{\overline{F}(x)}\right)^a = y^{-a\alpha},\ y > 1,$$

because of $F \in \mathcal{R}_\alpha$.

Consequently, to finish the proof of part (i) of Theorem 3 it is sufficient to derive that

$$\liminf_{x\to\infty} \frac{\overline{F^{(Z)}}(xy)}{\overline{F^{(Z)}}(x)} \geqslant y^{-a\alpha}$$

for each fixed $y > 1$, or equivalently,

$$\limsup_{x\to\infty} \frac{\overline{F^{(Z)}}(xy)}{\overline{F^{(Z)}}(x)} \leqslant y^{-a\alpha} \tag{50}$$

for each $0 < y < 1$.

Let now $y \in (0,1)$ be fixed. For any $\delta > 0$ and any $x > 0$ we have

$$\frac{\overline{F^{(Z)}}(xy)}{\overline{F^{(Z)}}(x)} = \frac{\int\limits_{[a,a+\delta)} \big(\overline{F}(xy)\big)^u \mathrm{d}F_Z(u)}{\int\limits_{[a,\infty)} \big(\overline{F}(x)\big)^u \mathrm{d}F_Z(u)} + \frac{\int\limits_{[a+\delta,\infty)} \big(\overline{F}(xy)\big)^u \mathrm{d}F_Z(u)}{\int\limits_{[a,\infty)} \big(\overline{F}(x)\big)^u \mathrm{d}F_Z(u)}$$
$$:= I_{1\delta} + I_{2\delta}. \tag{51}$$

Since $0 < y < 1$, we have

$$I_{1\delta} \leqslant \frac{\int\limits_{[a,a+\delta)} \big(\overline{F}(xy)\big)^u \mathrm{d}F_Z(u)}{\int\limits_{[a,a+\delta)} \big(\overline{F}(x)\big)^u \mathrm{d}F_Z(u)} \leqslant \left(\frac{\overline{F}(xy)}{\overline{F}(x)}\right)^{a+\delta}.$$

For the second term in (51), we obtain

$$\begin{aligned}
I_{2\delta} &\leqslant \frac{\int\limits_{[a+\delta,\infty)} \big(\overline{F}(xy)\big)^u \mathrm{d}F_Z(u)}{\int\limits_{[a,a+\delta/2)} \big(\overline{F}(x)\big)^u \mathrm{d}F_Z(u)} \\
&\leqslant \frac{\big(\overline{F}(xy)\big)^{a+\delta}\mathbb{P}(Z \geqslant a+\delta)}{\big(\overline{F}(x)\big)^{a+\delta/2}\,\mathbb{P}(a \leqslant Z < a+\delta/2)} \\
&= \left(\frac{\overline{F}(xy)}{\overline{F}(x)}\right)^{a+\delta/2} \big(\overline{F}(xy)\big)^{\delta/2}\, \frac{\mathbb{P}(Z \geqslant a+\delta)}{\mathbb{P}(a \leqslant Z < a+\delta/2)}.
\end{aligned}$$

Consequently, for each $\delta > 0$ we get

$$\begin{aligned}
\limsup_{x\to\infty} \frac{\overline{F^{(Z)}}(xy)}{\overline{F^{(Z)}}(x)} &\leqslant \limsup_{x\to\infty} \left(\frac{\overline{F}(xy)}{\overline{F}(x)}\right)^{a+\delta} \\
&\quad + \limsup_{x\to\infty} \left(\frac{\overline{F}(xy)}{\overline{F}(x)}\right)^{a+\delta/2} \frac{\big(\overline{F}(xy)\big)^{\delta/2}\,\mathbb{P}(Z \geqslant a+\delta)}{\mathbb{P}(a \leqslant Z < a+\delta/2)} \\
&= y^{-a\alpha-\alpha\delta},
\end{aligned}$$

because $F \in \mathcal{R}_\alpha$ and a is the greatest lower bound of the random effect Z. By passing δ to zero, we obtain the required inequality (50) for each fixed $y \in (0,1)$. The first part (i) of the Theorem 3 is proved. $\square$

***Proof of part* (ii)** From definition of the class $\mathcal{L}$ we have

$$F \in \mathcal{L} \;\;\Leftrightarrow\;\; \limsup_{x\to\infty} \frac{\overline{F}(x-1)}{\overline{F}(x)} \;\;\leqslant\;\; 1. \tag{52}$$

Let us choose $\varepsilon \in (0, \frac{1}{2})$ such that $\mathbb{P}\Big(Z \in \big[0, \frac{1}{\varepsilon}\big)\Big) > 0$. For sufficiently large x, say $x \geqslant x_2$, the following inequality holds:

$$\frac{\overline{F}(x-1)}{\overline{F}(x)} \leqslant 1 + \varepsilon^2.$$

If $x \geqslant x_2$, then we get

$$\begin{aligned}
\frac{\overline{F^{(Z)}}(x-1)}{\overline{F^{(Z)}}(x)} &= \frac{\int\limits_{[0,1/\varepsilon)} \big(\overline{F}(x-1)\big)^u \mathrm{d}F_Z(u)}{\int\limits_{[0,\infty)} \big(\overline{F}(x)\big)^u \mathrm{d}F_Z(u)} + \frac{\int\limits_{[1/\varepsilon,\infty)} \big(\overline{F}(x-1)\big)^u \mathrm{d}F_Z(u)}{\int\limits_{[0,\infty)} \big(\overline{F}(x)\big)^u \mathrm{d}F_Z(u)} \\
&\leqslant \max_{0\leqslant u\leqslant 1/\varepsilon} \left(\frac{\overline{F}(x-1)}{\overline{F}(x)}\right)^u + \frac{\int\limits_{[1/\varepsilon,\infty)} \big(\overline{F}(x-1)\big)^u \mathrm{d}F_Z(u)}{\int\limits_{[0,1/\varepsilon)} \big(\overline{F}(x)\big)^u \mathrm{d}F_Z(u)} \\
&\leqslant (1+\varepsilon^2)^{1/\varepsilon} + \frac{\big(\overline{F}(x-1)\big)^{1/\varepsilon}\,\mathbb{P}(Z \geqslant 1/\varepsilon)}{\big(\overline{F}(x)\big)^{1/\varepsilon}\,\mathbb{P}(Z \in [0, 1/\varepsilon))} \\
&\leqslant \left((1+\varepsilon^2)^{1/\varepsilon}\left(1 + \frac{\mathbb{P}(Z \geqslant 1/\varepsilon)}{\mathbb{P}(Z \in [0, 1/\varepsilon))}\right)\right).
\end{aligned}$$

Since ε can be as close to zero as desired, we get from the above inequality that

$$\limsup_{x\to\infty} \frac{\overline{F^{(Z)}}(x-1)}{\overline{F^{(Z)}}(x)} \leqslant 1,$$

which implies that $F^{(Z)} \in \mathcal{L}$. Part (ii) of the Theorem 3 is proved. □

***Proof of part* (iii)** For an arbitrary d.f. F we have

$$F \in \mathcal{OL} \;\;\Leftrightarrow\;\; \limsup_{x\to\infty} \frac{\overline{F}(x-1)}{\overline{F}(x)} < \infty \;\;\Leftrightarrow\;\; \sup_{x\in\mathbb{R}} \frac{\overline{F}(x-1)}{\overline{F}(x)} < \infty.$$

Therefore, for each positive K under condition $\mathbb{P}\big(Z \in [0,K)\big) \geqslant 1/2$, we get

$$\begin{aligned}
\sup_{x\in\mathbb{R}} \frac{\overline{F^{(Z)}}(x-1)}{\overline{F^{(Z)}}(x)} &\leqslant \sup_{x\in\mathbb{R}} \frac{\int\limits_{[0,K)} \big(\overline{F}(x-1)\big)^u \mathrm{d}F_Z(u)}{\int\limits_{[0,\infty)} \big(\overline{F}(x)\big)^u \mathrm{d}F_Z(u)} \\
&+ \sup_{x\in\mathbb{R}} \frac{\int\limits_{[K,\infty)} \big(\overline{F}(x-1)\big)^u \mathrm{d}F_Z(u)}{\int\limits_{[0,K)} \big(\overline{F}(x)\big)^u \mathrm{d}F_Z(u)} \\
&\leqslant \left(\sup_{x\in\mathbb{R}} \frac{\overline{F}(x-1)}{\overline{F}(x)}\right)^K + 2\mathbb{P}(Z \geqslant K) \sup_{x\in\mathbb{R}} \left(\frac{\overline{F}(x-1)}{\overline{F}(x)}\right)^K \\
&< \infty.
\end{aligned}$$

This means that $F^{(Z)} \in \mathcal{OL}$. Part (iii) of the Theorem 3 is proved. □

***Proof of part* (iv)** From definition of the class $\mathcal{D}$ we have that

$$F \in \mathcal{D} \;\Leftrightarrow\; \limsup_{x\to\infty} \frac{\overline{F}(\frac{x}{2})}{\overline{F}(x)} < \infty \tag{53}$$

for an arbitrary d.f. F. Let's choose positive K such that $\mathbb{P}\big(Z \in [0,K)\big) > 0$. For this K we get

$$\begin{aligned}
\frac{\overline{F^{(Z)}}(\frac{x}{2})}{\overline{F^{(Z)}}(x)} &= \frac{\int\limits_{[0,K)} \big(\overline{F}(\frac{x}{2})\big)^u \mathrm{d}F_Z(u)}{\int\limits_{[0,\infty)} \big(\overline{F}(x)\big)^u \mathrm{d}F_Z(u)} + \frac{\int\limits_{[K,\infty)} \big(\overline{F}(\frac{x}{2})\big)^u \mathrm{d}F_Z(u)}{\int\limits_{[0,\infty)} \big(\overline{F}(x)\big)^u \mathrm{d}F_Z(u)} \\
&:= J_{1K} + J_{2K}.
\end{aligned} \tag{54}$$

Since $F \in \mathcal{D}$, it follows from (53) that

$$\frac{\overline{F}(\frac{x}{2})}{\overline{F}(x)} \leqslant c_1$$

for some constant $c_1 \geqslant 1$ if x is sufficiently large, say $x \geqslant x_1$. Therefore,

$$J_{1K} \leqslant \frac{\int\limits_{[0,K)} \frac{\big(\overline{F}(\frac{x}{2})\big)^u}{\big(\overline{F}(x)\big)^u} \big(\overline{F}(x)\big)^u \mathrm{d}F_Z(u)}{\int\limits_{[0,K)} \big(\overline{F}(x)\big)^u \mathrm{d}F_Z(u)} \leqslant \max_{0\leqslant u\leqslant K} \left(\frac{\overline{F}(\frac{x}{2})}{\overline{F}(x)}\right)^u \leqslant c_1^K,\; x \geqslant x_1. \tag{55}$$

When $x \geqslant x_1$, we get

$$J_{2K} \leqslant \frac{\int\limits_{[K,\infty)} \big(\overline{F}(\frac{x}{2})\big)^u \mathrm{d}F_Z(u)}{\int\limits_{[0,K)} \big(\overline{F}(x)\big)^u \mathrm{d}F_Z(u)} \leqslant \frac{\big(\overline{F}(\frac{x}{2})\big)^K \, \mathbb{P}(Z \geqslant K)}{\big(\overline{F}(x)\big)^K \mathbb{P}(Z \in [0,K))}$$
$$\leqslant c_1^K \frac{\mathbb{P}(Z \geqslant K)}{\mathbb{P}(Z \in [0,K))}. \tag{56}$$

From estimates (55) and (56) we derive

$$\limsup_{x\to\infty} \frac{\overline{F^{(Z)}}(\frac{x}{2})}{\overline{F^{(Z)}}(x)} \leqslant c_1^K \left(1 + \frac{\mathbb{P}(Z \geqslant K)}{\mathbb{P}(Z \in [0,K))}\right),$$

which implies that $F^{(Z)} \in \mathcal{D}$. Part (iv) of Theorem 3 is proved. □

5.4 Some remarks

Remark 1. Theorem 1 provides us a large set of survival functions with a random effect. A new survival function can be constructed from any survival function S that meets its requirements (9)-(12) and any positive random variable Z. The few examples below show that the random effect preserves absolute continuity of the initial survival function but significantly changes its type.

For instance, let us consider the exponential survival function

$$S(x) = \mathrm{e}^{-0.03x}, \; x \geqslant 0.$$

According to Theorem 1, function $\widehat{S}(x) = \mathbb{E}S^Z(x)$ is a new absolutely continuous survival function in the case of the random effect Z under condition $\mathbb{P}(Z > 0) = 1$.

In particular, if Z is uniformly distributed on the interval $[a, b]$ with $0 \leqslant a < b < \infty$, then for $x \geqslant 0$

$$\widehat{S}(x) = \frac{1}{b-a}\int_a^b e^{-0.03xz} dz = \frac{300}{(b-a)x}(e^{-0.003ax} - e^{-0.03bx}).$$

If the random effect Z has exponential distribution with positive parameter λ

$$\mathbb{P}(Z \leqslant z) = (1 - e^{-\lambda z})\mathbb{1}_{[0,\infty)}(z),$$

then the function

$$\widehat{S}(x) = \lambda \int_0^{\infty} \mathrm{e}^{-(0.03x+\lambda)z} \mathrm{d}z = \frac{\lambda}{0.03x+\lambda},\ x \geqslant 0,$$

is a Pareto-type survival function.

If the random effect Z has the Bernoulli distribution

$$\mathbb{P}(Z=1) = (1-p),\ \mathbb{P}(Z=2) = p,\ p \in (0,1),$$

then the function with random effect

$$\widehat{S}(x) = \mathrm{e}^{-0.03x}(1-p) + \mathrm{e}^{-0.06x}p,\ x \geqslant 0,$$

is a mixture of exponential survival functions.

If the random effect Z distributed according to the shifted Poisson law with positive parameter λ, then

$$\mathbb{P}(Z=k) = \mathrm{e}^{-\lambda} \frac{\lambda^{k-1}}{(k-1)!},\ k \in \mathbb{N},$$

and, consequently,

$$\widehat{S}(x) = \mathrm{e}^{-\lambda} \sum_{k=1}^{\infty} (S(x))^k \frac{\lambda^{k-1}}{(k-1)!} = \mathrm{e}^{-0.03x} \mathrm{e}^{\lambda\left(\mathrm{e}^{-0.03x}-1\right)},\ x \geqslant 0.$$

Finally, if Z has the classical Peter and Paul distribution

$$\mathbb{P}(Z=2^k) = \frac{1}{2^k},\ k \in \mathbb{N},$$

then

$$\widehat{S}(x) = \sum_{k=1}^{\infty} \frac{1}{2^k} \mathrm{e}^{-2^k 0.03x},\ x \geqslant 0,$$

is the infinite mixture of exponential survival functions.

Remark 2. In the case of the integer valued random effect Z, the tail function $\overline{F^{(Z)}}$ is the tail function of a randomly stopped minimum, considered, for instance, in [24], [25], [26].

Suppose that $\{X_1, X_2, \ldots\}$ are independent copies of the r.v. X with d.f.

F and tail function $\overline{F}$. In the case of $\mathbb{P}(Z=1)=1$, we have that

$$\overline{F^{(Z)}}(x)=\mathbb{E}\left(\overline{F}(x)^Z\right)=\mathbb{P}(X_1>x).$$

If $\mathbb{P}(Z=2)=1$, then

$$\overline{F^{(Z)}}(x)=\mathbb{P}(X_1>x)\mathbb{P}(X_2>x)=\mathbb{P}\big(\min\{X_1,X_2\}>x\big).$$

If $\mathbb{P}(Z=3)=1$, then

$$\overline{F^{(Z)}}(x)=\mathbb{P}\big(\min\{X_1,X_2,X_3\}>x\big).$$

Finally, in the case of integer valued random effect Z such that $\mathbb{P}(Z\in\mathbb{N})=1$, we have

$$\begin{aligned}\overline{F^{(Z)}}(x)&=\sum_{k=1}^{\infty}\mathbb{P}\big(\min\{X_1,X_2,\ldots,X_k\}>x\big)\mathbb{P}(Z=k)\\&=\mathbb{P}\big(\min\{X_1,X_2,\ldots,X_Z\}>x\big),\end{aligned}$$

if random effect Z and the collection of random variables $\{X_1,X_2,\ldots\}$ are independent.

Remark 3. Theorem 1 justifies the use of a random effect in demographics to find the expression of the mortality force that fits empirical data the best. It is important to note that the random effect applies not only to the transformations of survival but also to other models used in various studies, for instance, in medical research (see [28, 29, 30, 31]) or statistical analysis of certain problems (see [32, 33, 34, 35, 36, 37]). In [38, 39], probabilistic objects with random effects are examined.

5.5 Several examples for Theorem 3.

Below we present some examples of transformed survival functions and their behaviour in terms of regularity classes. These examples illustrate Theorem 3 from Section 4.2.

Example 1. Let F be the Burr d.f. with parameters $c=1$ and $k=2$, i.e.

$$\overline{F}(x)=\frac{1}{(1+x)^2},\ x\geqslant 0,$$

and suppose that random effect Z has the uniform distribution, i.e. $Z \sim U[0,1]$. It is easy to find that in the case under consideration

$$\overline{F^{(Z)}}(x) = \mathbb{E}\Big(\big(\overline{F}(x)\big)^Z\Big) = \int\limits_0^1 \left(\frac{1}{(1+x)^2}\right)^u \mathrm{d}u = \frac{x(2+x)}{2(1+x)^2\ln(1+x)},\ x \geqslant 0.$$

Hence $F \in \mathcal{R}_2$ and $F^{(Z)} \in \mathcal{R}_0$ what is compatible with the statement of the Theorem 3(i). Additionally, since it was already proven earlier that $F \in \mathcal{L}$ (see Example 4 of 3.4) and $F^{(Z)} \in \mathcal{L}$, the example is also compatible with the statement of the Theorem 3(ii).

The differences of two tails $\overline{F}$ and $\overline{F^{(Z)}}$ are shown in the Figure 5 below.

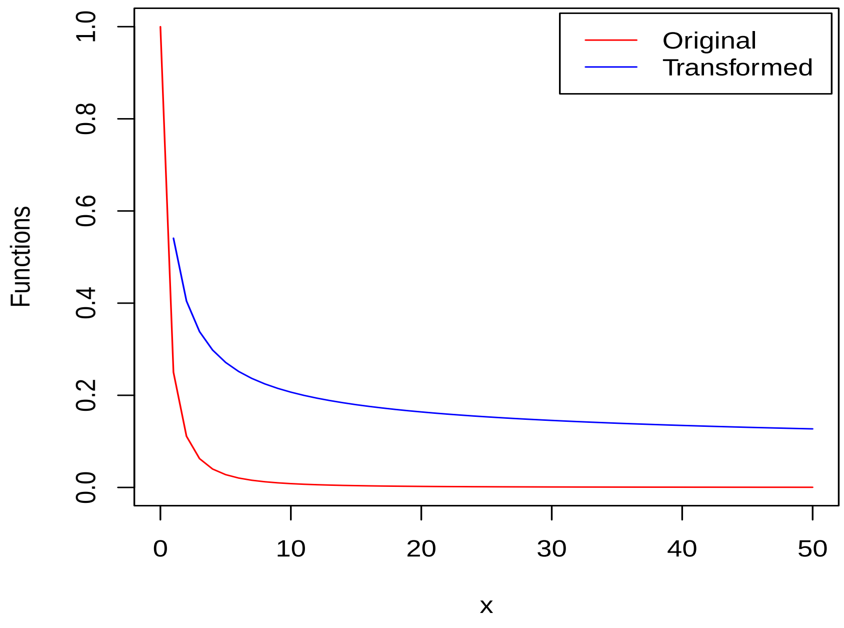

Figure 5: Comparison of tails $\overline{F}$ and $\overline{F^{(Z)}}$ from Example 1.

Example 2. Let F be the Peter and Paul d.f. with parameters $a = 1$ and $b = 3$. According to [74], the tail of such d.f., when $x \geqslant 1$ has the form

$$\overline{F}(x) = 3^{-\lfloor \log_3 x \rfloor}.$$

In addition, let's assume that random effect Z obtains values $\{1, 2, 3\}$ with equal probabilities.

In the described case,

$$\overline{F^{(Z)}}(x) = 3^{-1-\lfloor \log_3(x) \rfloor}\left(1 + 3^{-\lfloor \log_3(x) \rfloor} + 3^{-2\lfloor \log_3(x) \rfloor}\right),\ x \geqslant 1.$$

It is easy to check that in this case F and $F^{(Z)}$ belong to the class $\mathcal{D}$, which is compatible with the statement of the Theorem 3(iv). The tails of the functions F and $F^{(Z)}$ are plotted in the Figure 6.

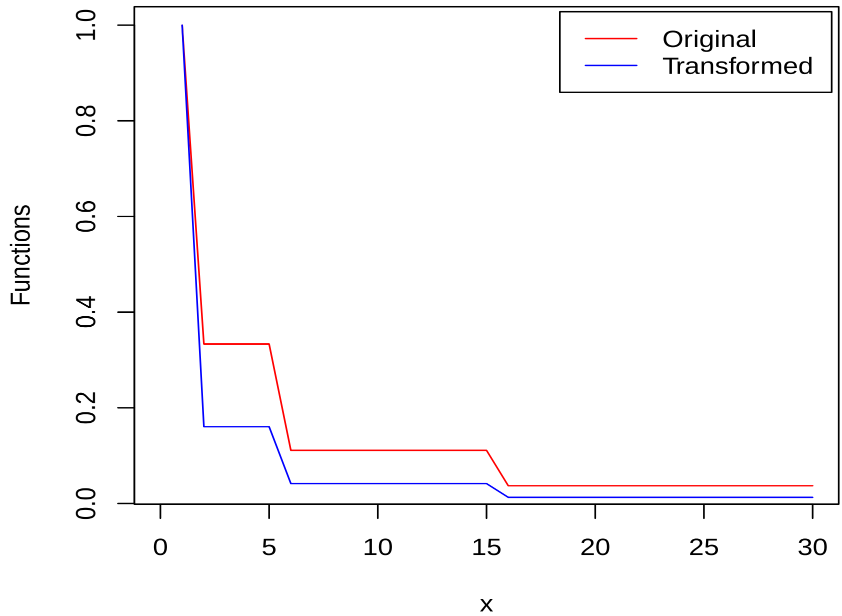

Figure 6: Comparison of tails $\overline{F}$ and $\overline{F^{(Z)}}$ from Example 2.

