СОФИЙСКИ УНИВЕРСИТЕТ "СВ. КЛИМЕНТ ОХРИДСКИ"
ФАКУЛТЕТ ПО МАТЕМАТИКА И ИНФОРМАТИКА
КАТЕДРА "ВЕРОЯТНОСТИ, ОПЕРАЦИОННИ ИЗСЛЕДВАНИЯ И СТАТИСТИКА"

Пламен Ивайлов Траянов

ОБЩИ РАЗКЛОНЯВАЩИ СЕ ПРОЦЕСИ НА КРЪМП–МОД–ЯГЕРС – МОДЕЛИ И ПРИЛОЖЕНИЯ В ДЕМОГРАФИЯТА

АВТОРЕФЕРАТ

на дисертация
за придобиване на образователна и научна степен
„Доктор“

Научен ръководител:
проф. д-р Марусия Славчова-Божкова

София, 2016

СОФИЙСКИ УНИВЕРСИТЕТ "СВ. КЛИМЕНТ ОХРИДСКИ"
ФАКУЛТЕТ ПО МАТЕМАТИКА И ИНФОРМАТИКА
КАТЕДРА "ВЕРОЯТНОСТИ, ОПЕРАЦИОННИ ИЗСЛЕДВАНИЯ И СТАТИСТИКА"

Пламен Ивайлов Траянов

ОБЩИ РАЗКЛОНЯВАЩИ СЕ ПРОЦЕСИ НА КРЪМП–МОД–ЯГЕРС – МОДЕЛИ И ПРИЛОЖЕНИЯ В ДЕМОГРАФИЯТА

АВТОРЕФЕРАТ

на дисертация
за придобиване на образователна и научна степен
„Доктор“

Професионално направление: *4.5 Математика*
Научна специалност: *Теория на вероятностите и математическата статистика (01.01.10)*

Научен ръководител:
проф. д-р Марусия Славчова-Божкова

София, 2016

Данни за дисертационния труд:
Обем на дисертацията: 118 стр.
Основен текст: 71 стр.
Литература: 113 заглавия.
Публикации по дисертацията: 5 заглавия.

Дисертационният труд е обсъден и препоръчан за започване на процедура по защита на разширено заседание на катедра "Вероятности, операционни изследвания и статистика"при Факултет по математика и информатика - СУ, проведено на 16 май 2016 г.

Автор: *Пламен Ивайлов Траянов*
Заглавие: *Общи разклоняващи се процеси на Кръмп–Мод–Ягерс – модели и приложения в демографията.*

Мотивация, цели и задачи на дисертацията

Теорията на разклоняващите се процеси започва своето начало през втората половина на 19-ти век със задачата да даде обяснение на изчезването на аристократичните фамилни линии в Европа. През 1845 г. Биенеме създава първият модел на разклоняващ се процес, като години по-късно, през 1874 г., независимо от него Галтон и Уотсън публикуват първата си научна работа върху тези процеси. Терминът „Разклоняващ се процес“ е въведен по-късно от Колмогоров, през 1947 г. (виж А. Колмогоров и Н. Дмитриев [31]). Разклоняващият се модел, известен още като процес на Биенеме–Галтон–Уотсън (БГУ), оттогава се използва за моделиране на системи, чиито компоненти (клетки, частици, индивиди) се възпроизвеждат, трансформират и умират. Интересно е да се отбележи, че развитието на теорията на разклоняващите се процеси в Русия след края на Втората световна война е било засекретено поради приложението на тази теория в разработката на атомната бомба. Така статиите на редица учени, сред които А. Колмогоров, Н. Дмитриев и Б. Севастьянов, биват засекретени (виж Н. Янев [72], А. Колмогоров и Н. Дмитриев [31], А. Колмогоров и Б. Севастьянов [32]). Разклоняващият се процес на БГУ, в който частиците живеят единица време и в края на своя живот се възпроизвеждат, по-късно е обобщен от процеса на Белман–Харис, при който продължителността на живот на частиците се моделира със случайна величина. През 1968 г. – 1969 г. Кръмп, Мод и Ягерс публикуват паралелно модел на разклоняващ се процес (виж [11, 12, 27]), при който продължителността на живота на частиците и моментите на раждане са случайни величини. Този модел става известен в литературата като разклоняващ се процес на Кръмп–Мод–Ягерс или още като Общ Разклоняващ се Процес (ОРП). През 1975 г. се публикува книгата на Ягерс, описваща модела на ОРП (виж [28]). ОРП се прилага за моделиране на растежа на популация от клетки (виж П. Грийн [23]), изследване на M/G/1 системи (виж С. Гришечкин [24]), изследване на генетиката на остаряването (виж П. Олофсон [45]), а в П. Ягерс [28] е представено демографско приложение, като е показано уравнението на Лотка, даващо общия брой раждания до момента, като следствие от теорията на ОРП.

В България изучаването на теорията на разклоняващите се процеси започва с учебниците на Обрешков (виж [1]) и Обретенов (виж [2]), като по-късно изучаването на тези стохастични процеси започва да става по-интензивно с трудовете на Н. Янев, П. Майстер, Б. Димитров и техните ученици. Както е отбелязано в [43], за период от 23 години, от 1985 г. до 2008 г., са публикувани повече от 150 статии, книги, глави от книги, доклади, дисертации и други. Издадена е книга за приложението на разклоняващите се процеси в биологията и медицината (виж А. Яковлев и Н. Янев [68]), учебник за студенти (виж М. Славчова-Божкова и Н. Янев [52]), глави от книги по разклоняващи се процеси (виж К. Митов и Н. Янев [42], П. Майстер [36], К. Митов, Г.

Митов и Н. Янев [39], Н. Янев [71], М. Гонзалес, Р. Мартинес и М. Славчова-Божкова [22]), както и книга по процеси на възстановяване (виж К. Митов и Е. Омей [40]), чиято теория е ключова за модела на ОРП. Първият Световен конгрес по разклоняващи се процеси е организиран във Варна през 1993 г. от българските учени с председател на организационния комитет Н. Янев, като резултатите са публикувани в [19]. Методи за оценка на параметрите и симулации на разклоняващ се процес на БГУ са изследвани от В. Стоименова [53] и В. Стоименова, Д. Атанасов и Н. Янев [54, 55, 56]. Методи за оценка на разпределението на поколението на един индивид, когато разполагаме с информация за част от реализираното дърво, са представени в Н. Даскалова [14, 15, 16].

Съвременната теория на разклоняващите се процеси има широко приложение в много различни по характера си дисциплини. Така например те ни дават възможност да изследваме развитието на епидемии в населението и контролирането им чрез ваксинация (виж [22, 51]). В медицината те също предоставят модел на развитието на ракови клетки в организма (виж [18]). В ядрената физика те предоставят модел за работата на ядрен реактор (виж [7]), в който контролираме критичността на процеса. В биологията разклоняващите се процеси могат да моделират делене на клетки, възпроизвеждане на вируси (виж [22, 30, 70]). В сферата на финансите могат да се използват за определяне на цената на опции (виж [37, 38]). Разклоняващите се процеси също биха могли да отразяват и влиянието на миграционните процеси върху числеността на популацията (виж [4, 50, 69]).