Example 3. Let us consider the Weibull distribution with parameters $\alpha = -1/2$ and $a = 1/2$. For such distribution

$$\overline{F}(x) = \mathrm{e}^{-\sqrt{x}},\ x \geqslant 0.$$

In addition, let us suppose that random effect in this example has the exponential distribution with parameter $\lambda = 1$, i.e. $Z \sim \operatorname{Exp}(1)$.

In the case under consideration,

$$\overline{F^{(Z)}}(x) = \int_0^\infty (\mathrm{e}^{-\sqrt{x}})^u \mathrm{e}^{-u} \mathrm{d}u = \frac{1}{\sqrt{x}+1},\ x \geqslant 0.$$

It is easy to see that in this case the tail of the d.f. $F^{(Z)}$ is significantly heavier than the tail $\overline{F}$. However, both functions F and $F^{(Z)}$ belong to classes $\mathcal{L} \subset \mathcal{OL}$, which is consistent with statements (ii) and (iii) Theorem 3. The graphs of t.f. $\overline{F}$ and $\overline{F^{(Z)}}$ of this example can be seen in Figure 7.

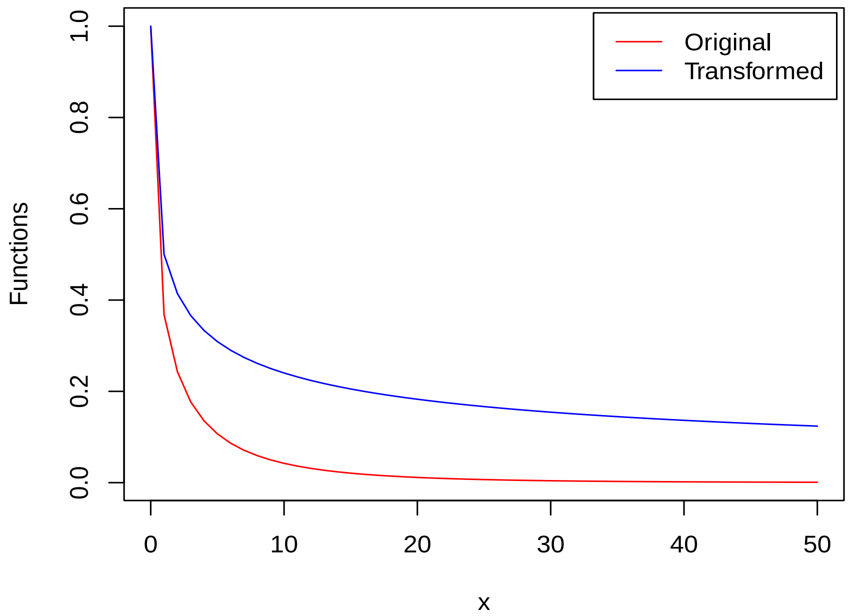

Figure 7: Comparison of tails $\overline{F}$ and $\overline{F^{(Z)}}$ from Example 3.

6 Some special survival functions

According to Theorem 1, we can construct new survival functions having a basic force of mortality μ_x and a positive random variable. In this section, several examples of popular survival functions and corresponding hazard rates will be presented. From each selected survival function, a new survival functions using well-known random effects will be constructed. For each selected pair of survival function and random effect, analytical expression of the new survival function and the analytical formula of the corresponding hazard rate is found.

Below we present a couple of examples of modified mortality models where random effect is applied to well-known Weibull and Gompertz hazard rate functions, which have been analyzed by many researchers in the past, including Juckett and Rosenberg [1] and Missov [2], among others.

6.1 Gamma - Weibull model

At the beginning, let us consider the Weibull force of mortality

$$\mu_x = ax^{\alpha}, \quad x \geqslant 0,$$

depending on two positive parameters a and α. By choosing a Gamma distributed random variable $Z \sim \Gamma(k, \lambda)$ for a random effect, we get the Gamma - Weibull model described in [40] and [41], among others.

Gamma distributed random variable $Z \sim \Gamma(k, \lambda)$ has the density

$$f_Z(x) = \frac{\lambda^k}{\Gamma(k)} x^{k-1} \mathrm{e}^{-\lambda x}, \quad x \geqslant 0,$$

where k and λ are positive parameters and $\Gamma(k) = \int\limits_0^\infty t^{k-1} e^{-t} dt$ is the standard Gamma function.

The information above implies the following expression of Gamma - Weibull survival function

$$\hat{S}_{GW}(x) = \mathbb{E}\left(\mathrm{e}^{-Z\int\limits_0^x \mu_t \mathrm{d}t}\right) = \int\limits_0^\infty \mathrm{e}^{-z\int\limits_0^x at^\alpha \mathrm{d}t} \frac{\lambda^k}{\Gamma(k)} z^{k-1}\mathrm{e}^{-\lambda z}\mathrm{d}z$$

$$= \frac{\lambda^k}{\Gamma(k)} \int\limits_0^\infty \mathrm{e}^{-z\left(\int\limits_0^x at^\alpha \mathrm{d}t+\lambda\right)} z^{k-1}\mathrm{d}z.$$

By denoting $w = z\left(\int\limits_0^x at^\alpha dt + \lambda\right)$, we get

$$\hat{S}_{GW}(x) = \frac{\lambda^k}{\Gamma(k)} \frac{1}{\left(\int\limits_0^x at^\alpha \mathrm{d}t + \lambda\right)^k} \int\limits_0^\infty \mathrm{e}^{-w} w^{k-1}\mathrm{d}w$$

$$= \frac{\lambda^k}{\left(\frac{ax^{\alpha+1}}{\alpha+1} + \lambda\right)^k} = \left(\frac{ax^{\alpha+1}}{(\alpha+1)\lambda} + 1\right)^{-k}.$$

It is clear that the obtained survival function has derivative

$$\hat{S}'_{GW}(x) = -\frac{akx^\alpha}{\lambda}\left(\frac{ax^{\alpha+1}}{\lambda(\alpha+1)} + 1\right)^{-k-1}.$$

By using (46) formula, we obtain the following force of mortality expression for the Gamma - Weibull model

$$\hat{\mu}_x = -\frac{\widehat{S}'_{GW}(x)}{\widehat{S}_{GW}(x)} = \frac{akx^\alpha}{\frac{ax^{\alpha+1}}{\alpha+1} + \lambda}. \tag{57}$$

6.2 Gamma - Gompertz model

In the Gamma – Gompertz model, it is assumed that the basic force of mortality has the Gompertz expression, i.e.

$$\mu_x = B\mathrm{e}^{\alpha x}, \quad x \geqslant 0, \tag{58}$$

with positive parameters B and α. This expression can be derived from the Gompertz - Makeham model (see [42], [43], [44] and [45], among others). It should be noted that Gompertz force of mortality belongs to Perk's family of hazard rate functions and assumes that mortality increases exponentially with age [46], [47].

In the Gamma - Gompertz model random effect Z has the Gamma distribution, identical as in the Gamma – Weibull model described in 6.1. Hence, in order to find the expression of model's survival function, identical calculations can be used to those that were performed while analyzing the Gamma – Weibull model. The only difference – expression in the integral ct^n should be changed by Gompertz expression $Be^{\alpha t}$. After the detailed calculations we get the following expressions

$$\hat{S}_{GG_4}(x) = \left(1 + \frac{B}{\alpha\lambda}(\mathrm{e}^{\alpha x} - 1)\right)^{-k},$$

$$\hat{S}'_{GG_4}(x) = -k\frac{B}{\lambda}\mathrm{e}^{\alpha x}\left(1 + \frac{B}{\alpha\lambda}(\mathrm{e}^{\alpha x} - 1)\right)^{-k-1},$$
$$\hat{\mu}_x = -\frac{\hat{S}'_{GG_4}(x)}{\hat{S}_{GG_4}(x)} = \frac{k\frac{B}{\lambda}\mathrm{e}^{\alpha x}}{\left(1 + \frac{B}{\alpha\lambda}(\mathrm{e}^{\alpha x} - 1)\right)}. \tag{59}$$

We can see that the above survival function $\hat{S}_{GG_4}$ and force of mortality $\widehat{\mu}_x$ depend on four parameters. In addition to this general case, a separate version of the Gamma - Gompertz model with three parameters can be considered which is obtained by supposing $k = \lambda$. It is obvious that for the Gamma - Gompertz model with three parameters we have:

$$\hat{S}_{GG_3}(x) = \left(1 + \frac{B}{\alpha\lambda}(\mathrm{e}^{\alpha x} - 1)\right)^{-\lambda},\ \hat{\mu}_x = \frac{B\mathrm{e}^{\alpha x}}{\left(1 + \frac{B}{\alpha\lambda}(\mathrm{e}^{\alpha x} - 1)\right)}. \tag{60}$$

6.3 Poisson - Gompertz model

In the Poisson – Gompertz model force of mortality function has the Gompertz expression (58), while random effect Z has the shifted Poisson distribution with parameter $\lambda > 0$, i.e.

$$\mathbb{P}(Z = k) = \mathrm{e}^{-\lambda}\frac{\lambda^{k-1}}{(k-1)!},\ k \in \mathbb{N} = \{1, 2, \ldots\}.$$

In the case of the integer-valued random variable Z, the expression of the survival function (45) gets the following form

$$\hat{S}(x) = \sum_{k=1}^{\infty} \mathbb{P}(Z = k) \exp\Big\{ - k \int_0^x \mu_t \mathrm{d}t \Big\}. \tag{61}$$

Therefore, for the Poisson - Gompertz model survival function has the following expression:

$$\begin{aligned}
\hat{S}_{PG}(x) &= \mathrm{e}^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \mathrm{e}^{-kB \int_0^x \mathrm{e}^{\alpha t} \mathrm{d}t} \\
&= \exp\Big\{ - \lambda - B \int_0^x \mathrm{e}^{\alpha t} \mathrm{d}t \Big\} \sum_{k=1}^{\infty} \frac{1}{(k-1)!} \left(\lambda \mathrm{e}^{-B \int_0^x \mathrm{e}^{\alpha t} \mathrm{d}t} \right)^{k-1} \\
&= \exp\left\{ \frac{B}{\alpha}(1 - \mathrm{e}^{\alpha x}) + \lambda \left(\mathrm{e}^{\frac{B}{\alpha}(1-\mathrm{e}^{\alpha x})} - 1 \right) \right\}
\end{aligned}$$

with positive parameters α, B and λ.

For the derivative of survival function we obtain

$$\begin{aligned}
\hat{S}'_{PG}(x) &= \hat{S}_{PG}(x) \left(\frac{B}{\alpha}(1 - \mathrm{e}^{\alpha x})' + \lambda \left(\mathrm{e}^{\frac{B}{\alpha}(1-\mathrm{e}^{\alpha x})} \right)' \right) \\
&= -\hat{S}_{PG}(x) B \mathrm{e}^{\alpha x} \left(\lambda \mathrm{e}^{\frac{B}{\alpha}(1-\mathrm{e}^{\alpha x})} + 1 \right).
\end{aligned}$$

Consequently, force of mortality for Poisson - Gompertz model is the following:

$$\widehat{\mu}_x = B \mathrm{e}^{\alpha x} \left(\lambda \mathrm{e}^{\frac{B}{\alpha}(1-\mathrm{e}^{\alpha x})} + 1 \right). \tag{62}$$

6.4 Geometric - Gompertz model

In the Geometric – Gompertz model, mortality force function has the Gompertz expression (58), and the random effect Z has the shifted Geometric distribution, i.e.

$$\mathbb{P}(Z = k) = p(1-p)^{k-1},\ k \in \mathbb{N}.$$

Since random variable Z is discrete, by using equality (61), we obtain

$$\begin{aligned}
\hat{S}_{GEG}(x) &= \sum_{k=1}^{\infty} p(1-p)^{k-1} \mathrm{e}^{-k \int_0^x B \mathrm{e}^{\alpha t} \mathrm{d}t} \\
&= p \mathrm{e}^{-B \int_0^x \mathrm{e}^{\alpha t} \mathrm{d}t} \sum_{k=1}^{\infty} \left((1-p) \mathrm{e}^{-B \int_0^x \mathrm{e}^{\alpha t} \mathrm{d}t} \right)^{k-1} \\
&= \frac{p \mathrm{e}^{\frac{B}{\alpha}(1-\mathrm{e}^{\alpha x})}}{1 - (1-p) \mathrm{e}^{\frac{B}{\alpha}(1-\mathrm{e}^{\alpha x})}}
\end{aligned}$$

with parameters $\alpha > 0$, $B > 0$ and $p \in (0,1)$.

Derivative of survival function is

$$\hat{S}'_{GEG}(x) = -\frac{pB\mathrm{e}^{\alpha x}\mathrm{e}^{\frac{B}{\alpha}(1-\mathrm{e}^{\alpha x})}}{\left(1-(1-p)\mathrm{e}^{\frac{B}{\alpha}(1-\mathrm{e}^{\alpha x})}\right)^2}.$$

and the hazard rate of the model has following expression:

$$\widehat{\mu}_x = \frac{B\mathrm{e}^{\alpha x}}{1-(1-p)\mathrm{e}^{\frac{B}{\alpha}(1-\mathrm{e}^{\alpha x})}}. \tag{63}$$

6.5 Discrete - Weibull model

In the Discrete - Weibull case, differently than in Subsection 6.1, we suppose that force of mortality has the Weibull expression with modal age of death, i.e. we suppose that

$$\mu_x = \frac{1}{\sigma}\left(\frac{x}{M}\right)^{\frac{M}{\sigma}-1}$$

with positive parameters M and σ.

Usually the parameter M is called the modal age of death, because at this age population has the largest number of deaths, for details see [2], [51], [52], [53], [54].

In the Discrete - Weibull model, random effect Z is supposed to be discrete with finite support. We consider the case when random effect Z acquires three different values $1, 2, 3$. More precisely, we consider the Three - Point - Discrete - Weibull model with Z having distribution $\mathbb{P}(Z=1)=p$, $\mathbb{P}(Z=2)=q$, $\mathbb{P}(Z=3)=1-p-q$, where $p, q \in [0,1]$ and $p+q \leqslant 1$.

For the model under consideration, by using expression (61) we obtain

$$\hat{S}_{DW}(x) = p\,\mathrm{e}^{-\left(\frac{x}{M}\right)^{M/\sigma}} + q\,\mathrm{e}^{-2\left(\frac{x}{M}\right)^{M/\sigma}} + (1-p-q)\mathrm{e}^{-3\left(\frac{x}{M}\right)^{M/\sigma}}.$$

Since

$$\hat{S}'_{DW}(x) = -\frac{1}{\sigma}\left(\frac{x}{M}\right)^{M/\sigma-1}\left(p\,\mathrm{e}^{-\left(\frac{x}{M}\right)^{M/\sigma}} + 2q\,\mathrm{e}^{-2\left(\frac{x}{M}\right)^{M/\sigma}} + 3(1-p-q)\mathrm{e}^{-3\left(\frac{x}{M}\right)^{M/\sigma}}\right)$$

we derive that

$$\hat{\mu}_x = \frac{1}{\sigma}\left(\frac{x}{M}\right)^{M/\sigma-1}\frac{p\,\mathrm{e}^{-\left(\frac{x}{M}\right)^{M/\sigma}} + 2q\,\mathrm{e}^{-2\left(\frac{x}{M}\right)^{M/\sigma}} + 3(1-p-q)\mathrm{e}^{-3\left(\frac{x}{M}\right)^{M/\sigma}}}{p\,\mathrm{e}^{-\left(\frac{x}{M}\right)^{M/\sigma}} + q\,\mathrm{e}^{-2\left(\frac{x}{M}\right)^{M/\sigma}} + (1-p-q)\mathrm{e}^{-3\left(\frac{x}{M}\right)^{M/\sigma}}}. \tag{64}$$

We note that when p is equal to 1 the Discrete - Weibull force of mortality (64) becomes the free Weibull force of mortality.

7 Applications to real mortality data of Baltic countries

In this section we apply the theoretical part to actual data - mortality models are applied to different data sets of Baltic population to evaluate survival probabilities and determine which of the presented models with random effect describe mortality and morbidity data the best.

We analyze the mortality data of Lithuanian, Latvian and Estonian populations in the XXI century. Chosen datasets were obtained from the Human Mortality Database. For simplicity purposes, we are going to use the unisex data, i.e. mortality for men and women will not be covered separately. Although it is obvious that mortality of genders differs, we must take into consideration the fact that in life insurance business companies are prohibited from using gender as a risk factor when deciding life insurance premium, therefore, same mortality tables are used for men and women. Our dataset will cover people aged from 0 to 110. Since there is very little exposure of people older than 110, all of these individuals were added to the age group of 110. The dataset contains both survival probabilities as well as central hazard rates for each group. Mortality tables are usually constructed by observing many years of data, therefore, we used average central hazard rates and survival probabilities recorded since 2000 and not rely only on one year of obtained information.

The sample of the dataset is provided in Table 5 in the Appendix. In this table $\widetilde{S}(x)$ denotes the empirical survival probabilities and $\widetilde{m}_x$ denotes the empirical central mortality rate values for each age $x \in \{0, 1, ..., 110\}$.

Moreover, we use previously mentioned assumption of mortality force being constant in the interval $[x, x+1)$. This means that $\mu_{x+t} = \mu_x$ for any $x \in \mathbb{N}_0 = \{0, 1, ...\}$ and $t \in [0, 1)$. Under such restriction the empirical central mortality rate can be compared with the empirical hazard rate, i.e. $\widetilde{m}_x = \widetilde{\mu}_x,\ x \in \mathbb{N}_0$.

Application of mortality models to each population was performed using calculations in programming language R with the help of package *Mortality-Laws* - based on chosen loss function, best estimate paramaters of each mortality model were calculated for each population separately by providing empirical hazard rates as an input. In this analysis, we will consider one of the most traditional and most often used metrics (loss functions) to estimate model parameters - the mean squared error:

$$MSE = \frac{\sum_{x=1}^{N} (\widetilde{m}(x) - \hat{\mu}(x))^2}{N},$$

where $\widetilde{m}(x)$ - empirical hazard rate, $\hat{\mu}(x)$ - modelled hazard rate and N - number of observations. In our case, $N = 110$.

We also use MSE when comparing different mortality models.

Comparison of models described in Section 6 for each country along with estimated parameters, MSE values and hazard rate graphs are provided below.

In subsection 7.1 we present results for Lithuanian population.

In subsection 7.2 we present results for Latvian population.

In subsection 7.3 we present results for Estonian population.

In subsection 7.4 we present conclusion of the analysis.

7.1 Lithuanian mortality results

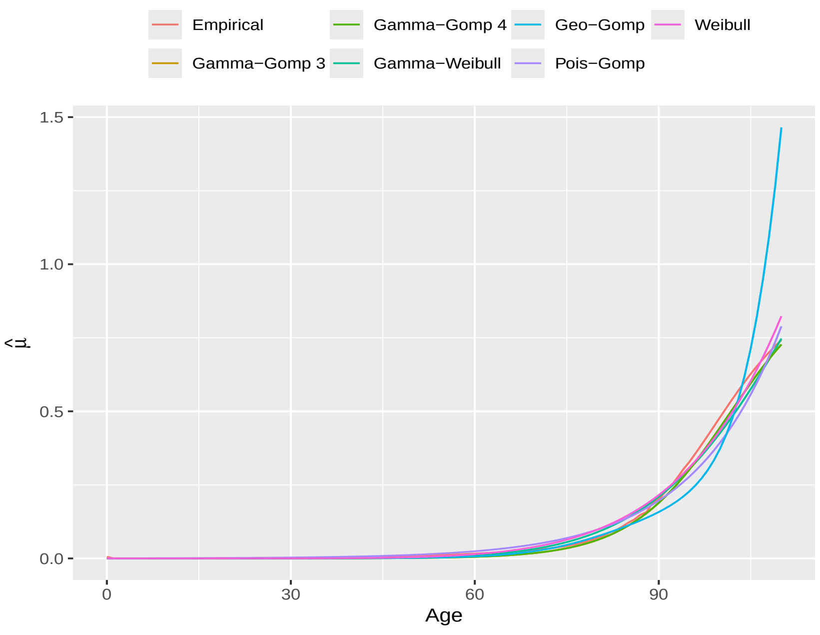

Figure 8: Mortality models with random effect applied to Lithuanian data (all ages)

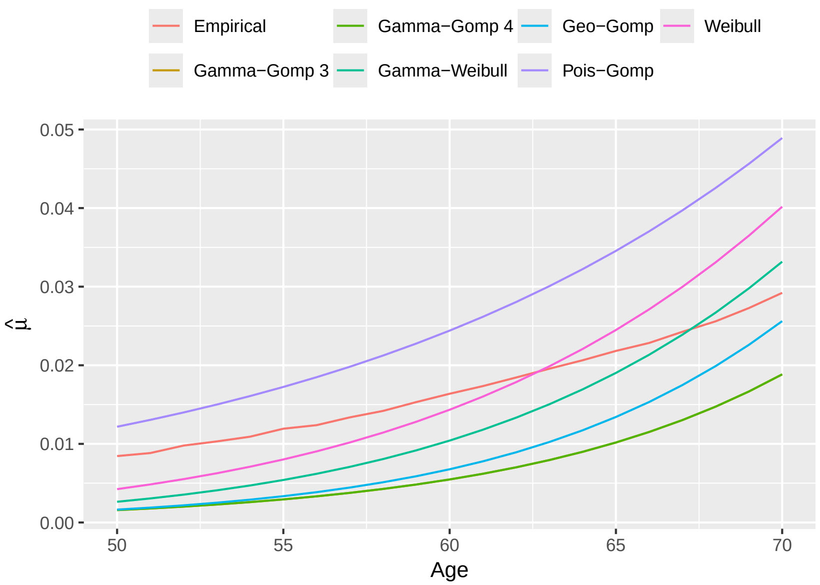

Figure 9: Mortality models with random effect applied to Lithuanian data (ages 50-70)

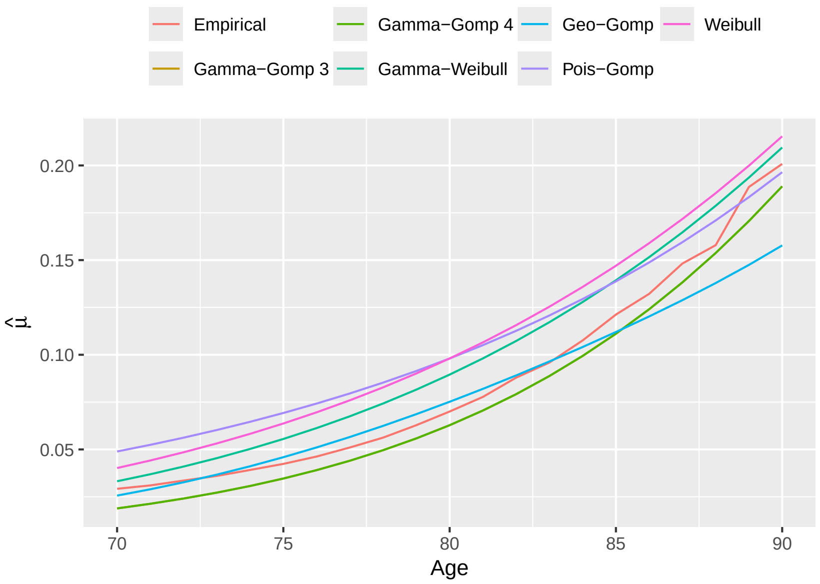

Figure 10: Mortality models with random effect applied to Lithuanian data (ages 70-90)

Table 1: MSE for different models for Lithuanian population

Model	Mean squared error (MSE)
Gamma - Gompertz with four parameters	0.0001820656
Gamma - Gompertz with three parameters	0.0001824057
Gamma - Weibull	0.0003312705
Weibull (Discrete - Weibull with $p = 1$)	0.0003361133
Poisson - Gompertz	0.0008037022
Geometric - Gompertz	0.01091893

The best parameters for each model's force of mortality function applied to Lithuanian data:

- Gamma - Gompertz with four parameters, see (59):
  $\{B = 5.136861 \times 10^{-7},\ \alpha = 0.1250529,\ \lambda = 1.320658,\ k = 7.800965\}$;

- Gamma - Gompertz with three parameters, see (60):
  $\{B = 3.028194 \times 10^{-6},\ \alpha = 0.1250734,\ \lambda = 7.798495\}$;

- Gamma - Weibull, see (57):
  $\{a = 5.497963 \times 10^{-13}, k = 9347.058,\ \lambda = 403176.2 \times 10^{12}, \alpha = 6.752323\}$;

- Weibull (Discrete - Weibull with $p = 1$), see (64):
  $\sigma = 10.38606,\ M = 79.78316,\ p = 1$;

- Poisson - Gompertz, see (62):
  $\{B = 0.0003768185,\ \alpha = 0.06951546,\ \lambda = 5.637439 \times 10^{-7}\}$;

- Geometric - Gompertz, see (63):
  $\{\alpha = 1.71132 \times 10^{-7},\ B = 0.1451152,\ p = 0.1467799\}$.

7.2 Latvian mortality results

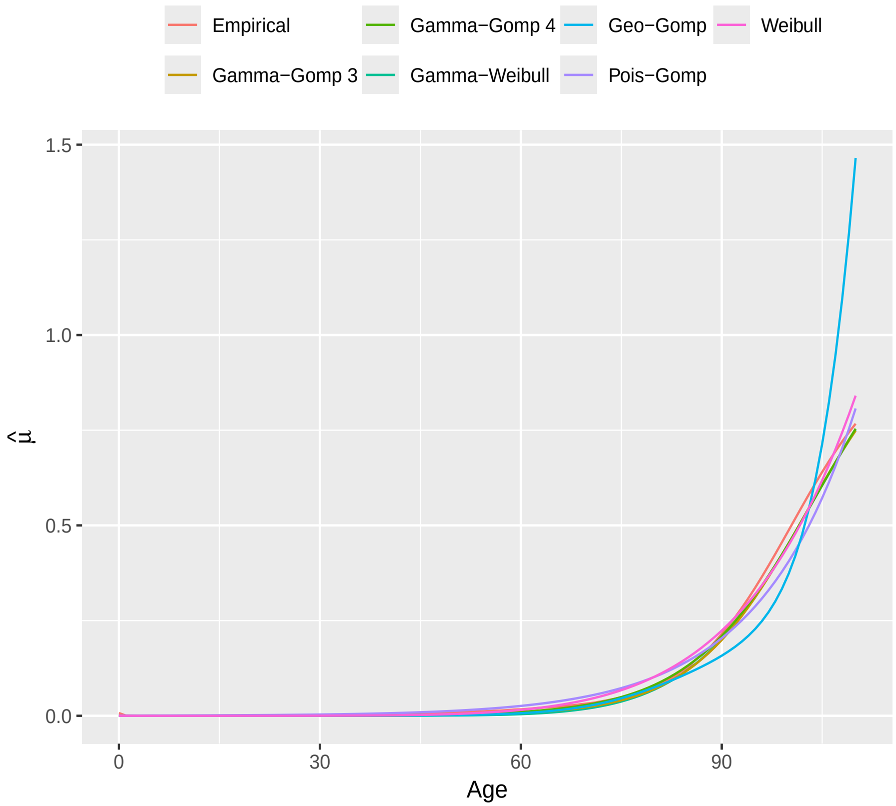

Figure 11: Mortality models with random effect applied to Latvian data (all ages)

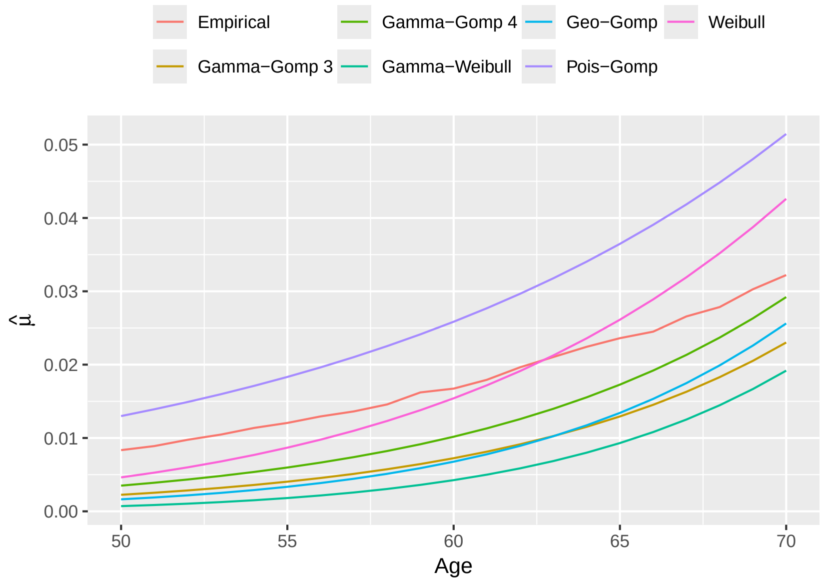

Figure 12: Mortality models with random effect applied to Latvian data (ages 50-70)

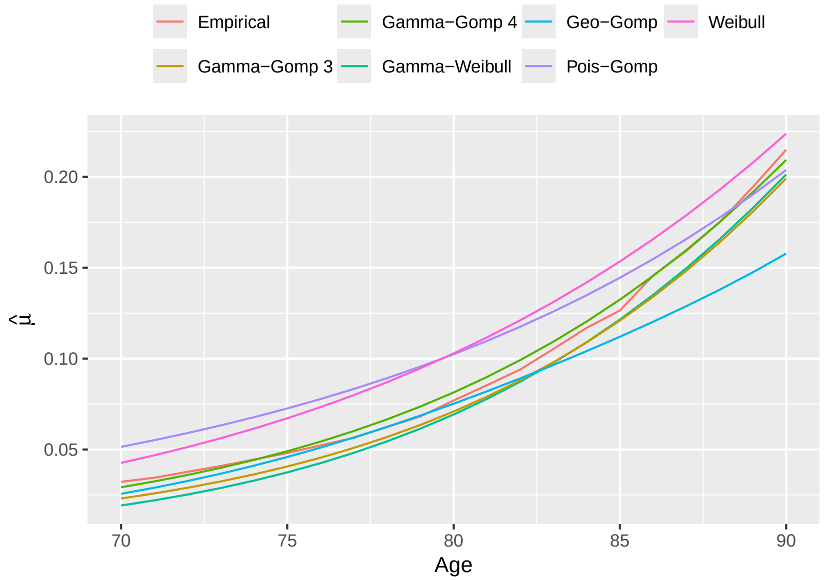

Figure 13: Mortality models with random effect applied to Latvian data (ages 70-90)

Table 2: MSE for different models for Latvian population

Model	Mean squared error (MSE)
Gamma - Gompertz with four parameters	0.0001455127
Gamma - Gompertz with three parameters	0.0001830891
Gamma - Weibull	0.0002027425
Weibull (Discrete - Weibull with $p = 1$)	0.0002679813
Poisson - Gompertz	0.0007349811
Geometric - Gompertz	0.01031875

The best parameters for each model's force of mortality function applied to Latvian data:

- Gamma - Gompertz with four parameters, see (59):
$\{B = 0.003359441,\ \alpha = 0.1071846,\ \lambda = 2184.861,\ k = 10.74998\}$;

- Gamma - Gompertz with three parameters, see (60):
$\{B = 6.45048 \times 10^{-6},\ \alpha = 0.1171701,\ \lambda = 9.04615\}$;

- Gamma - Weibull, see (57):
$\{a = 1.359224 \times 10^{-14},\ k = 14.06893,\ \lambda = 13115012,\ \alpha = 9.822215\}$;

- Weibull (Discrete - Weibull with $p = 1$), see (64):
$\sigma = 10.4192,\ M = 79.16829,\ p = 1$;

- Poisson - Gompertz, see (62):
$\{B = 0.0004160514,\ \alpha = 0.06882173,\ \lambda = 1.653197 \times 10^{-5}\}$;

- Geometric - Gompertz, see (63):
$\{\alpha = 1.71132 \times 10^{-7},\ B = 0.1451152,\ p = 0.1467799\}$.

7.3 Estonian mortality results

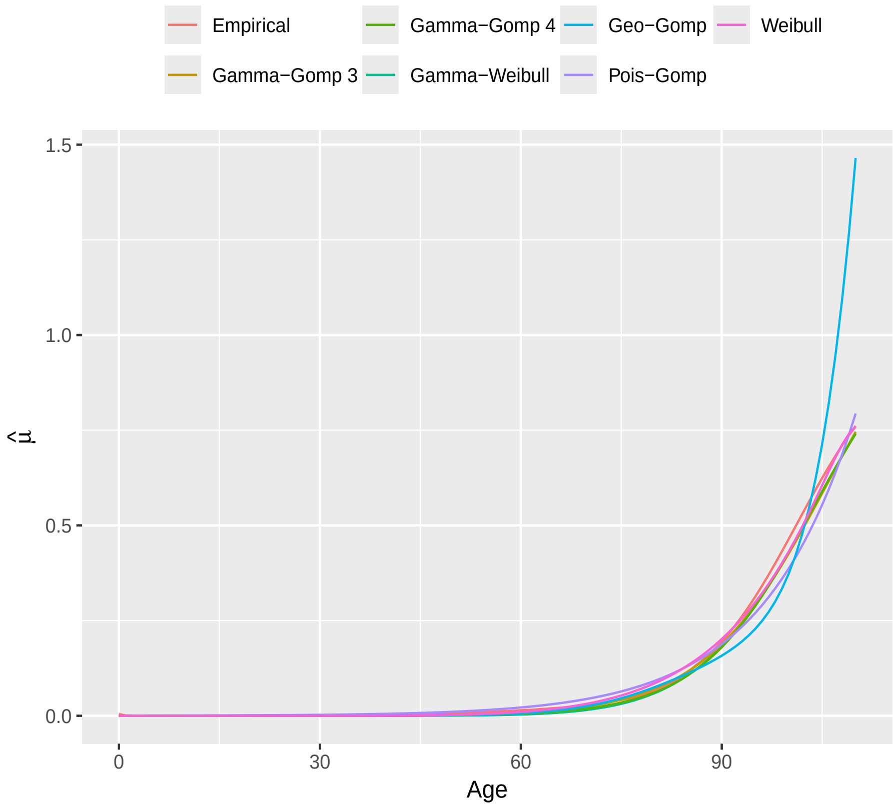

Figure 14: Mortality models with random effect applied to Estonian data (all ages)

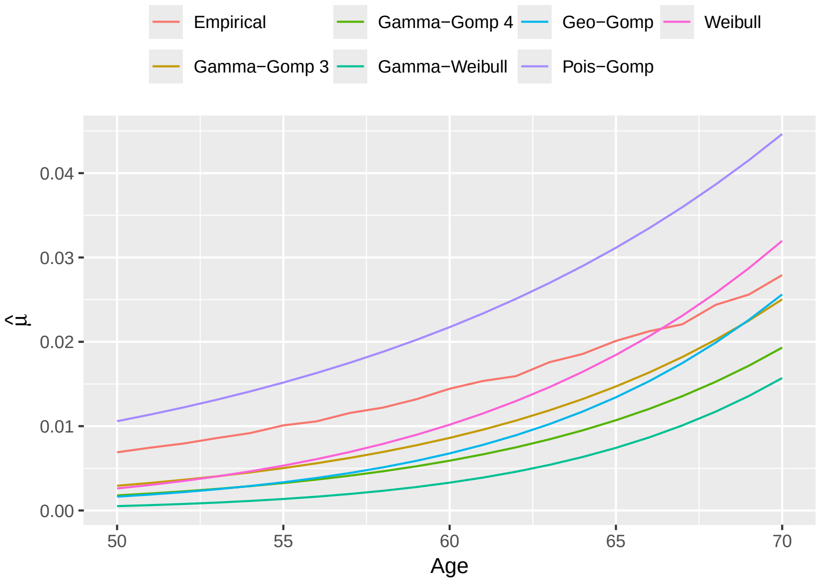

Figure 15: Mortality models with random effect applied to Estonian data (ages 50-70)

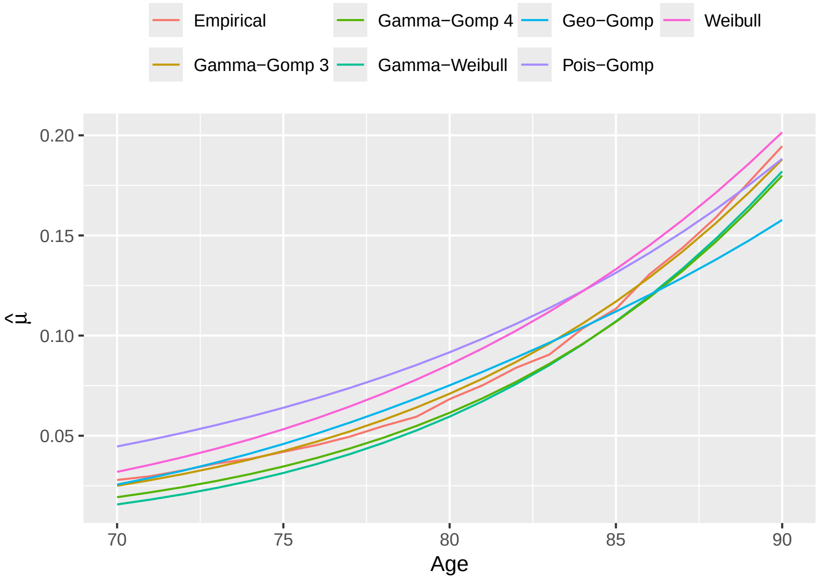

Figure 16: Mortality models with random effect applied to Estonian data (ages 70-90)

Table 3: MSE for different models for Estonian population

Model	Mean squared error (MSE)
Weibull (Discrete - Weibull with $p < 1$)	0.0001198733
Gamma - Gompertz with three parameters	0.0001817575
Gamma - Gompertz with four parameters	0.0001832729
Gamma - Weibull	0.0001985111
Poisson - Gompertz	0.0006680722
Geometric - Gompertz	0.01013277

The best parameters for each model's force of mortality function applied to Estonian data are following:

- Gamma - Gompertz with four parameters, see (59):
  $\{B = 0.03077626,\ \alpha = 0.1198032,\ \lambda = 61698.92,\ k = 8.989365\}$;

- Gamma - Gompertz with three parameters, see (60):
  $\{B = 1.317173 \times 10^{-5},\ \alpha = 0.1081533,\ \lambda = 11.23144\}$;

- Gamma - Weibull, see (57):
  $\{a = 0.9190396,\ k = 0.004814763,\ \lambda = 1.311606 \times 10^{-5},\ \alpha = 15.05727\}$;

- Weibull (Discrete - Weibull with $p < 1$), see (64):
  $\sigma = 10.57929,\ M = 89.5228,\ p = 0.001071196,\ q = 0.8634049$;

- Poisson - Gompertz, see (62):
  $\{B = 0.0002898264,\ \alpha = 0.07195851,\ \lambda = 9.621506 \times 10^{-7}\}$;

- Geometric - Gompertz, see (63):
  $\{\alpha = 1.71132 \times 10^{-7},\ B = 0.1451152,\ p = 0.1467799\}$.

7.4 Conclusion

Based on the results obtained, a conclusion can be made that for Lithuanian and Latvian population the Gamma - Gompertz (with 3 or 4 parameters) models fit the empirical mortality data the best, whereas the Discrete - Weibull model (with $p < 1$) is the best fit for mortality of Estonian population. For Lithuanian data, Gamma - Gompertz models with 3 and 4 parameters have very similar mean squared errors and forecasted hazard rates, therefore, their curves are almost identical and cannot be distinguished in the graphs. As a result of sufficiently small mean square error, all mortality force functions provided above are suitable to approximate mortality of populations under consideration, except the Geometric - Gompertz model, since force of mortality of this model is similar to the step function which is not usual to describe mortality of real population. For the Gamma - Gompertz model, the forecasted mortality is lower compared to statistics. Other analysis performed by Missov [48] and by Wang and Brown [49] also suggest that the Gamma-Gompertz model increases human life duration compared to actual statistics. Such tendency is observed also in other Gompertz frailty models (see, for instance, the Poisson - Gompertz model results in [41, 50], however, for Baltic states such trend for Poisson - Gompertz was not observed). For the Discrete - Weibull model, a conclusion is made that introduction of additional parameters into force of mortality does not reduce the error of approximation of real data.

8 Morbidity analysis

For morbidity part, the application of models with random effect is relevant for both medical and life insurance fields. Researchers in the medicine can apply these models to determine the projection of a lifespan of people who suffer from specific diseases, while life insurance market can use this information to construct mortality or morbidity tables which might be used for dread disease insurance, where the insured people pay premiums and receive payouts in case of a diagnosed disease or death caused by a specific disease. The benefits of an insured person may come in the form of a single payout after the diagnosis or scheduled benefits, which cover hospital stays and treatments. The payouts and the premiums heavily depend on survival probabilities computed from actual statistics.
One of the most common dread diseases for males is prostate cancer. According to Rebello et. al. (see [91]), it is a second leading cause of cancer diagnosis and deaths, resulting in more than million new cases and 350 000 deaths per year. Prostate cancer has been a significant topic of research in recent years as investments for this purpose during 2016-2020 exceeded 1.2 billion US dollars (see McIntosh et. al. [92]). Part of the research includes potential diagnosis without any signs or symptoms. This is done through prostate cancer screening process, which involves blood tests to check the levels of prostate specific antigen (PSA). According to Rebello et. al. (see [91]), some studies have shown that prostate cancer screening reduces the prostate cancer mortality rates. Such trend is observed based on Lithuanian data as well. In this section, analysis is performed using the prostate cancer screening information data obtained during 10-year span from January 1, 2006, until December 31, 2016 for Lithuanian men aged $50-75$. Data was observed and shared with thesis author by National Institue of Cancer. Data was split into two separate samples:

(i) men, who participated in the prostate cancer screening program and who had PSA (prostate specific antigen) tests performed,

(ii) men, who did not participate in prostate cancer screening program, but are still observed during the specified time interval.

The sample of raw data used is provided in Table 6 in the Appendix (full data not provided since it consists of more than 650 000 records).

Analysis is split into two parts: analysis for prostate cancer deaths only and analysis for all deaths (including prostate cancer). Based on the data set, containing birth, death, departure, first PSA tests, start and end dates of observation, we were able to calculate empirical values of 1-year death probabilities $q_x = {}_1q_x$ (see formula (15)), where x denotes the time moment of observation, which was constructed as follows:

(i) for screened population, the first moment $x = 0$, was set at the year of first PSA test of each individual,

(ii) for unscreened population, the first moment $x = 0$, was set at the year of start date when individual was started to be observed (i.e. the year when person turned 50 years old or year 2006 (if person turned 50 before that time)). Empirical death probabilities were used as an input for the Weibull mortality model with $p = 1$, described in Subsection 6.5. For calculation of the model parameters M and σ, *R* statistical program and the *MortalityLaws* library were used, where empirical mortality probabilities were specified with the following binomial loss function

$$LF = -\sum_{x=0}^{10} \Big(D_x \ln\left(1 - \mathrm{e}^{-\mu_x}\right) - (E_x - D_x)\mu_x\Big),$$

where D_x is the estimated death counts at time moment x, E_x is the estimated population exposed to risk at time moment x, μ_x is the estimated model hazard rate at moment x. All estimates are based on empirical death probabilities. For the prostate cancer mortality, we get the following Weibull model parameters:

(i) for screened population: $M = 20.93804$, $\sigma = 4.150694$,

(ii) for unscreened population: $M = 28.58124$, $\sigma = 8.154898$.