Този дисертационен труд е съсредоточен върху приложението на ОРП в демографията. Съществуват известни теоретични ограничения на класическия ОРП, които трябва да бъдат преодолени, за да се моделира население. От една страна ОРП е изключително гъвкав модел, в непрекъснато време, в който всеки индивид има случайна продължителност на живота, всяка жена ражда случаен брой деца през живота си и през случайни интервали от време. От друга страна моделът разглежда единствено процес, започващ от един единствен индивид на възраст 0 в момента 0, както и приема, че стохастичните закони на смъртността и раждаемостта не се променят във времето. Тези две ограничения на модела се преодоляват в публикациите [59, 60, 61], за да се моделира население, което се състои от множество индивиди на различни възрасти в нулевия момент от време и в което законите на смъртността и раждаемостта се променят с течение на времето.

Възрастовата структура на населението в ОРП се явява решение на уравнения на възстановяване. Използването на числен метод за решението на тези уравнения е единствен възможен подход, предвид това, че не разполагаме с аналитичен вид на разпределението на смъртността и раждаемостта. Такива числени методи са дискутирани в Д. Бартоломю [5], С. Фром [21], М. Кси [67], К. Митов и Е. Омей [41]. В публикация [60] също е предложен изчислително ефективен числен метод за решаване на тези уравнения на възстановяване. Оказва се, че демографският метод за проекция на населението с матрица на Лесли е всъщност частен случай на този числен метод. В същата публикация е дискутирано и численото пресмятане на очакваната възрастова структура на ОРП, започващ от множество индивиди на случайни възрасти в началния момент.

Бъдещата възрастова структура на населението зависи едновременно от стохастиката на самия разклоняващ се процес и от променливите стохастични закони на раждаемостта и смъртността. Взаимодействието на тези два различни източника на несигурност за прогнозата определя доверителните интервали за бъдещото население. Оказва се, че основен дял от риска на прогнозата за общия брой на населението носи стохастиката на законите на смъртността и раждаемостта, докато стохастиката на ОРП представлява малък дял от общия риск. Когато обаче се интересуваме от прогнози на малки групи хора, като тези на определена възраст или живеещите в малки населени места, тогава стохастиката на ОРП става по-съществена за прогнозата на населението. Доверителните интервали ни позволяват да придобием по-ясна представа за тенденциите на процесите, случващи се в населението, както и да направим качествени изводи за неговото бъдеще.

Демографските и статистическите методи предоставят необходимите параметри, участващи в модела на ОРП и поради това представляват важна част от моделирането на населението. Чрез тях се моделира разпределението на продължителността на живота и вероятностите за раждане по възрасти, които участват и са съществена част от модела на ОРП. Богатството от съществуващи и утвърдени методологии предоставя на изследователя възможност да избере най-подходящия инструмент за решаване на специфичните проблеми, които възникват при моделиране на параметрите на ОРП. Настоящият дисертационен труд се стреми да опише в детайли процеса на моделиране и прогнозиране на населението чрез ОРП - от демографските данни, с които разполагаме, до крайния модел на ОРП, както и се стреми да представи възможно най-подходящите и съвременни демографски методи, възникнали през последните години, и предимствата им спрямо класическите демографски подходи.

Една от класическите книги, описваща демографската теория и широко използвана и днес, е тази на Н. Кийфиц и Х. Касуел (виж [29]). Тя описва класическата математическата теория за пресмятането на таблиците за смъртност, част от които са и коефициенти за смъртност и раждаемост по възрасти и пол. В процеса на изследване се установи, че съществуват редица подобрения на класическата методология, представени в методологията на Уилмот, Андреев, Жданов и Глей (виж HMD [65]), както и Школников (виж [49]), които внасят допълнително подобрение в емпиричната оценка на демографските коефициенти за смъртност. Връзката между коефициентите за смъртност/раждаемост и вероятностите за смъртност/раждаемост е изведена от Чианг (виж [9, 10]). Използвайки формулата на Чианг можем да пресметнем емпиричните вероятности за смъртност и раждаемост по възрасти (виж параграф 1.3), като от тях се забелязва, че съществува функционална зависимост от възрастта на индивида.

Един от класическите демографски модели на смъртността е функцията на Гомперц–Макеам (виж [44]), която се построява върху логаритъма от емпиричните вероятности за смъртност по възрасти. По-късно Мод предлага обобщение на съществуващите дотогава аналитични моделни функции за възрастовата структура на вероятностите за смъртност (виж Ч. Мод [44]). Един от проблемите, забелязани от Мод, е че емпиричните вероятности за смъртност не се описват достатъчно добре от съществуващите аналитични моделни функции (на места се наблюдават твърде големи средноквадратични грешки) и това налага да се използват „по-гъвкави“ функции при моделирането на смъртността. Идеята на Мод е, че комбинирайки различни добре известни аналитични функции (като тази на Гомперц, Макеам, Вайбул и др.), всяка от които е пригодена да описва добре специфична част от възрастовата структура на смъртността, може да се постигне по-добра моделна функция с по-малка средноквадратична грешка. Предложеният от него модел има общо 11 параметъра, което осигурява достатъчна гъвкавост. В същото време обаче това се оказва и недостатък на модела, тъй като на практика итеративната процедура на нелинейните най-малки квадрати не успява да намери устойчиви оценки за всички тези параметри. Дори редукцията на броя на параметрите до 7, предложена от Мод, не успява да се справи напълно с този практически проблем.

Един съвременен подход, този на изглаждащите сплайни, се оказва по-подходящ за моделиране на разпределението на продължителността на живота. Изглаждащите сплайни, като част от анализа на функционални данни, са представени в книгата на Д. Рамзи [48] и навлизат в практиката основно след труда на Ч. де Бор [17], който предлага ефективен метод за тяхното оценяване, като линейна комбинация на базисни сплайни. Средноквадратичната грешка на моделната сплайн функция е значително по-малка, както и не се наблюдава закономерност в остатъците, за разлика от модела на Мод. Често изтъкван недостатък на изглаждащите сплайни е „липсата на аналитичен вид“. Този недостатък обаче лесно се преодолява чрез анализ на главните компоненти (виж Д. Рамзи и Б. Силвърман [48]), който представя моделните сплайн функции като линейна комбинация от няколко (обикновено 3-4) базисни функции, а коефициентите пред тях са независими помежду си, за разлика от параметрите в модела на Мод. Това подобрява прогнозируемостта на функциите, от една страна поради по-стабилната оценка на коефициентите пред главните направления, а от друга поради независимостта им. Този подход за прогнозиране на функции в демографията е представен от Р. Хиндман (виж [66]). Прогнозирането на главните компоненти може да се осъществи например с ARIMA модели, за които съществува богата литература - Г. Бокс, Г. Дженкинс и Г. Райнсел [8], Д. Крайер и К. Чан [13], Т. Хасти [25], Д. Ало и Б. Спенсър [3] и други.

В процеса на решаване на практически задачи често възникват проблеми с липсващи данни. Често за някои години и за някои възрасти липсват данни за смъртността, за раждаемостта и за броя на населението над определена възраст, което налага използването на различни похвати за попълване на липсващи данни, там където това е възможно. Една от утвърдилите се съвременни методологии за моделиране на липсващи данни за смъртността е предложена от В. Канисто (виж [58]), като методът е известен като „Модел на Канисто за смъртност на високите възрасти". Един от стандартните методи за попълване на липсващи данни за броя на населението е така наречения „Survivor Ratio" метод (виж [57, 65]).

Целта на дисертационния труд е да представи необходимите нови резултати в теорията на разклоняващите се процеси, които преодоляват някои от съществуващите ограничения на класическия ОРП, като позволяват чрез него да се моделира население. Една от задачите, които е необходимо да се решат, е да се моделира население, състоящо се от множество индивиди на различни възрасти в нулевия момент от време и в което законите на смъртността и раждаемостта се променят с течение на времето. Теорията на разклоняващите се процеси представя възрастовата структура на населението като решение на уравнения на възстановяване. Използването на числен метод за решението на тези уравнения в случая е единствен възможен подход, предвид това, че не разполагаме с аналитичен вид на разпределението на смъртността и раждаемостта. Поради това възниква необходимост да се намери изчислително ефективен числен метод, който намира решенията на необходимите уравнения на възстановяване и очакваната възрастова структура на ОРП, започващ от множество индивиди на случайни възрасти в началния момент. Оказва се, че демографският метод за проекция на населението с матрица на Лесли всъщност представлява частен случай на разработения числен метод. Прилагайки теорията на ОРП в непрекъснато време можем да получим оценка за грешката на този демографски подход, която произтича от дискретизация на времето. Бъдещата възрастова структура на населението зависи едновременно от стохастиката на самия разклоняващ се процес и от променливите стохастични закони на раждаемостта и смъртността. Така естествено възниква задачата да се изследва взаимодействието на тези два различни източника на несигурност за прогнозата, техният принос към общия риск, както и да се определят доверителните интервали за бъдещото население. Дисертационният труд също се стреми да опише в детайли процеса на моделиране и прогнозиране на населението - от демографските данни, с които разполагаме, до крайния модел на ОРП, както и се стреми да представи възможно най-подходящите и съвременни демографски методи, възникнали през последните години, и предимствата им спрямо класическите демографски подходи.

Структура на дисертацията

• В Глава 1 е разгледана съществуващата съвременна методология за моделиране на законите на смъртността и раждаемостта. Представени са избрани демографски методи за оценка на вероятностите за смъртност и раждаемост по възрасти, които са необходими за модела на ОРП.

Практическата приложимост на модела на населението зависи в голяма степен от това колко добре моделираме стохастичните закони на смъртността и раждаемостта. Колкото по-точно те отразяват реалните данни с които разполагаме, толкова по-релевантен ще бъде и моделът на ОРП, в който те участват. Съществуващата демографска теория се оказва изключително полезна в това отношение, предлагайки методология базирана на реални данни за населението. В тази глава са представени известните съвременни демографски техники за пресмятане на разпределението на продължителността на живота и вероятностите за раждаемост по възрасти. Разгледани са предимствата на някои от съвременните демографски похвати спрямо предходни такива. Въведени са понятията кохортна и периодна вероятност за доживяване, разгледана е формулата на Чианг, която показва връзката между демографските коефициенти за смъртност и вероятностите за смъртност по възрасти. Разгледани са емпиричните оценки на периодните коефициенти и вероятности за смъртност по възрасти. Разгледан е модела на Мод за периодната функция на доживяване, както и алтернатива на този модел, базирана на анализ на функционални данни. Разгледани са съвременни методологии за моделиране на липсващи данни за смъртността и броя на населението и моделирането на вероятностите за раждаемост по възрасти.

• В Глава 2 са представени някои от известните твърдения от теория на възстановяването и теория на ОРП. Представени са и доказателствата на новите резултати, които позволяват моделирането на населението от реални данни и същевременно представят някои демографски методи за прогнозиране на населението като естествен резултат в теорията на ОРП.

В параграф 2.1 са дискутирани необходимите твърдения от теория на възстановяването, която се използва за намиране на очакваното бъдещо население. Теорията на възстановяването, разгледана в книгата на К. Митов и Е. Омей [41], се оказва важен елемент при моделирането на ОРП. Уравнението, описващо очакваното бъдещо население, представлява уравнение на възстановяване, чието решение и асимптотични свойства се описват от теорията на възстановяването (виж О. Хайриен и Н. Янев [26]). Същевременно миграционните процеси в населението биха могли да се моделират чрез процес на възстановяване, в който през случайни моменти от време пристигат/напускат групи хора със случаен размер.

Определение 2.1.1. Нека разгледаме процес на възстановяване $N(t)$ с времена между възстановяванията с разпределение $F(t)$ и нека $F^{n*}(t)$ означава n-тата конволюция на функцията $F(t)$. Функция на възстановяване се дефинира като

$$U(t) = \mathbb{E}(N(t)) + 1 = \sum_{n=0}^{\infty} F^{n*}(t)$$

и представлява средния брой възстановявания до момента $t > 0$, като броим момента $t = 0$ за момент на възстановяване.

От определението следва, че $U(t)$ е растяща функция, а може да се покаже, че също представлява Радонова мярка (крайна върху ограничените множества). Поради това, че $F(t)$ и $U(t)$ представляват мерки над реалната права, съсредоточени над $\mathbb{R}_+$, можем да представим $U((a, b]) = U(b) - U(a)$ като мярка на интервала $(a, b]$ и тогава $U((0, t]) = U(t)$.

Теорема 2.1.1 (Елементарна теорема на възстановяването, виж К. Митов и Е. Омей [40])**.** Ако $F(0) < 1$ и $F(\infty) = 1$, то

$$\lim_{t\to\infty} \frac{U(t)}{t} = \begin{cases} \mu^{-1} & \text{, ако } \mu < \infty \\ 0 & \text{, ако } \mu = \infty. \end{cases}$$

Определение 2.1.2. Уравнение на възстановяване наричаме интегрално уравнение от вида

$$Z(t) = z(t) + \int_0^t Z(t-u)\,\mathrm{d}F(u), \quad t \geq 0, \tag{2.1.1}$$

където $z : \mathbb{R} \to \mathbb{R}$, $z(t) = 0$ за $t < 0$ и $F(t)$ е функция на разпределение на $[0, \infty]$. Допуска се и несобствена функция на разпределение F.

Следната теорема ни дава решението на уравнение (2.1.1):

Теорема 2.1.2. Нека $z(t)$ е ограничена върху всеки краен интервал и $z(t) = 0$ за $t < 0$. Тогава уравнението на възстановяване (2.1.1) има единствено решение в класа на функциите ограничени върху крайните интервали и равни на 0 в $(-\infty, 0)$, което се задава с

$$Z(t) = U * z(t) = \int_0^t z(t-u)\,\mathrm{d}U(u), \quad t \geq 0,$$

където U е функцията на възстановяване, дефинирана с Определение 2.1.1.

Граничното поведение на решението на уравнението на възстановяване се дава от възловата теорема на възстановяването:

Теорема 2.1.3. Нека $z(t)$ е непосредствено интегруема по Риман и $z(t) = 0$ за $t < 0$. Нека $F(t)$ е собствена функция на разпределение, несъсредоточена в нулата и с крайно математическо очакване. Тогава, ако $F(t)$ е неаритметична, то

$$Z(t) = U * z(t) \to \frac{1}{\mu} \int_0^\infty z(u)\,\mathrm{d}u, \quad t \to \infty.$$

За аритметична функция $F(t)$ е изпълнено съответно твърдение.