For the first 10 years $x \in [0, 10]$, curves of survival functions $S(x)$ are shown in Figure 17.

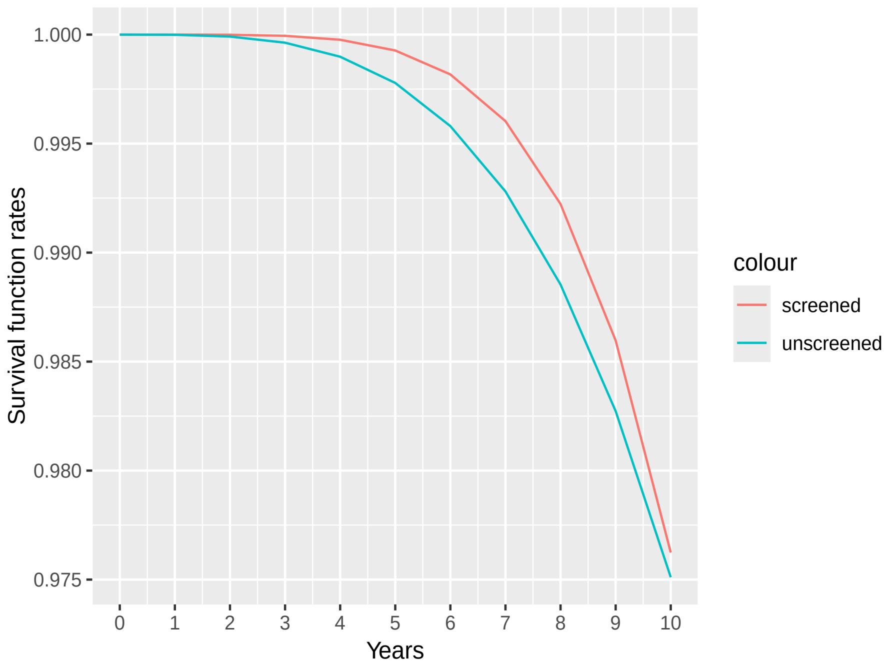

Figure 17: Survival functions for prostate cancer mortality

Using the relations from Section 3.2, we get the following particular characteristics:

(i) For screened population: probability to live at least 6 years $S(6) = 0.998174$, probability to live at least 8 years $S(8) = 0.992229$, the remaining 10-year life expectancy $\mathring{e}_{0,\overline{10}|} = 9.96048$.

(ii) For unscreened population: probability to live at least 6 years $S(6) = 0.995802$, probability to live at least 8 years $S(8) = 0.988535$, the remaining 10-year life expectancy $\mathring{e}_{0,\overline{10}|} = 9.94444$.

For mortality in general (including the prostate cancer mortality), we obtain the following Weibull model parameters:

(i) for screened population: $M = 10.34769$, $\sigma = 1.956164$,

(ii) for unscreened population: $M = 9.568636$, $\sigma = 3.62983$.

In the case of mortality in general, curves of survival functions $S(x)$ are shown in Figure 18.

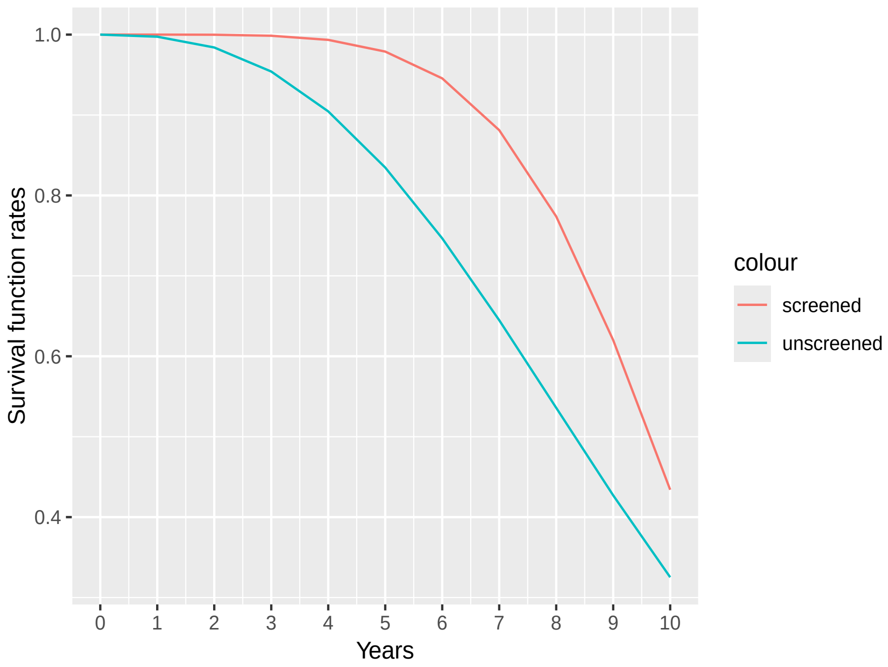

Figure 18: Survival functions for mortality in general

In the case under consideration, using the relations from Section 3.2 again, we get the following particular characteristics:

(i) For screened population: probability to live at least 6 years $S(6) = 0.945141$, probability to live at least 8 years $S(8) = 0.773109$, the remaining 10-year life expectancy $\mathring{e}_{0,\overline{10}|} = 8.92152$.

(ii) For unscreened population: probability to live at least 6 years $S(6) = 0.746628$, probability to live at least 8 years $S(8) = 0.535927$, the remaining 10-year life expectancy $\mathring{e}_{0,\overline{10}|} = 7.69986$.

Based on the results computed with the Discrete - Weibull mortality model, a conclusion can be made that the prostate screening program in Lithuania has been somewhat effective and that screened individuals have higher partial remaining life expectancies and survival rates compared to unscreened individuals.

9 Concluding remarks

In this thesis, author analyzes the tail of distribution function of random variable X, defined on a probability space $(\mathbf{W}, \mathcal{F}, \mathbf{P})$:

$$\overline{F}_X(x) := \mathbf{P}(X > x) = 1 - F_X(x),\ x \in \mathbb{R},$$

where

$$F_X(x) = \mathbf{P}(X \leqslant x) = \mathbf{P}(w \in \mathbf{W} : X(w) \leqslant x),\ x \in \mathbb{R}$$

is a distribution function of X. The characteristics of F_X and $\overline{F}_X$ are described and analyzed by providing the main properties of the functions as well as providing some examples. Additionally, the behaviour of tail functions is explained by splitting them into various categories, also known as the regularity classes.

The main idea of the thesis is to check how and if the behaviour and properties of the tail functions change if they are transformed. Transformation is obtained by applying a positive random variable Z to a standard tail function of F_X:

$$\overline{F^{(Z)}}(x) := \mathbb{E}\Big(\overline{F}^Z(x)\Big) = \mathbb{E}\Big(\big(\overline{F}(x)\big)^Z\Big),\ x \in \mathbb{R}.$$

The d.f. for transformed t.f. is

$$F^{(Z)}(x) = 1 - \overline{F^{(Z)}}(x) = 1 - \mathbb{E}\Big(\overline{F}^Z(x)\Big),\ x \in \mathbb{R}.$$

We remark only that in general the random variable Z is defined on a possible another probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Therefore, we use the symbol $\mathbb{E}$ for expectation in this probability space.

Thesis author constructs a theorem, stating that $F^{(Z)}$ is a new distribution function, satisfying all the required properties. Moreover, transformed tail function maintains its behaviour in terms of specific regularity classes. Therefore, author constructs and proves a theorem, stating that $\overline{F^{(Z)}}$ belongs to its original regularity class before transformation:

(i) $F \in \mathcal{R}_\alpha \ \Rightarrow\ F^{(Z)} \in \mathcal{R}_{\alpha a}$, where a is the greatest lower bound of r.v. Z, i.e:

$$\mathbb{P}(Z \geqslant a) = 1 \text{ and } \mathbb{P}(Z \geqslant a+\delta) < 1 \text{ for any } \delta > 0;$$

(ii) $F \in \mathcal{L} \ \Rightarrow \ F^{(Z)} \in \mathcal{L}$;
(iii) $F \in \mathcal{OL} \ \Rightarrow \ F^{(Z)} \in \mathcal{OL}$;
(iv) $F \in \mathcal{D} \ \Rightarrow \ F^{(Z)} \in \mathcal{D}$,

where F is a d.f. of X.

Tail functions can be analyzed when describing human mortality. If T is a non-negative and absolutely continuous random variable, then the function

$$S(x) = S_T(x) = \mathbf{P}(T > x) = 1 - F_T(x) = \overline{F}_T(x),$$

is called a survival function, which is equal to the tail function of F_T. Survival function is used in other mortality characteristics, such as the hazard rate:

$$\mu_x = \tfrac{f(x)}{S(x)} = -\tfrac{S'(x)}{S(x)}, \ x \geqslant 0.$$

Thesis author defines a transformed survival function by formula

$$\widehat{S}(x) = \mathbb{E}\left(S^Z(x)\right) = \mathbb{E}\left(\exp\left\{ - Z \int_0^x \mu_u \mathrm{d}u\right\}\right).$$

In thesis, author constructs a theorem stating that $\widehat{S}(x)$ is a new survival function which satisfies all the required properties meaning that both transformed survival and hazard rate functions can be used when describing human mortality.

After the theoretical part of the thesis, practical part is provided with the application of transformed survival and hazard rate functions to two sets of actual data. The first set covers Baltic population mortality since year 2000 split into Estonian, Latvian and Lithuanian data. Thesis author applies a group of different transformed mortality models to the data and calculates model's parameters in order to evaluate which of the models describe certain population's mortality the best. The fit is validated through mean squared error metric. After the analysis, it is determined that out of analyzed models, the Gamma - Gompertz model provided the best fit for Lithuanian and Latvian mortality data with the following survival function expression:

$$\hat{S}_{GG_4}(x) = \left(1 + \frac{B}{\alpha\lambda}(\mathrm{e}^{\alpha x} - 1)\right)^{-k}.$$

For Estonia's mortality, most precise results are obtained with the Discrete - Weibull model (with parameter $p < 1$):

$$\hat{S}_{DW}(x) = p\,\mathrm{e}^{-\left(\frac{x}{M}\right)^{M/\sigma}} + q\,\mathrm{e}^{-2\left(\frac{x}{M}\right)^{M/\sigma}} + (1-p-q)\mathrm{e}^{-3\left(\frac{x}{M}\right)^{M/\sigma}}.$$

The Discrete - Weibull model with $p = 1$ is used for the second dataset, which covers the prostate screening program in Lithuania for males aged 50 - 75 between years 2006 - 2016. With the application of the model to the mortality data of screended and unscreened individuals, 10-year survival probabilities are obtained. With the help of the model it is determined that screened individuals have higher survival rate from both cancer mortality and mortality in general compared to unscreened individuals.

Both practical and theoretical parts show that with the transformation it is possible to obtain a variety of new survival and hazard rate functions. Transformed functions can be thoroughly analyzed in terms of all available regularity classes and not only the ones covered in this thesis. Moreover, it is shown that the new functions can be applied to practical problems in medicine, finance and insurance. It is important to note that models, described in the thesis, cannot be considered as unambiguously best applicable for mortality forecasting, since the choice of model is highly dependent on the population that we are studying. This work only includes a small amount of mortality models. Therefore, the search for the unambiguously best applicable model for mortality forecasting remains one of the unsolved tasks for mathematicians and the life insurance market.

10 Disertacijos santrauka

10.1 Temos aktualumas

Atsitiktinių dydžių (a.d.), taip pat žinomų kaip atsitikinių efektų, savybių analizė yra viena svarbiausių temų tikimybių teorijoje ir statistikoje. A.d. įvairovė suteikia galimybę konstruoti įvairias pasiskirstymo funkcijas su ypatingomis savybėmis, kurios plačiai naudojamos ne tik teorinėje matematikoje, bet ir finansų sektoriuje: bankuose, gyvybės draudime, rizikos vertinime. Pasiskirstymo funkcijų uodegos yra labai svarbios konstruojant naujas pasiskirstymo funkcijas su norimomis savybėmis. Pagal uodegų elgesį pasiskirstymo funkcijų aibė skirstoma į reguliarumo klases. Pasiskirstymo funkcijos uodega yra pagrindinė išgyvenamumo charakteristika išgyvenamumo analizėje. Analizuojant populiacijos išgyvenamumą, įprastai pasiskirstymo funkcijos uodega vadinama išgyvenamumo funkcija. Be to, išgyvenamumo analizėje įprastai laikoma, kad ta pagrindinė charakteristika - išgyvenamumo funkcija - yra absoliučiai tolydi. Tokia prielaida yra susijusi su individo teise numirti bet kuriuo laiko momentu. Išgyvenamumo analizėje, žinant išgyvenamumo funkcijos išraišką, galima suskaičiuoti įvairias mirtingumo charakteristikas: mirties tikimybę, mirtingumo galią, likusią gyvenimo trukmę ir pan. Turint konkrečios populiacijos mirtingumo funkcijos išraišką, galima sukonstruoti mirtingumo lenteles, kurios tiesiogiai naudojamos įvertinant gyvybės draudimo įmokas, techninius atidėjinius ar kitas finansines charakteristikas. Be to, tos pačios išgyvenamumo funkcijos gali būti naudojamos skirtingoms populiacijoms, jeigu pritaikome naujas parametrų reikšmes. Taigi, išgyvenamumo funkcijos ir pasiskirstymo funkcijos suteikia galimybę sukurti įvairius vadinamuosius mirtingumo modelius, kurie bus viena pagrindinių šios disertacijos temų.

10.2 Disertacijos tikslas

Disertacija turi dvi dalis, teorinę ir praktinę. Teorinės dalies tikslas yra atsitiktinio efekto naudojimo pateisinimas išgyvenamumo analizėje. Teorinėje dalyje įrodoma, kad išgyvenamumo funkciją (pasiskirstymo funkcijos uodegą) paveikus atsitiktiniu efektu gaunama nauja išgyvenamumo funkcija (pasiskirstymo funkcijos uodega). Taip pat parodoma, kad atsitiktiniu efektu paveikta pasiskirstymo funkcija daugeliu atvejų pasilieka toje pačioje reguliarumo klasėje, kuriai pasiskirstymo funkcija priklausė prieš transformaciją.

Praktinėje dalyje naudojame realius Baltijos šalių mirtingumo duomenis. Praktinės disertacijos dalies tikslas yra nustatyti, kuris mirtingumo modelis su atsitiktiniu efektu labiausiai tinka pasirinktos populiacijos išgyvenamumui aprašyti. Papildomas praktinės dalies tikslas yra nustatymas, kiek veiksminga yra Lietuvos prostatos vėžio stebėsenos programa. Pastarajam tikslui pasiekti yra naudojami NVI duomenys ir specialios struktūros mirtingumo modelis.

10.3 Pagrindinės sąvokos

10.3.1 Pasiskirstymo, uodegos ir išgyvenamumo funkcijos

Tarkime, kad X yra atsitiktinis dydis, apibrėžtas tikimybinėje erdvėje $(\mathbf{W}, \mathcal{F}, \mathbf{P})$.

Apibrėžimas 1. *Funkcija, nusakanti a.d. $X : \mathbf{W} \to \mathbb{R}$ elgesį, yra vadinama pasiskirstymo funkcija (p.f.). Ši funkcija nusakoma tokia lygybe:*

$$F_X(x) := \mathbf{P}(X \leqslant x) = \mathbf{P}(w \in \mathbf{W} : X(w) \leqslant x),\ x \in \mathbb{R}. \tag{65}$$

Dėl tikimybinės erdvės savybių bet kuri p.f. F tenkina tokias savybes:

(i) $0 \leqslant F(x) \leqslant 1, \quad x \in \mathbb{R},$ (66)

(ii) $\lim\limits_{x \to -\infty} F(x) = 0, \quad \lim\limits_{x \to +\infty} F(x) = 1,$ (67)

(iii) *F yra nemažėjanti aibėje $\mathbb{R}$,* (68)

(iv) *F yra tolydi iš dešinės ir turi baigtinę ribą iš kairės kiekviename $a \in \mathbb{R}$, t.y. $\lim\limits_{x \to a+} F(x) = F(a),$ ir egzistuoja $\lim\limits_{x \to a-} F(x)$ su $a \in \mathbb{R}$.* (69)

Diskretaus a.d. atveju, pasiskirstymo funkcija turi išraišką

$$F_X(x) = \sum_{x_i \leqslant x} \mathbf{P}(X = x_i), \tag{70}$$

kur $x_i \in \mathcal{N}$ ir $\mathcal{N} = \{x_i\}$ yra baigtinė arba skaiti realių skaičių aibė.

Jeigu X yra absoliučiai tolydus, tada

$$F_X(x) = \int\limits_{-\infty}^{x} f_X(t)\mathrm{d}t,\ x \in \mathbb{R}, \tag{71}$$

kur f_X yra tankio funkcija, t.y. mati, neneigiama funkcija, tokia, kad

$$\int\limits_{-\infty}^{\infty} f_X(t)\mathrm{d}t = 1.$$

.

Formulė (71) reiškia, kad $f_X(t) = F_X^{'}(t)$ beveik visur aibėje $\mathbb{R}$.

Apibrėžimas 2. *Tarkime, kad F yra kokio nors atsitiktinio dydžio pasiskirstymo funkcija. Tada*

$$\overline{F}(x) = 1 - F_X(x) = \mathbf{P}(X > x),\ x \in \mathbb{R}, \tag{72}$$

vadinama pasiskirstymo funkcijos F uodega.

10.3.2 Išgyvenamumo analizės sąvokos

Tarkime, kad $T \geqslant 0$ yra neneigiamas absoliučiai tolydus atsitiktinis dydis, nusakantis naujagimio tikėtiną gyvenimo trukmę. Pagrindinė funkcija, nusakanti populiacijos mirtingumą yra išgyvenamumo funkcija:

$$S(x) = S_T(x) = \mathbf{P}(T > x).$$

Kadangi $S(x) = S_T(x) = 1 - F_T(x) = 1 - \mathbf{P}(T \leqslant x)$ ir $T \geqslant 0$, tai išgyvenamumo funkcijos tenkina sąlygas:

$$\text{(i)} \quad S(x) \text{ yra nedidėjanti su } x \in [0, \infty); \tag{73}$$

$$\text{(ii)} \quad S(0) = 1; \tag{74}$$

$$\text{(iii)} \quad S(\infty) := \lim_{x \to \infty} S(x) = 0; \tag{75}$$

$$\text{(iv)} \quad S(x) \text{ yra absoliučiai tolydi intervale } [0, \infty). \tag{76}$$

Pagal absoliutaus tolydumo apibrėžimą, pasirinktam $\varepsilon > 0$ egzistuoja $\delta > 0$, toks, kad kiekvienam baigtiniam nesikertančių intervalų porų $\{(a_1, b_1), \ldots, (a_n, b_n)\}$, kur

$$\sum_{k=1}^{n} (b_k - a_k) < \delta,$$

rinkiniui tenkinama sąlyga

$$\sum_{k=1}^{n} (S(a_k) - S(b_k)) < \varepsilon.$$

Funkcija $S(x)$ yra absoliučiai tolydi tada ir tik tada, kai:

$$S(x) = \int_x^{\infty} f_T(u) \mathrm{d}u,\ x \geqslant 0, \tag{77}$$

kur f_T yra tankio funkcija, t.y. tokia neneigiama ir integruojama funkcija, kad

$$\int_0^\infty f_T(u)\mathrm{d}u = 1.$$

Žinant išgyvenamumo funkciją, galima paskaičiuoti papildomas charakteristikas, nusakančias mirtingumą. Pavyzdžiui, tikimybė, jog x metų amžiaus individas gyvens bent n metų, yra lygi

$$_np_x = \frac{S(x+n)}{S(x)}, \tag{78}$$

tikimybė, jog x metų amžiaus individas negyvens n metų, yra lygi

$$_nq_x = 1 - \frac{S(x+n)}{S(x)}, \tag{79}$$

dalinė likusi gyvenimo trukmė x metų amžiaus individui per ateinančius n metų

$$\mathring{\mathrm{e}}_{x,\overline{n}|} = \frac{1}{S(x)}\int_x^{x+n} S(u)\mathrm{d}u. \tag{80}$$

Dar viena charakteristika, naudojama mirtingumo analizėje, yra vadinamoji mirtingumo galia, žymima μ_x. Mirtingumo galia apskaičiuojama pagal formulę:

$$\mu_x = \frac{f(x)}{S(x)} = -\frac{S'(x)}{S(x)}, \tag{81}$$

kur $x \in [0, \infty)$, toks, kad $S(x) > 0$.

Remiantis aukščiau esančia formule, išgyvenamumo funkcijos reikšmė taške x užrašoma lygybe

$$S(x) = \exp\Big\{ -\int\limits_0^x \mu_u \mathrm{d}u \Big\}. \tag{82}$$

Mirtingumo galios funkcijos elgesys intervale $[x, x+1)$ yra susiję̨s su vadinamuoju centriniu mirtingumo dažniu:

$$m_x = \frac{\int\limits_0^1 S(x+u)\mu_{x+u}\mathrm{d}u}{\int\limits_0^1 S(x+u)\mathrm{d}u} = \frac{S(x) - S(x+1)}{\int\limits_0^1 S(x+u)\mathrm{d}u}. \tag{83}$$

Jeigu mirtingumo galia intervale $[x, x+1)$ yra pastovi, t.y.

$$\mu_{x+t} = \mu_x \text{ bet kuriam } x \in \mathbb{N}_0 = \{0, 1, ...\} \text{ ir } t \in [0, 1),$$

tai centrinis mirtingumo dažnis yra lygus mirtingumo galiai:

$$m_x = \mu_x, \ x \in \mathbb{N}_0.$$

10.3.3 Reguliarumo klasės

Pasiskirstymo funkcijos yra suskirstytos į reguliarumo klases atsižvelgiant į tų funkcijų uodegų elgesį begalybės aplinkoje. Žemiau pateikiame pagrindinių reguliarumo klasių apibrėžimus. Pavyzdžius klasių atstovų galima rasti pagrindiniame disertacijos tekste.