В параграф 2.2 е описана необходимата теория (виж П. Ягерс [28]), която се използва за моделиране на раждаемостта.

Определение 2.2.1. Точков процес ξ върху $\mathbb{R}_+$ наричаме изображение от едно вероятностно пространство $(\Omega, \mathcal{F}, \mathbb{P})$ към целочислените бореловите мерки над $\mathbb{R}_+$, такива, че за всяко ограничено множество B в Бореловата σ-алгебра мярката $\xi(B)$ представлява целочислена крайна случайна величина. Мярката на неограничените множества се допуска да бъде и безкрайност.

Точковият процес представлява всъщност случайна мярка. Интересно е да се отбележи, че процесът на възстановяване е точков процес, понеже дефинира целочислена случайна мярка на всеки интервал като броя на възстановяванията, които са се случили в този интервал. В случая ξ описва броя на ражданията на една жена, като $\xi((a,b])$ е случайна величина, означаваща броя на ражданията на жената във възрастовия интервал $(a,b]$. Тъй като $\xi(-\infty, 0) = 0$, можем да означим $\xi(t) = \xi([0,t])$.

Определение 2.2.2. Функция на възпроизводството (reproduction function) наричаме

$$\mu(t) = \mathbb{E}[\xi(t)].$$

Определение 2.2.3. Вероятностна пораждаща функция на възпроизводството (reproduction generating function) наричаме функцията

$$f(s) = \mathbb{E}[s^{\xi(\infty)}].$$

В параграф 2.3 е изложена дефиницията и основните теореми от теорията на ОРП, които са представени в П. Ягерс [28]. Нека разгледаме процес, породен от един индивид - прародител. Означаваме този индивид с (0). С индекси $(1), (2), \ldots$ означаваме всички възможни индивиди, които биха могли да бъдат родени от прародителя (0). Означаваме с $(i,1), (i,2), \ldots$ множеството от всички възможни индивиди, които биха могли да бъдат родени от (i), където i е естествено число. Или най-общо казано, ако $x = (j_1, \ldots, j_n)$, то с $(x, j_{n+1}) = (j_1, \ldots, j_n, j_{n+1})$ означаваме j_{n+1}-то дете на майката x. Не всички от тези n-орки отговарят на реално роден индивид. Те включват и онези индивиди, които никога не са били раждани.

Нека I е индексното множество от всички възможни индивиди в популацията, т.е. I е биективно на множеството на всички наредени n-орки, $n = 1, 2, \ldots$. За всеки индивид $x \in I$ нека λ_x да бъде случайна величина, описваща продължителността на живота му, а ξ_x да бъде точков процес, описващ възпроизводството на x. Нека означим с $N(\mathbb{R}_+)$ множеството от целочислените мерки над $\mathbb{R}_+$. Едно от важните предположения в класическата дефиниция на ОРП е, че двойките (λ_x, ξ_x) са независими, еднакво разпределени (λ_x и ξ_x обаче могат да бъдат зависими помежду си). Нека с Q означим съвместното им разпределение. Тогава Q е мярка в $\mathbb{R}_+ \times N(\mathbb{R}_+)$. Означаваме с $L(u) = \mathbb{P}(\lambda_x \leq u) = Q(u, \infty)$ маргиналното разпределение на λ_x, а $S(u) = 1 - L(u)$ е вероятността за доживяване до възраст u. Ще предполагаме също, че $\mathbb{P}(\xi_x(\lambda_x, \infty) = 0) = 1$, т.е. жената ражда само докато е жива.

Означаваме с $\tau_x(k) = \inf\{t : \xi_x(t) \geq k\}$ възрастта, на която майката ражда k-тото си дете. Допуска се и $\tau_x(k) = \infty$, ако жената е родила по-малко от k деца през живота си. Тогава

$$\sigma_x = \tau_0(j_1) + \tau_{j_1}(j_2) + \ldots + \tau_{(j_1, \ldots, j_{n-1})}(j_n)$$

е рождената дата на индивида $x = (j_1, \ldots, j_n)$ спрямо летоброене, започнало с раждането на (0), т.е. $\sigma_0 = 0$.

Определение 2.3.1. Казваме, че индивидът (x, k) е реализиран, ако x е реализиран и $\xi_x(\infty) \geq k$.

Макар, че I има мощност на континуум, само част от индивидите са реализирани. В демографския случай, ще се окаже, че само краен брой индивиди се реализират за крайно време.

Индивидът x е жив в момента t, ако $t - \sigma_x \geq 0$ и $t - \sigma_x \leq \lambda_x$, или по-накратко $\sigma_x \leq t < \sigma_x + \lambda_x$.

Определение 2.3.2. Дефинираме индикаторна случайна величина, която означава дали индивидът x е жив и на възраст по-малка от a в момента t, по следния начин:

$$z_t^a(x) = \begin{cases} 1, & \text{ако } t - a < \sigma_x \leq t < \sigma_x + \lambda_x, \\ 0, & \text{иначе.} \end{cases}$$

Определение 2.3.3. Дефинираме броя на живите индивиди в момента t, на възраст по-малка от a, като случайна величина

$$z_t^a = \sum_{x \in I} z_t^a(x).$$

Определение 2.3.4. Стохастичният процес $z_t^a, t \geq 0, a > 0$ се нарича общ разклоняващ се процес (ОРП).

Означаваме $z_t = z_t^a$, за $a > t$. Случайната величина z_t представлява броя на всички индивиди в популацията и е дефинирана коректно, тъй като максималната възможна възраст на индивид в момента t е самото t.

Определение 2.3.5. Нека да означим с $I_n = \{(n, x); x \in I\}$ поколението с родител (n). Тогава броят на индивидите в популацията, произлезли от (n), е

$$z_t^{[n]a} = \sum_{x \in I_n} z_t^a(x).$$

Следователно z_t^a може да се представи във вида

$$z_t^a = z_t^a(0) + \sum_{n=1}^{\xi_0(t)} z_t^{[n]a}.$$

Процесът започнал с раждането на (n) е нов разклоняващ се процес, вероятностно копие на оригиналния процес. Случайните величини $z_t^{[n]a}$ са независими, еднакво разпределени, със същото разпределение като z_t^a.

В сила е следната теорема, която гласи, че при определени условия общия брой на индивидите в популацията е краен във всеки един момент от време.

Теорема 2.3.1. Ако $\mu(0) < 1$ и $\mu(t) < \infty$ за някое t, то $\mathbb{P}(z_t < \infty) = 1$ за всяко t.

Една от най-важните теореми, представена в книгата на Ягерс (виж [28]), гласи, че очакваният бъдещ размер на популацията се получава като решение на уравнение на възстановяване.

Теорема 2.3.2. Ако $\mu(0) < 1$ и $\mu(\infty) < \infty$, то $m_t = \mathbb{E}[z_t] < \infty$ за всяко t и $m_t^a = \mathbb{E}[z_t^a]$ удовлетворява

$$m_t^a = 1_{[0,a)}(t)\{1 - L(t)\} + \int_0^t m_{t-u}^a \, \mu(\mathrm{d}u). \tag{2.3.1}$$

За аритметична функция μ е изпълнено съответно твърдение.

В демографския случай условията на Теорема 2.3.2 са изпълнени, понеже $f(s)$ е крайна функция. Теоремата предоставя уравнение на възстановяване за очаквания брой на населението след време t, а от теория на възстановяването знаем теоретичното решение на това уравнение. Теорема 2.1.2 гласи, че решението на уравнение (2.3.1) има вида

$$m_t = S * U = \int_0^t S(t-u) \, \mathrm{d}U(u),$$

където $U(u) = \sum\limits_{n=0}^{\infty} \mu^{n*}(u)$ означава функцията на възстановяване, породена от мярката μ.

Определение 2.3.6. Параметърът $m = \mu(\infty)$ се нарича критичен параметър, като ОРП се нарича докритичен при $m < 1$, критичен при $m = 1$, и надкритичен при $m > 1$.

В параграф 2.4 са представени дефиницията и основните свойства на Малтусовия параметър.

Определение 2.4.1. Малтусовият параметър α се дефинира като решение на уравнението $\hat{\mu}(\alpha) = 1$, където $\hat{\mu}(\alpha) = \int\limits_0^{\infty} e^{-\alpha u}\, \mathrm{d}\mu(u)$ е Лаплас–Стилтесовата трансформация на функцията μ.

Следната гранична теорема ни дава по-ясна интерпретацията на този параметър.

Теорема 2.3.2 (виж П. Ягерс [28])**.** Нека условията на Теорема 2.3.2 са изпълнени, нека съществува решение на уравнението $\hat{\mu}(\alpha) = 1$ и нека μ е неаритметична. Тогава за $0 \leq a < \infty$ са изпълнени следните асимптотични свойства при $t \to \infty$:
Ако $\mu(\infty) = 1$ и съществува Малтусов параметър α, то $\alpha = 0$ и

$$m_t^a \to \frac{\int\limits_0^a \{1 - L(u)\}\, \mathrm{d}u}{\int\limits_0^{\infty} u\, \mu(\mathrm{d}u)}.$$

Ако $\mu(\infty) > 1$ и съществува Малтусов параметър α, то $\alpha > 0$ и

$$m_t^a \sim e^{\alpha t} \cdot \frac{\int\limits_0^a e^{-\alpha u}\{1 - L(u)\}\, \mathrm{d}u}{\int\limits_0^{\infty} u e^{-\alpha u}\, \mu(\mathrm{d}u)}.$$

Ако $\mu(\infty) < 1$ и съществува Малтусов параметър α, то $\alpha < 0$ и

$$m_t^a \sim e^{\alpha t} \cdot \frac{\int\limits_0^a e^{-\alpha u}\{1 - L(u)\}\, \mathrm{d}u}{\int\limits_0^{\infty} u e^{-\alpha u}\, \mu(\mathrm{d}u)}.$$

За $a = \infty$ твърдението остава в сила, при условие, че $\int\limits_0^{\infty} u e^{-\alpha u}\, L(\mathrm{d}u) < \infty$. За аритметични функции важат съответни твърдения.

В параграф 2.5 са описани получените нови резултати, които допълват класическата теория на ОРП, за да се моделира население. Необходимо е първо да изследваме разклоняващ се процес, започващ от индивид на произволна възраст, както и от множество индивиди на различни възрасти (включително случайни). Също така и да намерим ефективен числен метод за проектиране на населението. Това са и основните теоретични приноси на този дисертационен труд по отношение на ОРП.

Теорията на ОРП представя очакваното бъдещо население като решение на уравнение на възстановяване. От теория на възстановяването (виж К. Митов и Е. Омей [40]) знаем, че теоретичното решение на това уравнение е конволюция на две функции. Едната от тях се нарича функция на възстановяването и представлява безкрайна сума от конволюции с нарастващ ред. Пресмятането на тези конволюции е изключително тежка изчислителна задача. В допълнение, за да се намери проекцията на цялата възрастова структура на населението, е необходимо да се решат 100 уравнения на възстановяване (по едно за всеки възрастов интервал). Затова е необходимо да се приложи изчислително ефективен подход за решаване на тези уравнения на възстановяване, описан в публикацията [60].

Оказва се, че частен случай на представения в публикацията [60] числен метод е добре известния метод в демографията за проекция на възрастовата структура на населението - с матрица на Лесли. Интересен резултат е, че проекцията с матрица на Лесли дава приблизително решение на уравнението на възстановяване (2.3.1). Теорията на разклоняващите се процеси в непрекъснато време ни дава оценка за грешката на този демографски подход, която възниква от дискретизацията на времето. Същевременно, погледнато от друга гледна точка, се оказва, че матрицата на Лесли представлява числен метод за решаване на интегрални уравнения на възстановяване.

Ако разполагаме с аналитична форма на функциите S и μ, тогава е възможно да намерим теоретичното решение на уравнението на възстановяване (2.3.1), но в общия случай решението няма удобен аналитичен вид. Например, когато моделираме население на базата на реални данни, функциите S и μ могат да представляват изглаждащи сплайни и тогава решението на уравнението на възстановяване няма „удобна“ аналитична форма. Друг проблем е, че дори да направим допускане за аналитичния вид на функциите $S(t)$ и $\mu(t)$, намирането на теоретично решение не винаги е възможно. Същевременно на практика известните аналитични модели на смъртността не описват достатъчно добре наблюдаваните емпирични вероятности за смъртност по възрасти и имат проблеми с оценката на параметрите (виж параграф 1.5). Поради това, в стремежа да създадем математически модел на населението, който максимално отговаря на реалността, се налага да работим с изглаждащи сплайни и числени методи за намиране на решението на уравнение (2.3.1).

Нека означим с ${}_bz_t$ разклоняващ се процес, започващ от една жена на възраст b в момента $t = 0$, ${}_b\xi$ да е съответстващият й точков процес, ${}_b\mu$ да е очакването на точковия процес и ${}_bS$ да е нейната функция на оцеляването (survivability function).

Имаме, че ${}_b\mu(t) = \mathbb{E}(\xi(t+b) - \xi(b) \mid \lambda > b), t > 0$ е очакваният брой деца на една жена след възраст b, ако е доживяла до тази възраст и ${}_bS(t) = \mathbb{P}(\lambda > b + t \mid \lambda > b)$ е вероятността тя да оцелее още t години, ако е доживяла възраст b. Нека ${}_bz_t^a$ е броят на индивидите на възраст по-малка от a в момента t, започнали от жена на възраст b в момента 0. Ако $a = \infty$ и $b = 0$ ще ги пропускаме в означенията. Нека $1_{[0,c]}(t)$ е индикаторна функция: $1_{[0,c]}(t) = 1$, ако $c \geq 0$, $t \in [0, c]$ и $1_{[0,c]}(t) = 0$, ако $c < 0$ или $t > c$.