Apibrėžimas 3. *P.f. F vadinama sunkiauodege, žymime $F \in \mathcal{H}$, jei bet kuriam fiksuotam $\delta > 0$*

$$\limsup_{x \to \infty} \mathrm{e}^{\delta x}\, \overline{F}(x) = \infty. \tag{84}$$

Apibrėžimas 4. *Sakoma, kad p.f. F priklauso eksponentinių uodegų klasei $\mathcal{L}(\gamma)$ su parametru $\gamma > 0$, jei bet kuriam $y > 0$*

$$\lim_{x \to \infty} \frac{\overline{F}(x+y)}{\overline{F}(x)} = e^{-\gamma y}. \tag{85}$$

Apibrėžimas 5. *Sakoma, kad p.f. F turi reguliariai kintančią uodegą su parametru $\alpha \geqslant 0$, žymime $F \in \mathcal{R}_\alpha$, jei bet kuriam $y > 0$ tenkinama savybė*

$$\lim_{x \to \infty} \frac{\overline{F}(xy)}{\overline{F}(x)} = y^{-\alpha}. \tag{86}$$

Visų reguliariai kintančių p.f. aibė žymima

$$\mathcal{R} := \bigcup_{\alpha \geqslant 0} \mathcal{R}_\alpha.$$

Apibrėžimas 6. *Sakoma, kad p.f. F turi dominuojančiai kintančią uodegą ($F \in \mathcal{D}$), jeigu*

$$\limsup_{x \to \infty} \frac{\overline{F}(xy)}{\overline{F}(x)} < \infty. \tag{87}$$

kiekvienam (ekvivalenčiai, su kokiu nors) $y \in (0, 1)$.

Apibrėžimas 7. *Sakome, kad pasiskirstymo funkcija F turi ilgą uodegą, rašome $F \in \mathcal{L}$, jeigu*

$$\overline{F}(x-y) \underset{x\to\infty}{\sim} \overline{F}(x). \tag{88}$$

Apibrėžimas 8. *Sakome, kad p.f. F priklauso $\mathcal{O}$-eksponentinių skirstinių klasei $\mathcal{OL}$, jei su kiekvienu (arba, analogiškai, su kokiu nors) $y > 0$*

$$\limsup_{x\to\infty} \frac{\overline{F}(x-y)}{\overline{F}(x)} < \infty. \tag{89}$$

Apibrėžimas 9. *Neneigiamo atsitiktinio dydžio X p.f. $F = F_X$ yra subeksponentinė ($F \in \mathcal{S}$), jeigu*

$$\lim_{x\to\infty} \frac{\overline{F*F}(x)}{\overline{F}(x)} = 2. \tag{90}$$

Apibrėžimas 10. *Sakome, kad p.f. F priklauso nuosaikiai kintančių uodegų klasei $\mathcal{C}$, jeigu*

$$\lim_{y\uparrow 1} \limsup_{x\to\infty} \frac{\overline{F}(xy)}{\overline{F}(x)} = 1. \tag{91}$$

Aukščiau išvardintos reguliarumo klasės tenkina šiuos sąryšius:

$$\mathcal{R} \subset \mathcal{C} \subset \mathcal{L} \cap \mathcal{D} \subset \mathcal{S} \subset \mathcal{L} \subset \mathcal{H}, \qquad \mathcal{D} \subset \mathcal{H}, \qquad \mathcal{D} \subset \mathcal{OL}.$$

Iš apibrėžimų taip pat nesunku pastebėti, kad

$$\mathcal{L} \subset \mathcal{OL}, \qquad \bigcup_{\gamma>0} \mathcal{L}(\gamma) \subset \mathcal{OL}.$$

Svarbu pažymėti, kad pasiskirstymo funkcijos uodegos neprivalo būti absoliučiai tolydžios, tačiau išgyvenamumo analizėje išgyvenamumo funkcija (arba uodegos funkcija) šią savybę tenkinti privalo.

Išgyvenamumo funkcijas (p.f. uodegas) (analogiškai ir pačias pasiskirstymo funkcijas) galima analizuoti jas transformavus. Vadinamoji atsitiktinio efekto transformacija, taikoma išgyvenamumo analizėje, gaunama mirtingumo galią dauginant iš teigiamo atsitiktinio dydžio Z, kuris išgyvenamumo analizėje vadinamas atsitiktiniu efektu:

$$\mu_{x|Z} = Z\mu_x. \tag{92}$$

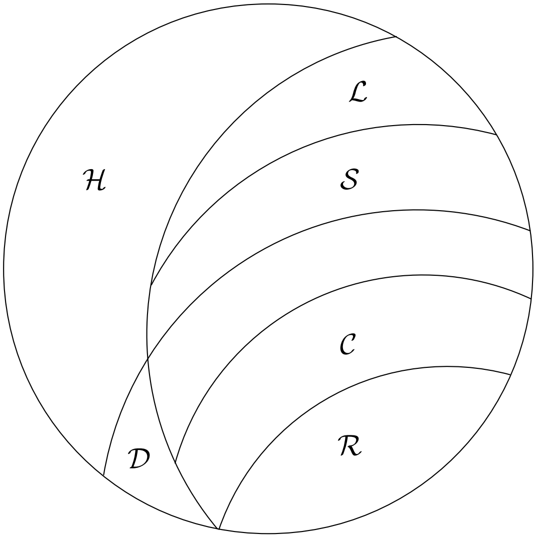

Figure 19: Sunkiauodegių skirstinių reguliarumo klasės

Išgyvenamumo analizėje modeliai, turintys tokį pavidalą, vadinami modeliais su atsitiktiniu efektu. Šiuose modeliuose atsitiktinis dydis Z yra apibrėžtas tikimybinėje erdvėje $(\Omega, \mathcal{A}, \mathbb{P})$, o atsitiktinis dydis T, nusakantis likusią naujagimio gyvenimo trukmę ir naudojamas mirtingumo galios funkcijos μ_x skaičiavimui, yra apibrėžtas tikimybinėje erdvėje $(\mathbf{W}, \mathcal{F}, \mathbf{P})$.

Tiesa, norint naudoti modelius su atsitiktiniu efektu, privalu įsitiktinti, kad transformuotos funkcijos tenkins visas originalias išgyvenamumo funkcijos sąlygas. Jeigu transformuotos mirtingumo galios funkcija turi (92) pavidalą, tada funkcija

$$\widehat{S}(x) = \mathbb{E}\left(S^Z(x)\right) = \mathbb{E}\left(\exp\Big\{-Z\int_0^x \mu_u \mathrm{d}u\Big\}\right) \tag{93}$$

turėtų būti nauja išgyvenamumo funkcija, o

$$\widehat{\mu}(x) = -\frac{\widehat{S}'(x)}{\widehat{S}(x)} = -\ln\left(\mathbb{E}\left(S^Z(x)\right)\right)' \tag{94}$$

turėtų būti nauja mirtingumo galios funkcija. Šie teiginiai bus teisingi tik tokiu atveju, jeigu bus tenkinamos (73) - (76) sąlygos. Kitame skyrelyje pateiksime disertacijoje suformuluotas teoremas, įskaitant ir teoremą, kuri teigia, jog funkcija $\hat{S}(x)$, nusakyta formule (93), yra nauja išgyvenamumo funkcija.

10.4 Pagrindiniai darbo teiginiai

Teorema 1. *Tarkime, kad $S = S(x)$ yra išgyvenamumo funkcija su mirtingumo galia $\mu = \mu_x$, o Z yra teigiamas atsitiktinis dydis. Tada funkcija $\widehat{S} = \widehat{S}(x)$, apibrėžta formulėje* (93)*, yra nauja išgyvenamumo funkcija.*

Iš (93) formulės matyti, jog atsitiktinio efekto transformacija gali būti pritaikyta ne tik išgyvenamumo funkcijai, kuri yra absoliučiai tolydi, bet ir bet kokiai uodegos funkcijai, nepriklausomai nuo to, ar ta funkcija yra absoliučiai tolydi. Remiantis (93) formule, bet kuriai uodegos funkcijai $\overline{F}$ galima apibrėžti

$$\overline{F^{(Z)}}(x) := \mathbb{E}\Big(\overline{F}^Z(x)\Big) = \mathbb{E}\Big(\big(\overline{F}(x)\big)^Z\Big),\ x \in \mathbb{R}, \tag{95}$$

o pasiskirtymo funkcija šiai uodegai lygi

$$F^{(Z)}(x) = 1 - \overline{F^{(Z)}}(x) = 1 - \mathbb{E}\Big(\overline{F}^Z(x)\Big),\ x \in \mathbb{R}. \tag{96}$$

Teorema 2. *Tarkime, kad $F = F(x)$ yra pasiskirstymo funkcija, o Z yra teigiamas atsitiktinis dydis. Tada funkcija $F^{(Z)}$, apibrėžta formulėje* (96)*, yra nauja pasiskirstymo funkcija.*

Negana to, kai kuriais atvejais atsitiktinis efektas nepakeičia pasiskirstymo funkcijų savybių kalbant apie specifines reguliarumo klases.

Teorema 3. *Tarkime, kad F yra pasiskirstymo funkcija, Z - teigiamas atsitiktinis dydis, ir $F^{(Z)}$ - p.f. su atsitiktiniu efektu Z, apibrėžta* (96) *formule. Tokiu atveju teisingi šie teiginiai:*

(i) $F \in \mathcal{R}_\alpha \ \Rightarrow\ F^{(Z)} \in \mathcal{R}_{\alpha a}$, *kur a yra didžiausiais a.d. Z apatinis rėžis, t.y.:*

$$\mathbb{P}(Z \geqslant a) = 1 \text{ ir } \mathbb{P}(Z \geqslant a + \delta) < 1 \text{ bet kuriam } \delta > 0;$$

(ii) $F \in \mathcal{L} \ \Rightarrow\ F^{(Z)} \in \mathcal{L}$;

(iii) $F \in \mathcal{OL} \ \Rightarrow\ F^{(Z)} \in \mathcal{OL}$;

(iv) $F \in \mathcal{D} \ \Rightarrow\ F^{(Z)} \in \mathcal{D}$.

Detalius teoremų įrodymus galima rasti skyrelyje 5.1-5.3.

10.5 Specialios išgyvenamumo funkcijos

10.5.1 Gamma - Weibull modelis

Weibull mirtingumo galios funkcija yra

$$\mu_x = ax^{\alpha}, \quad x \geqslant 0,$$

su teigiamais parametrais a ir α.

Atsitiktinis dydis Z turi Gamma skirstinį, t.y. $Z \sim \Gamma(k, \lambda)$, su tankio funkcija

$$f_Z(x) = \frac{\lambda^k}{\Gamma(k)} x^{k-1} \mathrm{e}^{-\lambda x}, \quad x \geqslant 0,$$

kur k ir λ yra teigiami parametrai, o $\Gamma(k) = \int\limits_0^{\infty} t^{k-1} e^{-t} dt$ yra standartinė Gamma funkcija.

Atlikus skaičiavimus, remiantis formule (93), Gamma - Weibull išgyvenamumo funkcija turės pavidalą

$$\begin{aligned} \hat{S}_{GW}(x) &= \frac{\lambda^k}{\Gamma(k)} \frac{1}{\left(\int\limits_0^x at^{\alpha}\mathrm{d}t + \lambda\right)^k} \int\limits_0^{\infty} \mathrm{e}^{-w} w^{k-1} \mathrm{d}w \\ &= \frac{\lambda^k}{\left(\frac{ax^{\alpha+1}}{\alpha+1} + \lambda\right)^k} = \left(\frac{ax^{\alpha+1}}{(\alpha+1)\lambda} + 1\right)^{-k}. \end{aligned}$$

Naudojantis (94) formule, randama modelio mirtingumo galios funkcija:

$$\hat{\mu}_x = -\frac{\widehat{S}'_{GW}(x)}{\widehat{S}_{GW}(x)} = \frac{akx^{\alpha}}{\frac{ax^{\alpha+1}}{\alpha+1} + \lambda}. \tag{97}$$

10.5.2 Gamma - Gompertz modelis

Gamma - Gompertz modelyje mirtingumo galios funkcija turi Gompertz pavidalą:

$$\mu_x = B\mathrm{e}^{\alpha x}, \quad x \geqslant 0, \tag{98}$$

su teigiamais parametrais B ir α, o Z turi Gamma skirstinį, tokį patį kaip skyrelyje 10.5.1. Modelio transformuotos išgyvenamumo ir mirtingumo galios funkcijos yra tokios:

$$\hat{S}_{GG_4}(x) = \Big(1 + \frac{B}{\alpha\lambda}(\mathrm{e}^{\alpha x} - 1)\Big)^{-k},$$

$$\hat{\mu}_x = -\frac{\hat{S}'_{GG_4}(x)}{\hat{S}_{GG_4}(x)} = \frac{k\frac{B}{\lambda}\mathrm{e}^{\alpha x}}{\Big(1 + \frac{B}{\alpha\lambda}(\mathrm{e}^{\alpha x} - 1)\Big)}. \tag{99}$$

Atveju $k = \lambda$, turėsime ne keturių, o trijų parametrų modelį:

$$\hat{S}_{GG_3}(x) = \Big(1 + \frac{B}{\alpha\lambda}(\mathrm{e}^{\alpha x} - 1)\Big)^{-\lambda}, \ \hat{\mu}_x = \frac{B\mathrm{e}^{\alpha x}}{\Big(1 + \frac{B}{\alpha\lambda}(\mathrm{e}^{\alpha x} - 1)\Big)}. \tag{100}$$

10.5.3 Poisson - Gompertz modelis

Poisson – Gompertz modelyje mirtingumo galia turi Gompertz išraišką (98), o atsitiktinis dydis Z turi paslinktą Poisson skirstinį su parametru $\lambda > 0$, t.y.

$$\mathbb{P}(Z = k) = \mathrm{e}^{-\lambda}\frac{\lambda^{k-1}}{(k-1)!}, \ k \in \mathbb{N} = \{1, 2, \ldots\}.$$

Diskretaus a.d. Z atveju, išgyvenamumo funkcija (93) bus skaičiuojama remiantis formule

$$\hat{S}(x) = \sum_{k=1}^{\infty} \mathbb{P}(Z = k) \exp\Big\{ - k \int_0^x \mu_t \mathrm{d}t\Big\}. \tag{101}$$

Taigi, Poisson - Gompertz modelio išgyvenamumo funkcija bus lygi:

$$\begin{aligned}
\hat{S}_{PG}(x) &= \mathrm{e}^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \mathrm{e}^{-kB\int_0^x \mathrm{e}^{\alpha t}\mathrm{d}t} \\
&= \exp\Big\{ - \lambda - B \int_0^x \mathrm{e}^{\alpha t}\mathrm{d}t\Big\} \sum_{k=1}^{\infty} \frac{1}{(k-1)!} \Big(\lambda \mathrm{e}^{-B\int_0^x \mathrm{e}^{\alpha t}\mathrm{d}t}\Big)^{k-1} \\
&= \exp\left\{ \frac{B}{\alpha}(1 - \mathrm{e}^{\alpha x}) + \lambda\Big(\mathrm{e}^{\frac{B}{\alpha}(1-\mathrm{e}^{\alpha x})} - 1\Big)\right\}
\end{aligned}$$

su teigiamais parametrais α, B ir λ.

Atitinkamai, Poisson - Gompertz modelio mirtingumo galios funkcija yra

$$\widehat{\mu}_x = Be^{\alpha x}\left(\lambda e^{\frac{B}{\alpha}(1-e^{\alpha x})} + 1\right). \tag{102}$$

10.5.4 Geometrinis - Gompertz modelis

Geometriniame – Gompertz modelyje, mirtingumo galios funkcija turi Gompertz išraišką (98), o a.d. Z - paslinktą geometrinį skirstinį, t.y.

$$\mathbb{P}(Z = k) = p(1-p)^{k-1},\ k \in \mathbb{N}.$$

Kadangi Z yra diskretus, naudojantis formule (101), gauname

$$\begin{aligned}\hat{S}_{GEG}(x) &= \sum_{k=1}^{\infty} p(1-p)^{k-1} e^{-k\int_0^x Be^{\alpha t}dt} \\ &= pe^{-B\int_0^x e^{\alpha t}dt} \sum_{k=1}^{\infty}\left((1-p)e^{-B\int_0^x e^{\alpha t}dt}\right)^{k-1} \\ &= \frac{pe^{\frac{B}{\alpha}(1-e^{\alpha x})}}{1-(1-p)e^{\frac{B}{\alpha}(1-e^{\alpha x})}}\end{aligned}$$

su parametrais $\alpha > 0$, $B > 0$ ir $p \in (0, 1)$.

Modelio mirtingumo galios funkcijos pavidalas:

$$\widehat{\mu}_x = \frac{Be^{\alpha x}}{1-(1-p)e^{\frac{B}{\alpha}(1-e^{\alpha x})}}. \tag{103}$$

10.5.5 Diskretus - Weibull modelis

Diskretaus - Weibull atveju, skirtingai negu Gamma - Weibull modelyje, darysime prielaidą, kad mirtingumo galia turi Weibull funkcijos pavidalą su modaliniu mirties amžiumi, t.y.

$$\mu_x = \frac{1}{\sigma}\left(\frac{x}{M}\right)^{\frac{M}{\sigma}-1}$$

su teigiamais parametrais M and σ. M - modalinis mirties amžius, ties kuriuo populiacijoje įvyksta daugiausiai mirčių.

Diskrečiame - Weibull modelyje, atsitiktinis efektas Z yra diskretusis ir baigtinis. Nagrinėsime atvejį, kur Z įgyja reikšmes $1, 2, 3$ su tikimybėmis $\mathbb{P}(Z = 1) = p$, $\mathbb{P}(Z = 2) = q$, $\mathbb{P}(Z = 3) = 1 - p - q$, kur $p, q \in [0, 1]$ ir $p + q \leqslant 1$.

Remiantis formule (101), gauname modelio išgyvenamumo funkciją

$$\hat{S}_{DW}(x) = p\,\mathrm{e}^{-\left(\frac{x}{M}\right)^{M/\sigma}} + q\,\mathrm{e}^{-2\left(\frac{x}{M}\right)^{M/\sigma}} + (1-p-q)\mathrm{e}^{-3\left(\frac{x}{M}\right)^{M/\sigma}}.$$

Mirtingumo galia bus lygi

$$\hat{\mu}_x = \frac{1}{\sigma}\left(\frac{x}{M}\right)^{M/\sigma - 1} \frac{p\,\mathrm{e}^{-\left(\frac{x}{M}\right)^{M/\sigma}} + 2q\,\mathrm{e}^{-2\left(\frac{x}{M}\right)^{M/\sigma}} + 3(1-p-q)\mathrm{e}^{-3\left(\frac{x}{M}\right)^{M/\sigma}}}{p\,\mathrm{e}^{-\left(\frac{x}{M}\right)^{M/\sigma}} + q\,\mathrm{e}^{-2\left(\frac{x}{M}\right)^{M/\sigma}} + (1-p-q)\mathrm{e}^{-3\left(\frac{x}{M}\right)^{M/\sigma}}}. \quad (104)$$

Svarbu pažymėti, jog atveju $p = 1$ Diskretaus - Weibull modelio mirtingumo galia (104) pasidaro lygi klasikinei Weibull mirtingumo galios funkcijai su modaliniu mirties amžiumi, aprašytai šio skyrelio pradžioje.

10.6 Praktinis taikymas Baltijos šalims

10.6.1 Baltijos šalių mirtingumo analizė

Mirtingumo duomenims analizuoti buvo pasirinkti Lietuvos, Latvijos ir Estijos mirtingumo duomenys po 2000 metų. Duomenys yra viešai prieinami ir buvo gauti iš Human Mortality Database duombazės (www.mortality.org). Paprastumo dėlei buvo nagrinėjami tik bendri duomenys, t.y. neskaidant pagal lytį. Nors ir akivaizdu, kad moterų ir vyrų mirtingumas skiriasi, tačiau gyvybės draudimo rinkoje kompanijos negali naudoti skirtingų mirtingumo lentelių vyrams ir moterims norint paskaičiuoti draudimo įmoką, todėl analizėje bus naudojama bendra lentelė. Į duomenis įtraukti individai nuo 0 iki 110 metų amžiaus. Žmonės, esantys vyresni nei 110 metų amžiaus, buvo įtraukti į 110 metų grupę dėl mažos imties. Duomenyse buvo pateiktos šios charakteristikos kiekviena amžiaus grupei: išgyvenimo tikimybės ir centrinis mirtingumo dažnis. Mirtingumo lentelėms sudaryti dažniausiai naudojami atsižvelgiant į ilgesnio laikotarpio duomenis, taigi, analizėje buvo pasielgta analogiškai - išgyvenimo tikimybės ir centriniai mirtingumo dažniai buvo paskaičiuojami imant aritmetinį vidurkį iš visų surinktų duomenų nuo 2000 metų.

Nagrinėti duomenys pateikti 11 skyriaus 5 lentelėje. Duomenyse $\widetilde{S}(x)$ žymi empirines išgyvenamumo tikimybes, o $\widetilde{m}_x$ žymi empirines centrinio mirtingumo dažnio reikšmes kiekvienam amžiui $x \in \{0, 1, ..., 110\}$.

Norint panaudoti empirinius duomenis, analizėje naudosime anksčiau minėtą prielaidą, jog mirtingumo galia yra pastovi intervale $[x, x+1)$. Tai reiškia, kad $\mu_{x+t} = \mu_x$ su kiekvienu $x \in \mathbb{N}_0 = \{0, 1, ...\}$ ir $t \in [0, 1)$. Remiantis šiomis sąlygomis, centrinis mirtingumo dažnis gali būti prilygintas empirinei mirtingumo galiai, t.y. $\widetilde{m}_x = \widetilde{\mu}_x,\ x \in \mathbb{N}_0$.