Теорема 2.5.1. Ако $f(s) = \mathbb{E}(s^{\xi(\infty)}) < \infty, |s| \leq 1$ и $S(b) > 0$, тогава ${}_bm_t = \mathbb{E}({}_bz_t) < \infty$, за всяко t и очакваният брой на индивидите по-млади от a в момента t (или на възраст точно a), започнали от жена на възраст b в момента 0, ${}_bm_t^a = \mathbb{E}({}_bz_t^a)$, удовлетворява

$$
{}_bm_t^a = {}_bS(t)1_{[0,a-b]}(t) + \int\limits_0^t m_{t-u}^a \, {}_b\mu(\mathrm{d}u), \tag{2.5.1}
$$

където ${}_bS(t) = \frac{S(b+t)}{S(b)}$ означава вероятността жена на възраст b да доживее до $b + t$, ${}_b\mu(t) = \frac{\mu(t+b)-\mu(b)}{S(b)}$ е очакването на нейният точков процес, а m_{t-u}^a е решението на уравнение (2.3.1).

Ако $S(t)$ и $\mu(t)$ са два пъти диференцируеми по t, тогава ${}_bm_t^a$ е два пъти диференцируема и по t, за $t \neq a - b$, и по b, за $b \neq a$. Ако $a = \infty$, тогава ${}_bm_t$ е два пъти диференцируема за всяко $t > 0$.

Нека ω е максималната възраст в таблиците за смъртност, т.е. $\mathbb{P}(\lambda > \omega) = 0$. С други думи това е максималната възраст, до която може да доживее човек. Нека също за улеснение на записа да е кратна на h.

Нека $N_t(bh; (b+1)h]$ е броят на жените на възраст $(bh; (b+1)h]$ в момента t, т.е. това е възрастовата структура на населението. Например, ако разклоняващият се процес е започнал от една жена на възраст 0 в момента 0, то $z_t = N_t[0; h] + \sum\limits_{b=1}^{\omega/h-1} N_t(bh; (b+1)h]$ и $N_t(bh; (b+1)h] = z_t^{bh+h} - z_t^{bh}$.

Следната теорема показва, че очакваната възрастова структура в момента $t + h$ може да бъде пресметната приближено от очакваната възрастова структура в момента t. Теоремата се отнася за разклоняващи се процеси, започващи от един или краен брой индивиди на различни възрасти (включително случайни).

Теорема 2.5.2. Нека са изпълнени условията на Теорема 2.5.1 и $S(t)$ и $\mu(t)$ са два пъти диференцируеми в $\mathbb{R}$ и $\mu(t) = 0$ за $t \leq 0$. Нека ${}_u\mu''(t)$ е ограничена функция на t и u, а $S''(t)$ е ограничена функция на t. Тогава следното числено приближение е в сила при $h \to 0$:

$$\begin{aligned}
\mathbb{E}N_{t+h}((b+1)h;(b+2)h] &= \mathbb{E}N_t(bh,(b+1)h] \cdot \left[{}_{bh}S(h) + O(h^2)\right], \\
\mathbb{E}N_{t+h}[0;h) &= \sum_{b=0}^{\omega/h-1} \mathbb{E}N_t(bh,(b+1)h] \cdot {}_{bh}\mu(h) + \mathbb{E}N_t[0,\omega] \cdot O(h^2),
\end{aligned} \tag{2.5.2}$$

където b е цяло число, $b \geq 0$. Алтернативна формулировка на уравнение (2.5.2) е:

$$\mathbb{E}N_{t+h}[0;h) = \sum_{b=0}^{\omega/h-1} [\mathbb{E}N_t(bh,(b+1)h] \cdot {}_{bh}\mu'(0)h] + \mathbb{E}N_t[0,\omega] \cdot O(h^2).$$

Константата в члена $O(h^2)$ не зависи от избора на b.

Нека първо въведем следните матрични означения, които ще улеснят означенията в следващите твърдения:

$$\mathbf{A} = \begin{bmatrix} {}_0\mu(h) & {}_h\mu(h) & \dots & {}_{\omega-2h}\mu(h) & {}_{\omega-h}\mu(h) \\ {}_0S(h) & 0 & \dots & 0 & 0 \\ 0 & {}_hS(h) & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & {}_{\omega-2h}S(h) & 0 \end{bmatrix}_{\frac{\omega}{h}\times\frac{\omega}{h}},$$

$$\mathbf{B} = \begin{bmatrix} {}_0\mu'(0)h & {}_h\mu'(0)h & \dots & {}_{\omega-2h}\mu'(0)h & {}_{\omega-h}\mu'(0)h \\ {}_0S(h) & 0 & \dots & 0 & 0 \\ 0 & {}_hS(h) & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & {}_{\omega-2h}S(h) & 0 \end{bmatrix}_{\frac{\omega}{h}\times\frac{\omega}{h}},$$

$$\mathbb{E}\mathbf{N_t} = \begin{bmatrix} \mathbb{E}N_t[0;h] \\ \mathbb{E}N_t(h;2h] \\ \vdots \\ \mathbb{E}N_t(\omega-h;\omega] \end{bmatrix}_{\frac{\omega}{h}\times 1}, \quad \mathbf{O(h^2)} = \begin{bmatrix} O(h^2) \\ O(h^2) \\ \vdots \\ O(h^2) \end{bmatrix}_{\frac{\omega}{h}\times 1} \quad \text{и } \mathbf{1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{\frac{\omega}{h}\times 1}.$$

Следствие 2.5.1. Нека предположим, че жените могат да имат само по едно дете във възрастовия интервал $(b;b+1]$. Нека p_b означава условната вероятност една жена да роди в този интервал и s_b означава условната вероятност една жена да доживее до края на интервала, ако е била жива в началото му. Тогава Теорема 2.5.2 остава в сила, като матрицата $\mathbf{A}$ придобива следната форма:

$$\mathbf{A} = \begin{bmatrix} p_0 & p_1 & \dots & p_{\omega-2} & p_{\omega-1} \\ s_1 & 0 & \dots & 0 & 0 \\ 0 & s_2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & s_{\omega-2} & 0 \end{bmatrix}_{\omega\times\omega}.$$

Забележка. Тази матрица е известна в демографията като матрица на Лесли и се използва като стандартен метод за проекция на населението (Н. Кийфиц [29]).

Теорема 2.5.3. Нека първоначалното население в момента $t = 0$ е съставено от краен брой индивиди на случайни възрасти, с абсолютно непрекъснато разпределение. Нека $\mathbb{E}N_0[0,h]$, $\mathbb{E}N_0(bh,(b+1)h]$, за $b = 1,\ldots,(\omega/h-1)$, е очакваната структура на населението в момента $t = 0$. Нека условията на Теорема 2.5.1 са изпълнени и нека ${}_u\mu''(v)$ е ограничена в $[0,t]^2$ и $S''(v)$ е ограничена в $[0,t]$. Тогава за всяко $k \le t/h$ имаме, че

$$\mathbb{E}\mathbf{N_{kh}} = \mathbf{A}^k \cdot \mathbb{E}\mathbf{N_0} + \mathbf{O(h^2)} \tag{2.5.4}$$

и общия брой на населението е

$$\mathbf{1}^\intercal \cdot \mathbb{E}\mathbf{N_{kh}} = \mathbf{1}^\intercal \cdot \mathbf{A}^k \cdot \mathbb{E}\mathbf{N_0} + O(h). \tag{2.5.5}$$

Аналогично твърдение е в сила и ако използваме за приближение матрицата **B** вместо **A**.