Ankstesniuose skyreliuose nagrinėtų mirtingumo modelių pritaikymas kiekvienai populiacijai buvo atliktas naudojant R programavimo kalbą ir paketą *MortalityLaws*. Remiantis pasirinkta paklaidos funkcija, buvo paskaičiuoti tiksliausi kiekvieno modelio funkcijų parametrai programai pateikiant empirinius mirtingumo dažnio duomenis. Analizėje buvo naudota viena populiariausių metrikų (paklaidos funkcijų) - vidutinė kvadratinė paklaida (angl. *mean squared error*):

$$MSE = \frac{\sum\limits_{x=1}^{N} (\widetilde{m}(x) - \hat{\mu}(x))^2}{N},$$

kur $\widetilde{m}(x)$ - empirinė mirtingumo galia, $\hat{\mu}(x)$ - sumodeliuota mirtingumo galia ir N - stebėjimų skaičius. Šios analizės atveju, $N = 110$.

MSE taip pat naudojame lyginant gretimus mirtingumo modelius. Žemiau pateikiame 10.5 skyrelyje nagrinėtų mirtingumo modelių palyginimus, mirtingumo galių grafikus ir vidutines kvadratines paklaidas kiekvienos šalies populiacijai.

Pradžioje pateikiame Lietuvos populiacijos rezultatus. Latvijos ir Estijos populiacijų rezultatai pateikti 7.2 ir 7.3 skyreliuose.

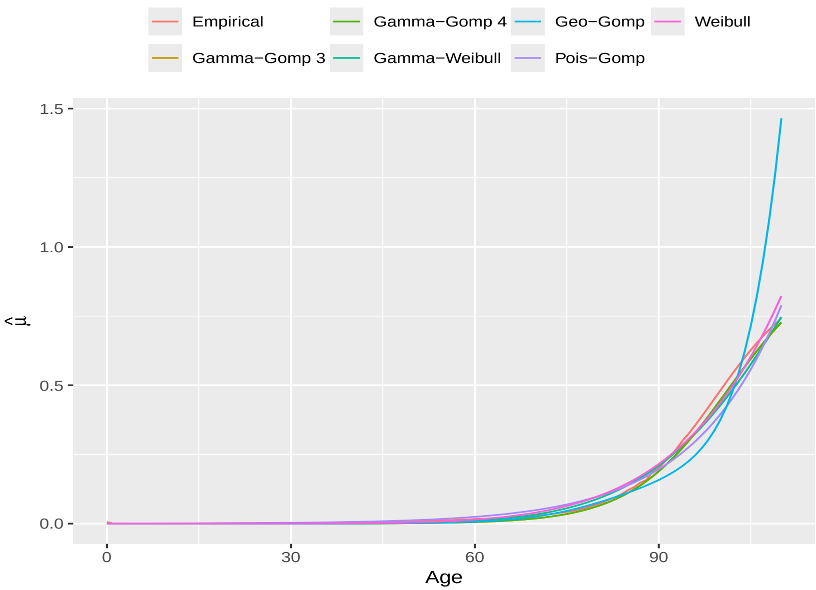

Figure 20: Atsitiktinio efekto mirtingumo modeliai pritaikyti Lietuvos duomenims (visos amžiaus grupės)

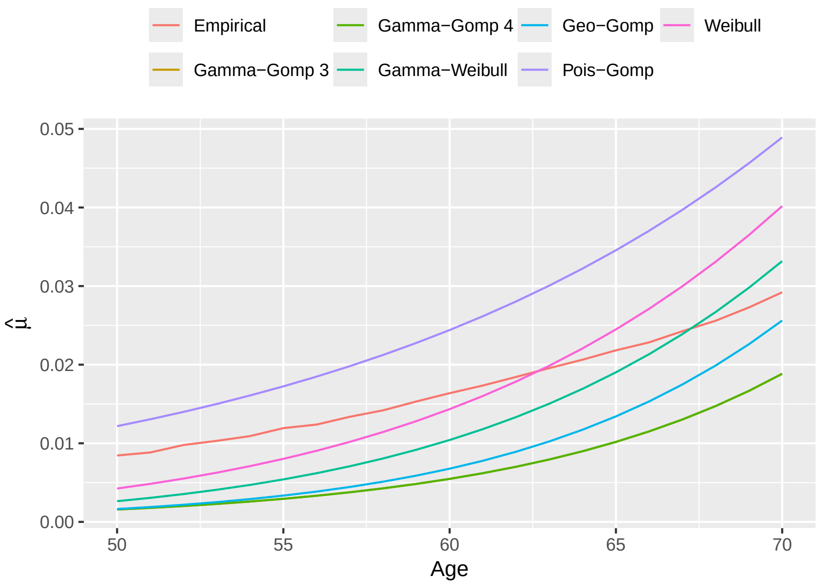

Figure 21: Atsitiktinio efekto mirtingumo modeliai pritaikyti Lietuvos duomenims (50-70 metų amžius)

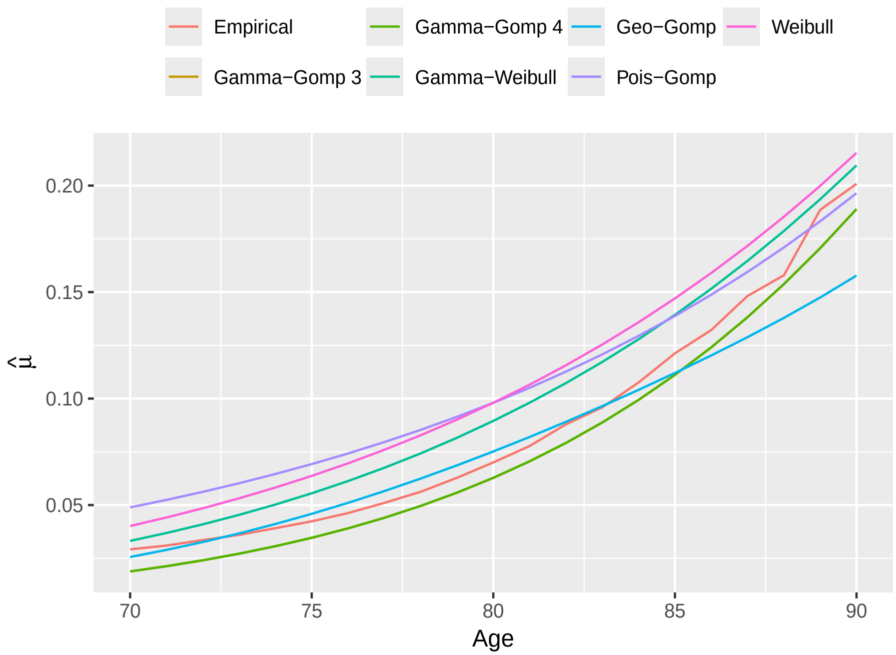

Figure 22: Atsitiktinio efekto mirtingumo modeliai pritaikyti Lietuvos duomenims (70-90 metų amžius)

Table 4: Kiekvieno modelio MSE Lietuvos populiacijai

Modelis	**Paklaidos funkcija (MSE)**
Gamma - Gompertz (keturi parametrai)	0.0001820656
Gamma - Gompertz (trys parametrai)	0.0001824057
Gamma - Weibull	0.0003312705
Weibull (Diskretus - Weibull su $p = 1$)	0.0003361133
Poisson - Gompertz	0.0008037022
Geometrinis - Gompertz	0.01091893

Modelių mirtingumo galios funkcijų parametrai Lietuvos duomenims:

• Gamma - Gompertz su keturiais parametrais (žr. (99)):
$\{B = 5.136861\times10^{-7},\ \alpha = 0.1250529,\ \lambda = 1.320658,\ k = 7.800965\}$;

• Gamma - Gompertz su trimis parametrais (žr. (100)):
$\{B = 3.028194 \times 10^{-6},\ \alpha = 0.1250734,\ \lambda = 7.798495\}$;

• Gamma - Weibull (žr. (97)):
$\{a = 5.497963 \times 10^{-13}, k = 9347.058, \ \lambda = 403176.2 \times 10^{12}, \ \alpha = 6.752323\}$;

• Weibull (Diskretus - Weibull su $p = 1$) (žr. (104)):
$\sigma = 10.38606, \ M = 79.78316, \ p = 1$;

• Poisson - Gompertz (žr. (102)):
$\{B = 0.0003768185, \ \alpha = 0.06951546, \ \lambda = 5.637439 \times 10^{-7}\}$;

• Geometrinis - Gompertz (žr. (103)):
$\{\alpha = 1.71132 \times 10^{-7}, \ B = 0.1451152, \ p = 0.1467799\}$.

Iš gautų rezultatų daroma išvada, jog Lietuvos ir Latvijos populiacijų mirtingumą geriausiai nusako Gamma - Gompertz (tiek 3, tiek 4 parametrų) modeliai, o Estijos populiacijai geriausius rezultatus duoda Diskretus - Weibull (su $p < 1$) modelis. Lietuvos atveju Gamma - Gompertz 3 ir 4 parametrų modelių vidutinės kvadratinės paklaidos ir modeliuotos mirtingumo dažnio reikšmės beveik sutampa, todėl grafikuose nubrėžtos kreivės yra beveik identiškos. Iš rezultatų matyti, jog su visais nagrinėtais modeliais gauta pakankamai maža vidutinė kvadratinė paklaida, taigi, visos lentelėse pateiktos transformuotos mirtingumo galios funkcijos yra tinkamos aproksimuoti nagrinėtų populiacijų mirtingumą, išskyrus Geometrinį - Gompertz modelį, kurio mirtingumo galios funkcija yra panaši į laiptinio pavidalo funkciją, o tai nėra tinkama apibūdinti mirtingumui. Taip pat įdomu pastebėti, jog Gamma - Gompertz modeliai linkę nuvertinti mirtingumo galią lyginant su realiais duomenimis.

10.6.2 Lietuvos prostatos vėžio stebėsenos programos rezultatai

Šiame poskyryje analizuojami prostatos vėžio stebėsenos (angl. screening) programos rezultatai Lietuvos duomenims, surinktiems 2006 m. sausio 1 d. ir 2016 m. gruodžio 31 d. laikotarpiu. Duomenis rinko ir jais pasidalinti sutiko Nacionalinis Vėžio Institutas. Programoje dalyvavo 50 - 75 metų amžiaus vyrai, o duomenys buvo skirstomi į dvi imtis:

(i) vyrai, kurie dalyvavo prostatos vėžio stebėsenos programoje ir kuriems buvo atliekami PSA (prostatos specifinio antigeno) testai,

(ii) vyrai, kurie nedalyvavo programoje, tačiau buvo stebimi nurodytu laikotarpiu.

Naudotų duomenų pavyzdys yra pateiktas 11 skyriaus 6 lentelėje.

Analizė buvo dalinama į dvi dalis: pirmoje buvo nagrinėjamas prostatos vėžio mirtingumas, o antroje - bendras mirtingumas (įskaitant prostatos vėžį). Remiantis gimimo, mirties, pirmų PSA testų, išvykimo, stebėsenos pradžios ir pabaigos datomis, buvo išskaičiuotos 1 metų mirties tikimybės $q_x = {}_1q_x$ (pagal (79) formulę), kur x žymi laiko momentą, kuris buvo konstruojamas pagal dvi taisykles:

(i) dalyvavusiems programoje, pirmasis momentas $x = 0$ buvo nustatytas atsižvelgiant į pirmo PSA testo datą kiekvienam individui,

(ii) nedalyvavusiems programoje, pirmasis momentas $x = 0$ buvo nustatytas atsižvelgiant į datą, kada individas buvo pradėtas stebėti (t.y. metai, kada individui suėjo 50 metų arba kalendoriniai 2006 metai (jeigu 50 metų individui suėjo anksčiau šio laiko)).

Empirinės mirtingumo tikimybės buvo naudojamos Weibull mirtingumo modeliui su parametru $p = 1$, aprašytam 10.5.5 skyrelyje. Modelio parametrų M ir σ skaičiavimui buvo naudojama *R* programa ir *MortalityLaws* biblioteka, kuriose empirinės mirtingumo tikimybės buvo pateiktos įvertinimui su binomine paklaidos skaičiavimo funkcija

$$LF = -\sum_{x=0}^{10} \Big(D_x \ln\left(1 - \mathrm{e}^{-\mu_x}\right) - (E_x - D_x)\mu_x \Big),$$

kur D_x yra mirčių skaičius laiko momentu x, E_x yra populiacijos skaičius laiko momentu x, μ_x - mirtingumo galia laiko momentu x. Visi įverčiai buvo paskaičiuoti atsižvelgiant į empirines mirtingumo tikimybes.

Prostatos vėžio mirtingumui aproksimuoti, paskaičiuoti Weibull modelio parametrai yra:

(i) dalyvavusiems programoje: $M = 20.93804$, $\sigma = 4.150694$,

(ii) nedalyvavusiems programoje: $M = 28.58124$, $\sigma = 8.154898$.

Pirmiems 10 metų $x \in [0, 10]$, išgyvenamumo funkcijos $S(x)$ kreivės nubrėžtos 23 grafike.

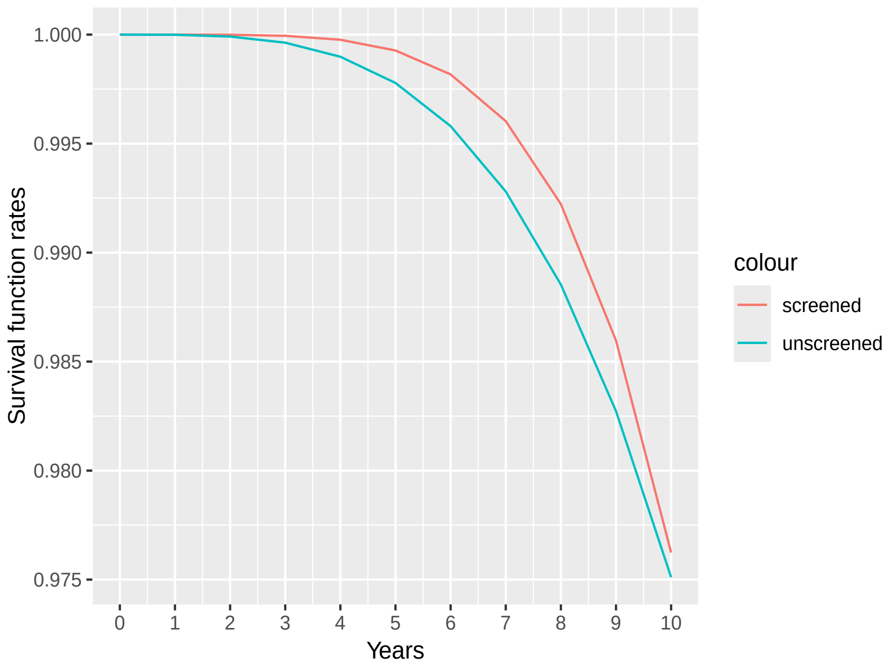

Figure 23: Išgyvenimo funkcijos prostatos vėžio mirtingumui

Naudojantis sąryšiais iš 10.3.2 skyrelio, gaunamos šios charakteristikos:

(i) Dalyvavusiems programoje: tikimybė išgyventi bent 6 metus $S(6) = 0.998174$, tikimybė išgyventi bent 8 metus $S(8) = 0.992229$, 10 metų likusi gyvenimo trukmė $\mathring{e}_{0,\overline{10}|} = 9.96048$.

(ii) nedalyvavusiems programoje: tikimybė išgyventi bent 6 metus $S(6) = 0.995802$, tikimybė išgyventi bent 8 metus $S(8) = 0.988535$, 10 metų likusi gyvenimo trukmė $\mathring{e}_{0,\overline{10}|} = 9.94444$.

Bendram mirtingumui (įskaitant ir prostatos vėžį) įvertinti, paskaičiuoti Weibull modelio parametrai yra:

(i) dalyvavusiems programoje: $M = 10.34769$, $\sigma = 1.956164$,

(ii) nedalyvavusiems programoje: $M = 9.568636$, $\sigma = 3.62983$.

Pirmiems 10 metų $x \in [0, 10]$, išgyvenamumo funkcijos $S(x)$ kreivės nubrėžtos 24 grafike.

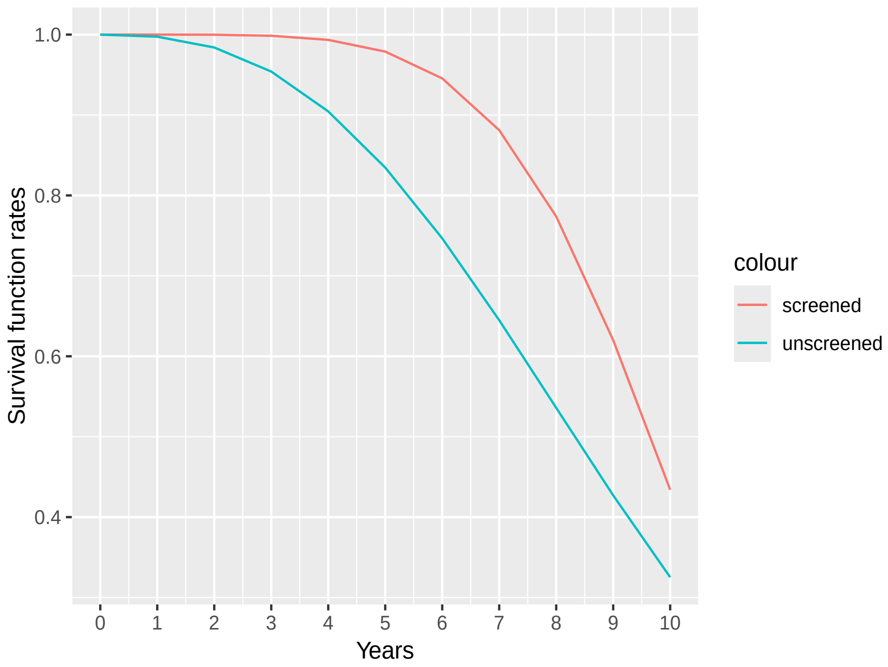

Figure 24: Išgyvenimo funkcijos bendram mirtingumui

Naudojantis sąryšiais iš 10.3.2 skyrelio, gaunamos šios charakteristikos:

(i) Dalyvavusiems programoje: tikimybė išgyventi bent 6 metus $S(6) = 0.945141$, tikimybė išgyventi bent 8 metus $S(8) = 0.773109$, 10 metų likusi gyvenimo trukmė $\overset{\circ}{e}_{0,\overline{10}|} = 8.92152$.

(ii) nedalyvavusiems programoje: tikimybė išgyventi bent 6 metus $S(6) = 0.746628$, tikimybė išgyventi bent 8 metus $S(8) = 0.535927$, 10 metų likusi gyvenimo trukmė $\overset{\circ}{e}_{0,\overline{10}|} = 7.69986$.

Remiantis rezultatais gautais su Diskrečiu - Weibull modeliu, daroma išvada, jog prostatos vėžio stebėsenos programa Lietuvoje gali būti laikoma sėkminga ir kad programoje dalyvavusių individų likusi dalinė gyvenimo trukmė yra ilgesnė lyginant su individais, kurie nedalyvavo programoje.

10.7 Išvados

Disertacijoje autorius analizuoja atsitiktinio dydžio X, apibrėžto tikimybinėje erdvėje $(\mathbf{W}, \mathcal{F}, \mathbf{P})$, pasiskirstymo funkcijos uodegą:

$$\overline{F}_X(x) := \mathbf{P}(X > x) = 1 - F_X(x),\ x \in \mathbb{R},$$

kur

$$F_X(x) = \mathbf{P}(X \leqslant x) = \mathbf{P}(w \in \mathbf{W} : X(w) \leqslant x),\ x \in \mathbb{R}$$

yra a.d. X pasiskirstymo funkcija. F_X ir $\overline{F}_X$ charakterisitkos yra aprašomos ir analizuojamos pateikiant jų savybes ir pavyzdžius. Funkcijų uodegų elgesys yra aprašomas skirstant funkcijas į atskiras kategorijas, vadinamas reguliarumo klasėmis.

Pagrindinis disertacijos tikslas yra patikrinti ar pasiskirstymo funkcijos uodegos elgesys ir savybės pasikeičia, jeigu uodega yra transformuojama. Transformacija yra atliekama uodegos funkcijai F_X pritaikant teigiamą atsitiktinį dydį Z:

$$\overline{F^{(Z)}}(x) := \mathbb{E}\Big(\overline{F}^Z(x)\Big) = \mathbb{E}\Big(\big(\overline{F}(x)\big)^Z\Big),\ x \in \mathbb{R}.$$

Transformuota pasiskirstymo funkcija lygi

$$F^{(Z)}(x) = 1 - \overline{F^{(Z)}}(x) = 1 - \mathbb{E}\Big(\overline{F}^Z(x)\Big),\ x \in \mathbb{R}.$$

Vidurkiui naudojamas simbolis $\mathbb{E}$, nes bendru atveju Z gali būti apibrėžtas kitoje tikimybinėje erdvėje $(\Omega, \mathcal{A}, \mathbb{P})$.
Disertacijos autorius pateikia teoremą, kuri teigia, jog $F^{(Z)}$ yra nauja pasiskirstymo funkcija, tenkinanti visus reikalavimus. Negana to, transformuota pasiskirstymo funkcijos uodega išlaiko savo savybes specifinių reguliarumo klasių kontekste. Autorius pateikia ir įrodo teoremą, kuri teigia, jog $\overline{F^{(Z)}}$ priklauso originaliai reguliarumo klasei prieš transformaciją:

(i) $F \in \mathcal{R}_\alpha \ \Rightarrow\ F^{(Z)} \in \mathcal{R}_{\alpha a}$, kur a yra didžiausias a.d. Z apatinis rėžis, t.y.:

$$\mathbb{P}(Z \geqslant a) = 1 \text{ ir } \mathbb{P}(Z \geqslant a + \delta) < 1 \text{ bet kuriam } \delta > 0;$$

(ii) $F \in \mathcal{L} \ \Rightarrow\ F^{(Z)} \in \mathcal{L}$;
(iii) $F \in \mathcal{OL} \ \Rightarrow\ F^{(Z)} \in \mathcal{OL}$;
(iv) $F \in \mathcal{D} \ \Rightarrow\ F^{(Z)} \in \mathcal{D}$,

kur F yra kažkokio atsitiktinio dydžio pasiskirstymo funkcija.