Забележка. Ако фиксираме периода t и намаляваме h, грешката на апроксимацията в уравнения (2.5.4) и (2.5.5) клони към нула. Интересно е да се отбележи обаче, че грешката зависи от t - ако фиксираме h и увеличаваме t, грешката нараства.

Теорема 2.5.4. Нека $N_0[0,+\infty) = N_0[b,b] = 1$, т.е. имаме само един индивид в началния момент и тя е на възраст b. Тогава за всяко $k = 1,\ldots,t/h$

$$ {}_bm_{kh} = \begin{bmatrix} 1 & \ldots & 1 \end{bmatrix}_{1\times\frac{\omega}{h}} \cdot \begin{bmatrix} {}_0\mu(h) & {}_h\mu(h) & \ldots & {}_{\omega-2h}\mu(h) & {}_{\omega-h}\mu(h) \\ {}_0S(h) & 0 & \ldots & 0 & 0 \\ 0 & {}_hS(h) & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & {}_{\omega-2h}S(h) & 0 \end{bmatrix}^k_{\frac{\omega}{h}\times\frac{\omega}{h}} \cdot \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{\frac{\omega}{h}\times 1} + O(h),$$

където единственият ненулев елемент в този вектор е на позиция $b/h + 1$, съответстваща на възрастовия интервал $(b, b + h]$. Като алтернатива можем да използваме за апроксимация матрицата **B** вместо **A** .

Следствие 2.5.2. Нека условията на Теорема 2.5.3 са изпълнени. Тогава решенията на уравненията на възстановяване (2.3.1) за всички възрасти a могат да се пресметнат едновременно с формулата:

$$\begin{bmatrix} m_{kh}^{h} \\ m_{kh}^{2h} - m_{kh}^{h} \\ \vdots \\ m_{kh}^{\omega} - m_{kh}^{\omega-h} \end{bmatrix} \approx \begin{bmatrix} {}_0\mu(h) & {}_h\mu(h) & \dots & {}_{\omega-2h}\mu(h) & {}_{\omega-h}\mu(h) \\ {}_0S(h) & 0 & \dots & 0 & 0 \\ 0 & {}_hS(h) & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & {}_{\omega-2h}S(h) & 0 \end{bmatrix}^{k}_{\frac{\omega}{h}\times\frac{\omega}{h}} \cdot \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{\frac{\omega}{h}\times 1}$$

за всяко $k = 1, \dots, t/h$. Очакваното население m_{kh}^{a} е приблизително равно на сумата на първите a/h на брой елементи на вектора по-горе.

Забележка. Интересно е да се отбележи, че ако $t \leq \omega$, то можем да използваме отрязаната матрица от първите (t/h) на брой редове и колони и ще получим същото приближение.

Ако разгледаме популация, започваща от индивид от мъжки пол, то $\mu(t) = 0$ за всяко t и уравнение (2.3.1) придобива вида ${}_bm_t^a = {}_bS(t)1_{[0,a-b]}(t)$. Това значи, че мъжете допринасят към броя на популацията единствено с техният живот, докато жените допринасят включително и с живота на техните деца. Другият факт, който трябва да се отчете в модела, е, че жените раждат и момчета, така че броят на живородените мъже зависи от очакваната възрастова структура на жените (която се изчислява с уравнение (2.5.1)). Тогава уравнение (2.5.2), записано за мъжете е:

$$\mathbb{E}N_{t+h}^{males}((b+1)h;(b+2)h] \approx \mathbb{E}N_t^{males}(bh,(b+1)h] \cdot {}_{bh}S^{males}(h),$$
$$\mathbb{E}N_{t+h}^{males}[0;h] \approx \sum_{b=0}^{\omega/h-1} \mathbb{E}N_t(bh,(b+1)h] \cdot r \cdot {}_{bh}\mu(h),$$

където r означава отношението между броя на родените мъже и жени, а $b \geq 0$ е цяло число. Можем да считаме, че $r = 1.05$, тъй като средно на всеки 100 момичета се раждат около 105 момчета или да получим оценка за това отношение от наличната статистика.

В параграф 2.6 са представени основните изводи и следствия от получените нови теоретични резултати и тяхното приложение. Следствие 2.5.2 ни дава числено решение на уравнението на възстановяване (2.3.1). Поради това, че се състои единствено от вдигане на степен на матрица, която освен това има ненулеви елементи единствено на първия си ред и един от диагоналите, численият метод работи бързо и ефективно. Освен това на всяка стъпка от умножението с матрицата $\mathbf{A}$ получаваме решенията на всички уравнения на възстановяване, за всички възрасти a едновременно.

Знаейки първоначалната възрастова структура на неселението и функциите $S(t)$ и $\mu(t)$ можем да пресметнем очакваната бъдеща структура на населението, прилагайки Теорема 2.5.2.

Теорема 2.5.3 показва, че разполагайки с началната възрастова структура на населението $\mathbf{N_0}$ и матрицата $\mathbf{A}$, можем да направим прогноза за бъдещия брой на населението. В Следствие 2.5.1 е показана връзката между матрицата $\mathbf{A}$ и матрицата на Лесли, използвана в демографията за проекция на населението. Оказва се, че този стандартен демографски метод за проекция на населението всъщност дава приближено числено решение на уравненията на възстановяване в Теорема 2.3.2, като грешката му е изведена с теорията на ОРП (виж Теорема 2.5.3). При всяко умножение на текущата възрастова структура на населението $\mathbf{N_0}$ с матрицата $\mathbf{A}$ получаваме приближение на очакваната структура на населението за следващия период от време.

Един от най-известните модели на населението е този Малтус (1798 г.), който показва експоненциален растеж на населението, без да отчита текущата му възрастова структура и стохастиката на процеса. Матричните модели за прогнозиране на населението са въведени още през 40-те години на миналия век от Бернардели (виж [6], 1941 г.), Левис (виж [35], 1942 г.) и Лесли (виж [33, 34], 1945–1948 г.), като първоначално те са детерминистични и не отчитат стохастиката на населението, но вече отчитат текущата възрастова структура. По-късно Полард (виж [46], 1966 г.) показва стохастичен модел на населението, който всъщност представлява многотипов разклоняващ се процес на Галтон–Уотсън (виж [28, 52, 64]), и извежда рекурентни формули за пресмятане на вариацията на този разклоняващ се процес. Ягерс представя общ разклоняващ се процес с непрекъснато време и прави първите стъпки към приложението му в демографията (виж [28]).

Едно от забележителните свойства на съвременните модели за население е, че те всъщност не изключват, а съдържат в себе си предишните модели за населението. Така например се оказва, че матричните детерминистични модели, както и разклоняващите се процеси, предсказват експоненциално нарастване на населението в безкрайността. Те подобряват Малтусовият модел, давайки по-прецизна прогноза за населението в краткосрочен хоризонт. Също така се оказва, че многотиповият разклоняващ се процес на Галтон–Уотсън дава същата прогноза за населението като матричния детерминистичен модел на Лесли, като същевременно обаче разглежда и стохастиката на населението. Моделът на население с ОРП също потвърждава това правило - очакваното население може да се приближи с матрица на Лесли и нараства/намалява с експоненциална скорост в безкрайността, както при модела на Малтус. Променяйки матрицата на Лесли за всеки бъдещ период от време ни позволява дори да преодолеем допускането за непроменливи закони на смъртността и раждаемостта и да моделираме разклоняващ се процес (население) в променлива среда. Същевременно изследвайки какво би станало, ако законите на смъртността и раждаемостта останат същите, получаваме също ценна информация за сегашното демографско състояние (виж параграф 3.3). Приложението на ОРП и практическите похвати, с които можем да преодолеем някои от привидните ограничения на модела, са разгледани в следващата Глава 3.