Pasiskirstymo funkcijos uodegos gali būti analizuojamos žmonių mirtingumo kontekste. Jeigu T yra neneigiamas ir absoliučiai tolydus atsitiktinis dydis, tada funkcija

$$S(x) = S_T(x) = \mathbf{P}(T > x) = 1 - F_T(x) = \overline{F}_T(x),$$

yra išgyvenamumo funkcija, kuri yra lygi uodegos funkcijai F_T. Išgyvenamumo funkcija naudojama nusakant ir kitas mirtingumo charakteristikas, pavyzdžiui, mirtingumo galią:

$$\mu_x = \tfrac{f(x)}{S(x)} = -\tfrac{S'(x)}{S(x)},\ x \geqslant 0.$$

Disertacijos autorius apibrėžia transformuotą išgyvenamumo funkciją pagal formulę

$$\widehat{S}(x) = \mathbb{E}\left(S^Z(x)\right) = \mathbb{E}\left(\exp\Big\{ - Z \textstyle\int_0^x \mu_u \mathrm{d}u\Big\}\right).$$

Autorius pateikia ir įrodo teoremą, teigiančią, jog $\widehat{S}(x)$ yra nauja išgyvenamumo funkcija, tenkinanti visus reikalavimus. Tai reiškia, jog transformuotos išgyvenamumo ir mirtingumo galios funkcijos gali būti naudojamos aprašant žmonių mirtingumą.

Po teorinės dalies, praktinėje dalyje buvo pateikiamas transformuotų išgyvenamumo ir mirtingumo galios funkcijų pritaikymas dviems duomenų tipams. Pirmajame pateikiamas Baltijos šalių populiacijos mirtingumas nuo 2000 metų skaidant duomenis į Lietuvos, Latvijos ir Estijos populiacijas. Disertacijos autorius minėtiems duomenims pritaiko skirtingus transformuotus mirtingumo modelius ir paskaičiuoja jų parametrus siekiant nustatyti, kurie iš modelių geriausiai aproksimuoja atitinkamų populiacijų mirtingumą. Modelio tinkamumas validuotas pasitelkus vidutinės kvadratinės paklaidos metriką. Atlikus analizę paaiškėjo, kad iš nagrinėtų modelių Lietuvos ir Latvijos mirtingumui aproksimuoti tinkamiausias yra Gamma - Gompertz modelis su šia išgyvenamumo funkcijos išraiška:

$$\hat{S}_{GG_4}(x) = \left(1 + \frac{B}{\alpha\lambda}(\mathrm{e}^{\alpha x} - 1)\right)^{-k}.$$

Estijos duomenims tiksliausi rezultatai gauti su Diskrečiu - Weibull modeliu (su parametru $p < 1$):

$$\hat{S}_{DW}(x) = p\,\mathrm{e}^{-\left(\frac{x}{M}\right)^{M/\sigma}} + q\,\mathrm{e}^{-2\left(\frac{x}{M}\right)^{M/\sigma}} + (1-p-q)\mathrm{e}^{-3\left(\frac{x}{M}\right)^{M/\sigma}}.$$

Diskretus - Weibull modelis su $p = 1$ taip pat naudojamas antrajam duomenų tipui, kuriame pateikta prostatos vėžio stebėsenos programos Lietuvoje statistika, apimanti 50 - 75 metų amžiaus vyrus 2006 - 2016 metais. Duomenims pritaikius mirtingumo modelį buvo paskaičiuotos 10 metų išgyvenimo tikimybės. Po atliktos analizės paaiškėjo, kad programoje dalyvaujantys individai turi didesnes išgyvenimo tikimybes tiek nuo prostatos vėžio, tiek nuo bendro mirtingumo lyginant su individais, kurie prostatos vėžio stebėsenos programoje nedalyvavo.

Tiek praktinė, tiek teorinė disertacijos dalys rodo, kad atlikus transformaciją galima gauti aibę naujų išgyvenamumo ir mirtingumo galios funkcijų. Transformuotos funkcijos gali būti toliau analizuojamos visose įmanomose reguliarumo klasėse (įskaitant ir klases, nagrinėtas šioje disertacijoje). Negana to, disertacijoje parodyta, jog naujos funkcijos gali būti plačiai pritaikomos praktiniams klausimams medicinoje, finansuose, draudime. Svarbu pažymėti, kad disertacijoje aprašyti mirtingumo modeliai negali būti laikomi vienareikšmiškai geriausiais mirtingumo prognozavimui, kadangi modelio pasirinkimas stipriai priklauso nuo nagrinėjamos populiacijos. Ši disertacija apima tik nedidelę dalį mirtingumo modelių. Taigi, tiksliausio mirtingumo prognozavimui modelio ieškojimas ir toliau išlieka vienu iš neišspręstų uždavinių matematikams ir gyvybės draudimo rinkai.

10.8 Rezultatų sklaida

Disertacijos rezultatai publikuojami šiuose moksliniuose straipsniuose:

- Šiaulys, J.; Puišys, R. Survival with random effect. *Mathematics* **2022**, *10*, 1097.
- Puišys, R.; Lewkiewicz, S.; Šiaulys, J. Properties of the random effect transformation. *Lith. Math. J.* **2024**, *64*, 177-189.
- Skučaitė, A.; Puvačiauskienė, A.; Puišys, R.; Šiaulys, J. Actuarial analysis of survival among breast cancer patients in Lithuania. *Healthcare* **2021**, *9*, 383.
- Levickytė, J.; Skučaitė, A.; Šiaulys, J.; Puišys, R.; Vincerževskienė, I. Actuarial analysis of survival after breast cancer diagnosis among Lithuanian females. *Healthcare* **2024**, *12*, 746.

Disertacijos rezultatai buvo paskelbti šiuose moksliniuose renginiuose:

- R. Puišys. *Kelios išgyvenamumo funkcijos su atsitiktiniu efektu*, Lietuvos Matematikų Draugijos konferencija, 2021 m. birželio 16-17 d.
- R. Puišys, J. Šiaulys, S.Lewkiewicz. *Kelios atsitiktinio efekto transformacijos savybės*. Lietuvos Matematikų Draugijos konferencija, 2023 m. birželio 21-22 d., Vilnius, Lietuva.
- R. Puišys, J. Šiaulys. *Survival with random effect*. Data Analysis Methods for Software Systems, 2023 m. lapkričio 30 d. - gruodžio 2 d., Druskinininkai, Lietuva.
- R. Puišys, J. Šiaulys. *Survival with random effect*. 27th International Congress on Insurance: Mathematics and Economics, 2024 m. liepos 8-11 d. Čikaga, Ilinojus, Jungtinės Amerikos Valstijos.

10.9 Trumpos žinios apie autorių

Išsilavinimas:

- 2014-2018 m., Vilniaus universitetas, Matematikos ir informatikos fakultetas, finansų ir draudimo matematikos bakalauras.
- 2018-2020 m., Vilniaus universitetas, Matematikos ir informatikos fakultetas, finansų ir draudimo matematikos magistras.
- 2020-2024 m., Vilniaus universitetas, Matematikos ir informatikos fakultetas, gamtos mokslų matematikos kryptis, doktorantūros studijos.

11 Appendix

11.1 Baltic population mortality table

	Estonia		Latvia		Lithuania	
Age	$\widetilde{m}_x$	$\widetilde{S}(x)$	$\widetilde{m}_x$	$\widetilde{S}(x)$	$\widetilde{m}_x$	$\widetilde{S}(x)$
0	0.0045	1	0.0071	1	0.0058	1
1	0.0005	0.9955	0.0006	0.9929	0.0005	0.9942
2	0.0003	0.995	0.0004	0.9923	0.0004	0.9937
3	0.0002	0.9946	0.0003	0.9919	0.0003	0.9933
4	0.0002	0.9944	0.0003	0.9916	0.0003	0.993
5	0.0003	0.9942	0.0003	0.9913	0.0002	0.9928
6	0.0002	0.9939	0.0002	0.9909	0.0002	0.9925
7	0.0002	0.9937	0.0003	0.9907	0.0002	0.9923
8	0.0002	0.9935	0.0002	0.9904	0.0002	0.9921
9	0.0001	0.9933	0.0002	0.9902	0.0002	0.9919
10	0.0002	0.9932	0.0002	0.9899	0.0002	0.9917
11	0.0002	0.993	0.0002	0.9897	0.0002	0.9915
12	0.0002	0.9928	0.0002	0.9895	0.0002	0.9913
13	0.0002	0.9927	0.0002	0.9893	0.0002	0.9911
14	0.0002	0.9924	0.0003	0.9891	0.0003	0.9909
15	0.0004	0.9922	0.0003	0.9888	0.0004	0.9905
16	0.0004	0.9918	0.0006	0.9884	0.0005	0.9901
17	0.0005	0.9915	0.0006	0.9879	0.0007	0.9896
18	0.0008	0.9909	0.0007	0.9873	0.0009	0.989
19	0.0008	0.9902	0.0009	0.9866	0.0011	0.988
20	0.0011	0.9894	0.001	0.9857	0.0011	0.987
21	0.001	0.9883	0.001	0.9848	0.0012	0.9859
22	0.0011	0.9874	0.0011	0.9838	0.0012	0.9847
23	0.0011	0.9862	0.0011	0.9828	0.0012	0.9835
24	0.0012	0.9851	0.0011	0.9817	0.0013	0.9823
25	0.0013	0.984	0.0013	0.9806	0.0014	0.981
26	0.0013	0.9827	0.0013	0.9793	0.0016	0.9797
27	0.0013	0.9814	0.0015	0.9781	0.0016	0.9781
28	0.0014	0.9801	0.0015	0.9767	0.0017	0.9766
29	0.0015	0.9787	0.0018	0.9752	0.0018	0.9749
30	0.0016	0.9773	0.0017	0.9734	0.0019	0.9731
31	0.0017	0.9757	0.0019	0.9718	0.0022	0.9712
32	0.0018	0.974	0.0021	0.9699	0.0023	0.9691
33	0.0018	0.9723	0.0023	0.9679	0.0024	0.9669
34	0.0018	0.9706	0.0026	0.9657	0.0026	0.9646

35	0.002	0.9689	0.0026	0.9632	0.0029	0.9621
36	0.0019	0.9669	0.0028	0.9607	0.0028	0.9593
37	0.0022	0.9651	0.0031	0.9581	0.0032	0.9566
38	0.0024	0.9629	0.0033	0.9551	0.0034	0.9536
39	0.0025	0.9606	0.0038	0.952	0.0037	0.9504
40	0.0028	0.9582	0.0039	0.9484	0.0039	0.9469
41	0.0031	0.9555	0.0039	0.9447	0.0044	0.9432
42	0.0034	0.9526	0.0044	0.941	0.0046	0.9391
43	0.0037	0.9494	0.0047	0.9369	0.0049	0.9348
44	0.0042	0.946	0.0053	0.9325	0.0051	0.9302
45	0.0045	0.942	0.0057	0.9276	0.0058	0.9255
46	0.0051	0.9378	0.006	0.9224	0.0061	0.9202
47	0.0054	0.9331	0.0066	0.9168	0.0068	0.9146
48	0.0056	0.9281	0.0069	0.9108	0.0072	0.9084
49	0.0062	0.9229	0.0082	0.9046	0.0076	0.9018
50	0.0069	0.9172	0.0083	0.8972	0.0084	0.895
51	0.0074	0.911	0.0089	0.8898	0.0088	0.8875
52	0.0079	0.9043	0.0098	0.8819	0.0098	0.8797
53	0.0086	0.8972	0.0105	0.8734	0.0103	0.8712
54	0.0092	0.8896	0.0114	0.8644	0.0109	0.8623
55	0.0101	0.8815	0.0121	0.8546	0.0119	0.853
56	0.0106	0.8727	0.0129	0.8444	0.0124	0.8429
57	0.0116	0.8637	0.0136	0.8336	0.0134	0.8326
58	0.0122	0.8538	0.0146	0.8224	0.0142	0.8215
59	0.0132	0.8436	0.0162	0.8105	0.0153	0.81
60	0.0144	0.8326	0.0167	0.7976	0.0164	0.7977
61	0.0154	0.8208	0.0179	0.7844	0.0174	0.7848
62	0.0159	0.8084	0.0196	0.7705	0.0184	0.7714
63	0.0176	0.7957	0.021	0.7556	0.0196	0.7573
64	0.0186	0.7819	0.0224	0.7399	0.0206	0.7427
65	0.0201	0.7677	0.0236	0.7236	0.0218	0.7276
66	0.0213	0.7526	0.0245	0.7068	0.0228	0.7119
67	0.0221	0.7369	0.0266	0.6898	0.0243	0.6959
68	0.0244	0.721	0.0279	0.6717	0.0256	0.6792
69	0.0256	0.7038	0.0303	0.6534	0.0273	0.6621
70	0.0279	0.6862	0.0322	0.634	0.0292	0.6443
71	0.0298	0.6675	0.0344	0.614	0.031	0.6258
72	0.0328	0.6482	0.0377	0.5934	0.0336	0.6067
73	0.0362	0.6276	0.0409	0.5715	0.036	0.5868
74	0.0385	0.6056	0.0445	0.5488	0.0392	0.5661
75	0.0419	0.5831	0.048	0.5251	0.0424	0.5444
76	0.0453	0.5596	0.0523	0.5007	0.0462	0.522

77	0.0496	0.5352	0.0564	0.4754	0.051	0.4985
78	0.0548	0.5098	0.0624	0.4496	0.0562	0.4738
79	0.0595	0.4831	0.0682	0.4226	0.0628	0.448
80	0.0683	0.4557	0.0769	0.395	0.07	0.4209
81	0.0753	0.4262	0.0853	0.3661	0.0777	0.3926
82	0.0839	0.3959	0.0939	0.3365	0.0879	0.3634
83	0.0905	0.3648	0.1053	0.3067	0.0959	0.333
84	0.1036	0.3339	0.1169	0.2764	0.1076	0.3026
85	0.1134	0.3017	0.1264	0.2463	0.1212	0.2719
86	0.1305	0.2701	0.1457	0.2174	0.1322	0.241
87	0.1438	0.2378	0.1592	0.1883	0.1482	0.2112
88	0.1589	0.2066	0.1751	0.161	0.1579	0.1822
89	0.1766	0.1769	0.1944	0.1354	0.1886	0.1557
90	0.1946	0.1488	0.2148	0.1118	0.2008	0.1289
91	0.2151	0.123	0.2375	0.0904	0.2247	0.1055
92	0.237	0.0996	0.2588	0.0715	0.246	0.0843
93	0.2607	0.079	0.2833	0.0553	0.2735	0.0659
94	0.2858	0.0611	0.3093	0.0417	0.3029	0.0501
95	0.3126	0.0461	0.3366	0.0306	0.3283	0.0369
96	0.3407	0.0339	0.3651	0.0219	0.3572	0.0265
97	0.3702	0.0242	0.3946	0.0152	0.3871	0.0185
98	0.4007	0.0168	0.425	0.0102	0.4177	0.0125
99	0.4321	0.0113	0.456	0.0066	0.4488	0.0082
100	0.4642	0.0073	0.4873	0.0042	0.4799	0.0052
101	0.4966	0.0046	0.5188	0.0025	0.5109	0.0032
102	0.5291	0.0028	0.55	0.0015	0.5415	0.0019
103	0.5615	0.0016	0.5809	0.0009	0.5712	0.0011
104	0.5933	0.0009	0.611	0.0005	0.5999	0.0006
105	0.6244	0.0005	0.6402	0.0003	0.6274	0.0003
106	0.6545	0.0003	0.6684	0.0001	0.6534	0.0002
107	0.6834	0.0001	0.6952	0.0001	0.678	0.0001
108	0.7109	0.0001	0.7207	0	0.7009	0
109	0.7369	0	0.7446	0	0.7223	0
110	0.7613	0	0.767	0	0.7422	0

Table 5: Empirical data table of Baltic population mortality in XXI century

11.2 Lithuanian prostate screening program data sample

Death	Birth	Departure	Death cause	Start	End	PSA	Age at PSA
-	1959-10-20	-	-	2009-10-20	2016-12-31	2010-01-21	51
-	1959-10-20	-	-	2009-10-20	2016-12-31	2011-04-13	52
-	1959-10-20	-	-	2009-10-20	2016-12-31	2012-10-18	53
-	1959-10-20	-	-	2009-10-20	2016-12-31	2009-12-04	50
-	1959-10-20	-	-	2009-10-20	2016-12-31	2014-01-27	55
-	1959-10-20	-	-	2009-10-20	2016-12-31	2010-06-03	51
-	1959-10-20	-	-	2009-10-20	2016-12-31	2013-02-05	54
2013-06-01	1959-10-20	-	X700	2009-10-20	2013-06-01	2011-07-07	52
-	1959-10-20	-	-	2009-10-20	2016-12-31	2014-07-11	55
-	1959-10-20	-	-	2009-10-20	2016-12-31	2011-12-05	52
-	1959-10-20	-	-	2009-10-20	2016-12-31	2010-05-28	51
-	1959-10-20	-	-	2009-10-20	2016-12-31	2014-10-02	55
2014-11-21	1959-10-20	-	C348	2009-10-20	2014-11-21	2011-11-15	52
-	1959-10-21	-	-	2009-10-21	2016-12-31	2011-05-20	52
-	1959-10-21	-	-	2009-10-21	2016-12-31	2010-12-10	51
-	1959-10-21	-	-	2009-10-21	2016-12-31	2010-11-23	51
-	1959-10-21	-	-	2009-10-21	2016-12-31	2012-05-28	53
-	1959-10-21	-	-	2009-10-21	2016-12-31	2011-06-09	52
-	1959-10-21	-	-	2009-10-21	2016-12-31	2010-02-04	51
-	1959-10-21	-	-	2009-10-21	2016-12-31	2010-02-09	51
-	1959-10-21	-	-	2009-10-21	2016-12-31	2011-12-02	52
-	1959-10-21	-	-	2009-10-21	2016-12-31	2014-12-03	55
-	1959-10-21	-	-	2009-10-21	2016-12-31	2014-09-05	55
-	1959-10-21	-	-	2009-10-21	2016-12-31	2012-04-24	53
-	1959-10-21	-	-	2009-10-21	2016-12-31	2010-04-01	51
-	1959-10-21	-	-	2009-10-21	2016-12-31	2010-06-17	51
-	1959-10-21	-	-	2009-10-21	2016-12-31	2012-07-31	53
-	1959-10-21	2017-09-30	-	2009-10-21	2016-12-31	2014-02-11	55
-	1959-10-21	-	-	2009-10-21	2016-12-31	2016-12-29	57
-	1959-10-21	-	-	2009-10-21	2016-12-31	2009-03-23	50
-	1959-10-21	-	-	2009-10-21	2016-12-31	2014-04-03	55
-	1959-10-21	-	-	2009-10-21	2016-12-31	2009-10-30	50
-	1959-10-21	-	-	2009-10-21	2016-12-31	2014-10-28	55
-	1959-10-21	-	-	2009-10-21	2016-12-31	2016-05-26	57
-	1959-10-21	-	-	2009-10-21	2016-12-31	2012-09-12	53
-	1959-10-21	-	-	2009-10-21	2016-12-31	2010-07-22	51
2013-10-21	1931-01-02	-	I251	2006-01-01	2006-01-02	2006-06-12	75
-	1931-01-02	-	-	2006-01-01	2006-01-02	2007-04-05	76
-	1931-01-02	-	-	2006-01-01	2006-01-02	2007-11-28	76
2016-04-04	1931-01-02	-	C250	2006-01-01	2006-01-02	2006-03-10	75
-	1931-01-02	-	-	2006-01-01	2006-01-02	2007-10-17	76
-	1931-01-02	-	-	2006-01-01	2006-01-02	2006-09-06	75
-	1931-01-02	-	-	2006-01-01	2006-01-02	2008-06-11	77
2014-07-25	1931-01-02	-	I259	2006-01-01	2006-01-02	2006-09-05	75
2010-10-17	1931-01-02	-	C20	2006-01-01	2006-01-02	2006-06-05	75
-	1931-01-02	-	-	2006-01-01	2006-01-02	2008-05-20	77
-	1931-01-02	-	-	2006-01-01	2006-01-02	2007-03-02	76
2011-07-28	1931-01-02	-	K550	2006-01-01	2006-01-02	2009-10-01	78
2013-05-22	1931-01-02	-	I635	2006-01-01	2006-01-02	2006-06-23	75
2015-05-25	1931-01-02	-	I210	2006-01-01	2006-01-02	2006-10-20	75
2016-12-31	1931-01-02	-	I259	2006-01-01	2006-01-02	2008-02-12	77
2014-07-21	1931-01-02	-	C187	2006-01-01	2006-01-02	2006-07-14	75
2014-01-16	1931-01-02	-	C61	2006-01-01	2006-01-02	2006-01-22	75
2016-04-28	1931-01-02	-	I251	2006-01-01	2006-01-02	2006-04-21	75
2013-02-14	1931-01-02	-	J449	2006-01-01	2006-01-02	2008-05-26	77
-	1931-01-02	-	-	2006-01-01	2006-01-02	2006-06-14	75
2008-08-17	1931-01-02	-	-	2006-01-01	2006-01-02	2006-03-20	75
2013-05-25	1931-01-02	-	C61	2006-01-01	2006-01-02	2007-09-21	76
2010-07-21	1931-01-02	-	I250	2006-01-01	2006-01-02	2006-06-23	75
2010-08-09	1931-01-02	-	C169	2006-01-01	2006-01-02	2009-09-28	78
-	1931-01-02	-	-	2006-01-01	2006-01-02	2006-12-05	75
-	1931-01-02	-	-	2006-01-01	2006-01-02	2006-12-20	75
2011-06-09	1931-01-02	-	Y34	2006-01-01	2006-01-02	2006-03-24	75
2017-02-03	1931-01-02	-	K801	2006-01-01	2006-01-02	2006-12-07	75
2013-04-27	1931-01-02	-	I693	2006-01-01	2006-01-02	2006-03-24	75
-	1931-01-02	-	-	2006-01-01	2006-01-02	2008-07-03	77
-	1931-01-03	-	-	2006-01-01	2006-01-03	2006-10-26	75

2009-02-23	1931-01-03	-	-	2006-01-01	2006-01-03	2006-06-09	75
2011-03-20	1931-01-03	-	I693	2006-01-01	2006-01-03	2006-03-09	75
2009-12-20	1931-01-03	-	-	2006-01-01	2006-01-03	2006-10-24	75
2016-06-27	1931-01-03	-	I255	2006-01-01	2006-01-03	2008-11-14	77
2015-02-01	1931-01-03	-	J449	2006-01-01	2006-01-03	2006-07-26	75
2017-09-01	1931-01-03	-	I258	2006-01-01	2006-01-03	2006-03-29	75
2008-02-09	1931-01-03	-	-	2006-01-01	2006-01-03	2006-04-19	75
2014-12-26	1931-01-03	-	I251	2006-01-01	2006-01-03	2008-11-05	77
2008-09-08	1931-01-03	-	I633	2006-01-01	2006-01-03	2006-10-30	75
-	1931-01-03	-	-	2006-01-01	2006-01-03	2006-02-14	75
-	1931-01-03	-	-	2006-01-01	2006-01-03	2007-11-29	76
2009-11-14	1931-01-03	-	C61	2006-01-01	2006-01-03	2006-09-08	75
2012-01-13	1931-01-04	-	I251	2006-01-01	2006-01-04	2008-05-14	77
-	1931-01-04	-	-	2006-01-01	2006-01-04	2006-08-22	75
-	1931-01-04	-	-	2006-01-01	2006-01-04	2006-07-27	75
-	1931-01-04	-	-	2006-01-01	2006-01-04	2006-06-14	75
2015-09-10	1931-01-04	-	A415	2006-01-01	2006-01-04	2006-10-23	75
2012-08-23	1931-01-04	-	I635	2006-01-01	2006-01-04	2007-04-20	76
-	1931-01-04	-	-	2006-01-01	2006-01-04	2006-07-03	75
-	1931-01-04	-	-	2006-01-01	2006-01-04	2007-01-19	76
2009-08-25	1931-01-05	-	I251	2006-01-01	2006-01-05	2006-02-15	75
2016-06-27	1931-01-05	-	I258	2006-01-01	2006-01-05	2006-09-28	75
2015-08-09	1931-01-05	-	C61	2006-01-01	2006-01-05	2006-03-29	75
2011-07-06	1931-01-05	-	K861	2006-01-01	2006-01-05	2006-05-23	75
-	1931-01-05	-	-	2006-01-01	2006-01-05	2006-04-21	75
-	1931-01-05	-	-	2006-01-01	2006-01-05	2006-05-16	75
-	1931-01-05	-	-	2006-01-01	2006-01-05	2006-07-07	75
2014-01-19	1931-01-05	-	I251	2006-01-01	2006-01-05	2008-02-27	77
2014-09-04	1931-01-06	-	D591	2006-01-01	2006-01-06	2006-12-08	75
-	1931-01-06	-	-	2006-01-01	2006-01-06	2006-09-12	75
-	1931-01-06	-	-	2006-01-01	2006-01-06	2006-06-27	75
2017-09-23	1931-01-06	-	C61	2006-01-01	2006-01-06	2006-06-27	75
-	1931-01-06	-	-	2006-01-01	2006-01-06	2009-09-28	78
-	1931-01-06	-	-	2006-01-01	2006-01-06	2006-11-10	75
2011-05-08	1931-01-06	-	C438	2006-01-01	2006-01-06	2006-03-02	75
2014-08-08	1931-01-06	-	I251	2006-01-01	2006-01-06	2006-02-22	75

Table 6: Small sample of empirical data table for prostate cancer screening program in Lithuania

11.3 List of distribution functions

Log-logistic distribution

$F(x) = \frac{x^\beta}{\alpha^\beta + x^\beta},\ x \geqslant 0,\ \alpha > 0,\ \beta > 0.$

Weibull distribution

$F(x) = 1 - e^{-\frac{a}{\alpha+1}x^{\alpha+1}},\ x \geqslant 0,\ a > 0,\ \alpha > 0.$

Peter and Paul distribution

$F(x) = 1 - (b^{-a})^{\lfloor log_b x \rfloor},\ x \geqslant 1,\ b > 1,\ a > 0.$

Burr distribution

$F(x) = 1 - (1 + x^c)^{-k},\ x \geqslant 0,\ c > 0,\ k > 0.$

Pareto distribution

$F(x) = 1 - (\frac{x_m}{1+x})^\alpha,\ x \geqslant x_m - 1,\ x_m > 0,\ \alpha > 0.$

References

[1] Juckett, D.A.; Rosenberg, B. Comparison of the Gompertz and Weibull functions as descriptions for human mortality distributions and their intersections. *Mech. Ageing Dev.* **1993**, *69*, 1-31.

[2] Missov, T.I.; Lenart, A.; Nemeth, L.; Canudas-Romo, V.; Vaupel, J.W. The Gompertz force of mortality in terms of the modal age at death. *Demog. Res.* **2015**, *32*, 1031-1048.

[3] Vaupel, J.W. Inherited frailty and longevity. *Demography* **1988**, *25*, 277-287.

[4] Lai, C.D. Constructions and applications of lifetime distributions. *Appl. Stoch. Models Bus. Ind.* **2013**, *29*, 127-140.

[5] Wienke, A. Frailty models. *MPIDR Working Paper WP 2003-32* **2003**.

[6] Manton, K.G.; Stallard, E.; Vaupel, J.W. Alternative Models for heterogeneity of mortality risks among the aged. *J. Am. Stat. Assoc.* **1986**, 635-644.

[7] Manton, K.G. Changing concepts of morbidity and mortality in the elderly population. *Milbank Mem. Fund Q. Health Soc.* **1982**, *60*, 183-244.

[8] Yashin, A.I.; Iachine, I.A.; Begun A.Z.; Vaupel, J.W. Hidden frailty: myths and reality. *Document de Travail* **2001**, *34*, 1-48.

[9] Butt, Z.; Haberman, S. Application of frailty-based mortality models using generalized linear models. *ASTIN Bull.* **2004**, 34, 175-197.

[10] Moger, T.A.; Alen, O.O. Hierarchical Levy frailty models and an frailty analysis of data on infant mortality in Norwegian siblings. *UW Biostatistics Working Paper Series* **2006**, Paper 290.

[11] Hougaard, P. Frailty models for survival data. *Lifetime Data Anal.* **1995**, *1*, 255-273.

[12] Finkelstein, M.S. Lifesaving explains mortality decline with time. *Math. Biosci.* **2005**, *196*, 187-197.

[13] Pitacco, E. From Halley to "frailty": a review of survival models for actuarial calculations. *Giornale dell'Istituto Italiano degli Attuari* **2005**, 1-27.

[14] Pitacco, E. High age mortality and frailty. Some remarks and hints for actuarial modelling. *CEPAR* **2016**, Working Paper 2016/19.

[15] Spreeuw, J.; Nielsen, J.P.; Jarner, S.F. A nonparametric visual test of mixed hazard models. *SORT* **2013**, *1*, 153-174.

[16] Assabil, S.E. Forecasting maternal mortality with modified Gompertz model, *J. Adv. Math. Comput. Sci.* **2019**, *32*, 1-7.

[17] Pitacco, E.; Denuit, M.; Haberman, S.; Olivieri, A. *Modelling Longevity Dynamics for Pensions and Annuity Business*; Oxford University Press: Oxford, 2009.

[18] Gavrilov, L.A.; Gavrilova, N.S. The reliability theory of aging and longevity, *J. Theor. Biol.* **2001**, *213*, 527-545.

[19] Henshaw, K.; Constantinescu, C.; Pamen, O.M. Stochastic mortality modelling for dependent coupled lives, *Risks* **2020**, *8*, 17.

[20] Royden, H.L. *Real Analysis*, Macmillan, 1969.

[21] Vaupel, J.M.; Manton, K.G.; Stallard, E. The impact of heterogeneity in individual frailty on the dynamics of mortality, *Demography* **1979**, *6*, 439 – 454.

[22] Fulla, S.; Laurent, J.P. Mortality fluctuations modelling with a shared frailty approach, *Life & Pensions* 2008, *October*, 39–44.

[23] Tuljapurkar, S.; Edwards, R.D. Variance in death and its implications for modelling and forecasting mortality, *Demogr. Res.* **2011**, *24*, 497-526.

[24] Danilenko, S.; Šiaulys, J.; Stepanauskas, G. Closure properties of O-exponential distributions, *Stat. Probab. Lett.* **2018**, *140*, 63-70.

[25] Ragulina, O.; Šiaulys, J. Randomly stopped minima and maxima with exponential-type distributions, *Nonlinear Anal. Model. Control* **2019**, *24*, 297-313.

[26] Sprindys, J.; Šiaulys, J. Regularly distributed randomly stopped sum, minimum, and maximum, *Nonlinear Anal. Model. Control* **2020**, *25*, 509-522.

[27] L. L. Campbell, A. L. McKellips and P. H. Wittke. Distributions and Expectations of Singular Random Variables, *J. Appl. Prob.* **32**, 1032-1040 (1995).

[28] Alamer, A.A; Almulhim,A.S; Alrashed, A.A; Abraham, I. Mortality, severity, and hospital admission among COVID-19 patients with ACEI/ARB use: A meta-analysis stratifying countries based on response to the first wave of the pandemic, *Healthcare* **2021**, *9*, 127.

[29] Chen, J.-J.; Kuo, G.; Lee, T.H.; Yang, H.-Y.;Wu, H.H.; Tu, K.-H.; Tian, Y.-C. Incidence of mortality, acute kidney injury and graft loss in adult kidney transplant recipients with coronavirus disease 2019: Systematic review and meta-analysis, *J. Clin. Med.* **2021**, *10*, 5162.

[30] Rivera-Izquierdo, M.; Pérez de Rojas, J.; Martínez-Ruiz, V.; Pérez-Gómez, B.; Sánchez, M.-J.; Khan, K.S.; Jiménez-Moleón, J.J. Obesity as a risk factor for prostate cancer mortality: A systematic review and dose-response meta-analysis of 280,199 patients, *Cancers* **2021**, *13*, 4169.

[31] Turner, L.; Burbanks, A; Cerasuolo, M. Mathematical insights into neuroendocrine trasndifferentiation of human prostate cancer cells. *Nonlinear Anal. Model. Control* **2021**, *5*, 884-913.

[32] Boucher, J.-P.; Turcotte, R. A longitudinal analysis of the impact of distance driven on the probability of car accidents, *Risks* **2020**, *8*, 91.

[33] Hostiuc, S.; Diaconescu, I.; Rusu, M.C.; Negoi, I. Age estimation using the cameriere methods of open apices: A meta-analysis, *Healthcare* **2021**, **9**, 237.

[34] Li, S.; Chen, J.; Chen, D. PQMLE of a partially linear, varying coefficient spatial autoregressive panel model with random effects, *Symmetry* **2021**, *13*, 2057.

[35] Koroleva, E.; Jigeer, S.; Miao, A.; i Skhvediani, A. Determinants affecting profitability of state-owned commercial banks: Case study of China, *Risks*, **2021**, *9*, 150.

[36] Młynarczyk, D.; Armero, C.; Gómez-Rubio, V.; Puig, P. Bayesian analysis of population health data, *Mathematics* **2021**, *9*, 577.

[37] Zimon, G.; Appolloni, A.; Tarighi, H.; Shahmohammadi, S.; Daneshpou, E. Earnings management, related party transactions and corporate performance: The moderating role of internal control, *Risks* **2021**, em 9, 146.

[38] Huang, Y.; Lu, Z.; Dai,W.; Zhang,W.;Wang, B. Remaining useful life prediction of cutting tools using an inverse Gaussian process model. *Appl. Sci.* **2021**, *11*, 5011.

[39] Sazonov, I.; Grebennikov, D.; Meyerhans, A; Bocharov, G. Markov chain-based stochastic modelling of HIV-1 life cycle in a CD4 T cell, *Mathematics* **2021**, *9*, 2025.

[40] Klakattawi, H.S. The Weibull–Gamma distribution: Properties and applications, *Entropy* **2019**, *21*, 438.

[41] Missov, T.I.; Lenart, A. Gompertz-Makeham life expectancies: Expressions and applications, *Theor. Pop. Biol.* **2013**, *90*, 29-35.

[42] Burger, O.; Missov, T.I. Evolutionary theory of ageing and the problem of correlated Gompertz parameters, *J. Theor. Biol.* **2016**, *408*, 34-41.

[43] Dotlaličova, P. Comparison of selected mortality models, *The 11-th International Days of Statistics and Economics, Prague, September 14-16, 2017*.

[44] Jarner, S.F.; Kryger, E.M. Modelling adult mortality in small populations: The SAINT model, *ASTIN Bull.* **2011**, *41*,377-418.

[45] Saika, P.; Borah, M. A comparative study of parametric models of old-age mortality, *Int. J. Sci. Res.* **2014**, *3* , 406-410.

[46] Beard, R.E. Some aspects of theories of mortality, cause of death analysis, forecasting and stochastic processes. In *Biological Aspects of Demography (ed. W. Brass)*; Taylor and Francis, London, 1971, 57-69.

[47] Pflaumer, P. Life table forecasting with Gompertz distribution. *JSM Proceedings, Social Statistics Section*; VA: American Statistical Association, Alexandria, 2007, 3564-3571.

[48] Missov, T.I. Gamma-Gompertz life expectancy at birth, *Demogr. Res.* **2013**, *28*, 59-270.

[49] Wang, S.S.; Brown, R.L. A frailty model for projection of human mortality improvements, *J. Actuar. Pract.* **1998**, *6*, 1993-2006.

[50] Missov, T.I.; Vaupel, J.W. Mortality implications of mortality plateaus, *SIAM Rev.* **2015**, *57*, 61-70.

[51] Horiuchi, S.; Ouellette, N.; Cheung, S.L.K.; Robine, J.M. Modal age at death: lifespan indicator in the era of longevity extension, *Vienna Yearb. Pop. Res.*, **2013**, *11*, 37–69.

[52] Rau, R.; Ebeling, M.; Peters, F.; Bohk-Ewald, C.; Missov, T.I. Where is the level of mortality plateau? *Living to 100, Society of Actuaries International Symposium*, 2017.

[53] Cohen, J.E.; Bohk-Ewald, C.; Rau, R. Gompertz, Makeham and Siler models explain Taylor's law in human mortality data, *Demog. Res.* **2018**, *38*, 773-842.

[54] Romo, V.C. The modal age of death and the shifting mortality hypothesis, *Demog. Res.* **2008**, *19*, 1179-1204.

[55] Šiaulys, J. and Puišys, R. Survival with random effect, *Mathematics*, **2022**, *10*, 1097.

[56] Puišys, R., Lewkiewicz, S. and Šiaulys, J. Properties of the random effect transformation. *Lith Math J* **(2024)**.

[57] J.M.P. Albin and M. Sunden. On the asymptotic behaviour of Lévy processes, part I: Subexponential and exponential processes. *Stoch. Process. Appl.*, *119*: 281 - 304, **2009**.

[58] Su, C. and Chen, Y. On the behavior of the product of independent random variables, *Sc. China Ser. A-Math.*, **2006**, *49*, 342 - 359.

[59] Cui, Z. and Wang, Y. On the long tail property of product convolution, *Lith. math. J.*, **2020**, *60*, 315 - 329.

[60] Xu, H. and Scheutzow, M. and Wang, Y. On a transformation between distributions obeing the principle of a single big jump, *J. Math. Anal. Appl.*, **2015** *430*, 672 - 684.

[61] Wang, S. and Gao, Y. Precise large deviations for aggregate claims of a compound renewal risk model with arbitrary dependence between claim sizes and waiting times, *Lith. math. J.*, **2022**, *62*, 542 - 552.

[62] Danilenko, S. and Šiaulys, J. Randomly stopped sums of not identically distributed heavy tailed random variables, *Stat. Probab. Lett.*, **2016**, *113*, 84 - 93.

[63] Bingham, N.H. and Goldie, C.M. and Teugels, J.L. Regular Variation, *Encycl. Math. Appl., Vol. 27, Cambridge Univ. Press, Cambridge*, **1987**.

[64] Leipus, R. and Šiaulys, J. and Konstantinides, D. Closure Properties for Heavy-Tailed and Related Distributions, **2023**, *Springer, Cham.*

[65] Foss, S. and Korshunov, D. and Zachary, S. An Introduction to Heavy-Tailed and Subexponential Distributions, *2nd ed., Springer, New York*, **2013**.

[66] Resnick, S.I. Exreme Values, Regular Variation and Point Process, *Springer, New York*, **1987**.

[67] Samorodnitsky, G. Stochastic Processes and Long Range Dependence, *Springer, New York*, **2016**.

[68] Seneta, E. Regularly Varying Functions, *Lecture Notes in Mathematics, Volume 508, Springer, Berlin, Heidelberg*, **1987**.

[69] Karamata, J. Sur un mode de croissance régulière. Théorèmes fundamentoux, *B. Soc. Math. Fr.*, *61*, 55 - 62, **1933**.

[70] Embrechts, P. and Klüppelberg, C. and Mikosch, T. Modelling Extremal Events for Insurance and Finance, *Springer, New York*, **1997**.

[71] Feller, W. An Introduction to Probability Theory and Its Applications, *Vol. II, Wiley, New York*, **1971** .

[72] Feller, W. One-sided analogues of Karamata's regular variation, *Enseign. Math.*, *15*, 107 - 121, **1969**.

[73] Cai, J. and Tang, Q. On max-type equivalence and convolution closure of heavy-tailed distributions and their applications, *J. Appl. Probab.*, *41*, 117 - 130, **2004**.

[74] Dirma, M. and Paukštys, S. and Šiaulys, J. Tails of the moments for sums with dominatedly varying random summands, *Mathematics*, *9*, 824, **2021**.

[75] Leipus, R. and Paukštys, S. and Šiaulys, J. Tails of higher-order moments of sums with heavy-tailed increments and application to the Haezendonck-Goovaerts risk measure, *Stat. Probab. Lett.*, *170*, 108998, **2021**.

[76] Leipus, R. and Šiaulys, J. and Vareikaitė, I. Tails of higher-order moments with dominatedly varying summands, *Lith. Math. J.*, *59*, 389 - 407, **2019**.

[77] Leipus, R. and Šiaulys, J. Closure of some heavy-tailed distribution classes under random convolution, *Lith. Math. J.*, *52*, 249 - 258, **2012**.

[78] Pakes, A.G. Convolution equivalence and infinite divisibility, *J. Appl. Probab.*, *41*, 407 - 424, **2004**.

[79] Shimura, T. and Watanabe, T. Infinite divisibility and generalized subexponentiality, *Bernoulli*, *11*, 445 - 469, **2005**.

[80] Tang, Q. and Yan, J. A sharp inequality for the tail probabilities of sums of i.i.d r.v.'s with dominatedly varying tails, *Sci. China Ser. A-Math.* , *45*,1006 - 1011, **2002**.

[81] Watanabe, T. and Yamamuro, K. Ratio of the tail of an infinitely divisible distribution on the line to that of its Lévy measure, *Electron. J. Probab.*, *15*, 44 - 74, **2010**.

[82] Watanabe, T. and Yamamuro, K. Local subexponentiality and self-decomposability, *J. Theor. Probab.*, *23*, 1039 - 1067, **2010**.

[83] Chistyakov, V.P. A theorem on sums of independent, positive random variables and its applications to branching processes, *Theory Probab. Appl.*, *9*, 640 - 648, **1964**.

[84] Beck, S. and Blath, J. and Scheutzow, M. A new class of large claim size distributions: definition, properties, and ruin theory, *Bernoulli*, *21*, 2475 - 2483, **2015**.

[85] Foss, S. and Korshunov, D. and Zachary, S. Convolution of long-tailed and subexponential distributions, *J. Appl. Probab.*, *46*, 756 - 767, **2009**.

[86] Embrechts, P. and Goldie, C.M. On closure and factorization properties of subexponential and related distributions, *J. Aust. Math. Soc. Ser. A*, *29*, 243 - 256, **2015**.

[87] Konstantinides, D. and Leipus, R. and Šiaulys, J. On the non-closure under convolution for strong subexponential distributions, *Nonlinear Anal. Model. Control*, *28*, 97 - 115, **2023**.

[88] Dirma, M. and Nakliuda, N. and Šiaulys, J. Generalized moments of sums with heavy-tailed random summands, *Lith. Math. J.*, *63*, 81 - 91, **2023**.

[89] Geng, B. and Liu, Z. and Wang, S. A Kesten-type inequality for randomly weighted sums of dependent subexponential random variables with applications to risk theory, *Lith. Math. J.*, *63*, 254 - 271, **2023**.

[90] Weibull, W. (**1939**) A Statistical Theory of the Strength of Materials. *Generalstabens Litografiska Anstalts Förlag, Stockholm*.

[91] Rebello, R.J. and Oing, C. and Knudsen, K.E. and Loeb, S. and Johnson, D.C. and Reiter, R.E. et al. **(February 2021)**. Prostate cancer. *Nature Reviews Disease Primers. 7 (1)*: 9

[92] McIntosh, S.A. and Alam, F. and Adams, L. and Boo, I.S. and Callaghan, J. and Conti, I. et al. **(June 2023)**. Global Funding for Cancer Research Between 2016 and 2020: a Content Analysis of Public and Philanthropic Investments. *Lancet Oncology. 24 (6)*: 636–645

NOTES

NOTES

NOTES

Vilnius University Press
9 Saulėtekio Ave., Building III, LT-10222 Vilnius
Email: info@leidykla.vu.lt, www.leidykla.vu.lt
bookshop.vu.lt, journals.vu.lt
Print Run 20