• В Глава 3 е представено приложението на ОРП в демографията за прогнозиране и моделиране на населението. Като пример е направена прогноза за населението на България с ОРП и съпоставка с други европейски държави, както и оценка на демографското им състояние.

В параграф 3.1 е разгледана последователността от стъпки, необходими за моделирането и прогнозирането на стохастичните закони на смъртността и раждаемостта от наличните демографски данни. За да приложим Теорема 2.5.3 за проекция на населението, първо се нуждаем от модел за разпределението на продължителността на живота, $S(t)$, и очаквания брой на децата на една жена, $\mu(t)$, до навършване на възраст t. Процесът на моделиране на тези две функции е многокомпонентен, включващ много демографски и статистически похвати. Те са описани подробно в Глава 1, като в този параграф е представена тяхната последователност като методология.

От данните, с които разполагаме (публикувани от Евростат [20]), бихме могли да намерим повъзрастовите коефициенти за смъртност и раждаемост, т.е. таблиците за смъртност (виж Н. Кийфиц [29], С. Престън [47]). Приложен е моделът на Канисто за смъртност на високите възрасти (виж [58]) за годините с липсващи данни за смъртността и методът „survivor ratio“ за годините с липсващи данни за броя на населението по възрасти. Изчерпателна и съвременна методология за пресмятане на таблиците за смъртност и попълване на липсващите данни е представена в „Human Mortality Database“ (виж [65]) и Школников (виж [49]). След прилагане на методологията и пресмятане на възрастовите коефициенти за смъртност и раждаемост можем да използваме формулата на Чианг за пресмятане на условните вероятностите за смъртност и раждаемост по възрасти (виж [10]). Пресметнатите емпирични вероятности съдържат шум, който се отстранява посредством анализ на функционални данни (виж Приложение А, Д. Рамзи и Б. Силвърман [48] и Р. Хиндман [66]).

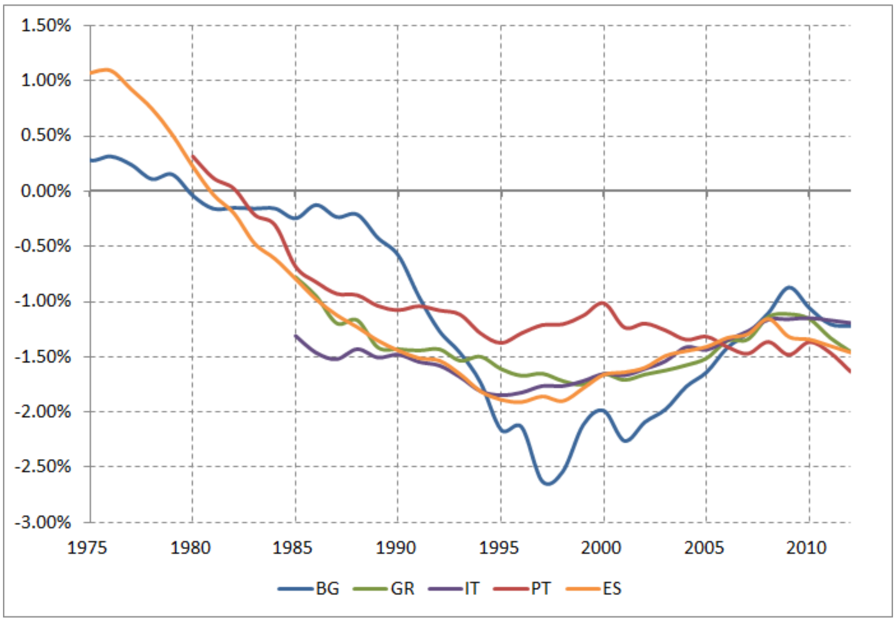
За всяка изминала година можем да намерим модели на функциите $S(t)$ и $\mu'(t)$, като така получаваме времеви ред за тяхната еволюция. След прилагане на анализ на главните компоненти (PCA), коефициентите пред базисните функции са моделирани чрез авторегресионен модел AR(1), както е предложено от Хиндман (виж [66]). Резултатите за България са представени на Фигури 3.4, 3.5 и 3.6 в дисертационния труд. Прогнозите са вероятностите за смъртност и раждаемост по възрасти са необходими за моделирането на ОРП. Ако симулираме всеки главен компонент с AR(1) модел, тогава линейната комбинация на базисните функции със симулираните главни компоненти ще ни даде симулации за функциите $S(t)$ и $\mu(t)$.

В параграф 3.2 е разгледано приложението на получените теоретични резултати за прогнозиране на населението с ОРП. Разгледана е симулационна процедура за ОРП, която отчита стохастиката на променливите закони на раждаемостта и смъртността, както и стохастиката на самия ОРП.

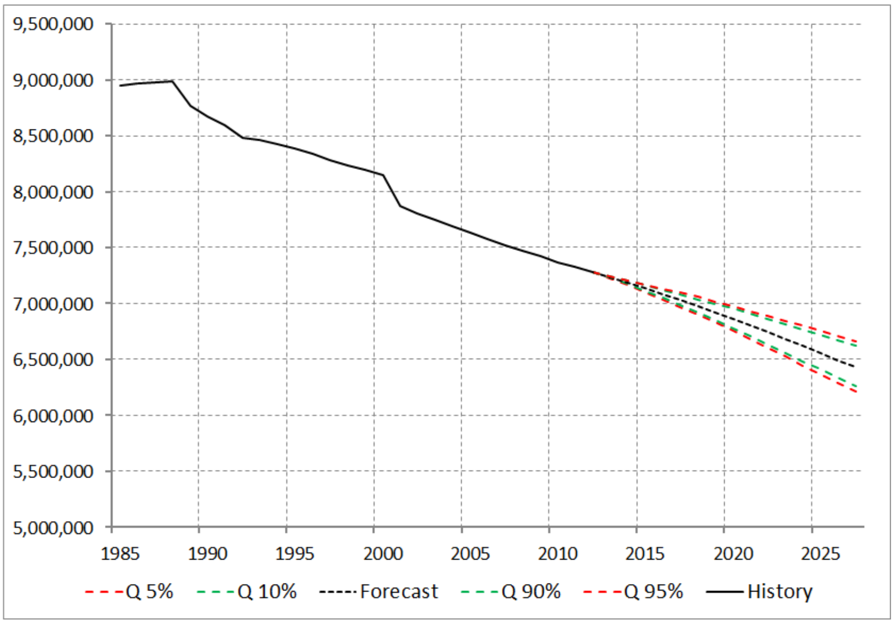
В параграф 3.3 са представени резултатите за Малтусовия параметър на ОРП и неговата динамика във времето за няколко европейски страни. Предвид това, че $S(t)$ и $\mu(t)$ са гладки функции, то такава функция е и решението m_t на уравнението на възстановяване (2.3.1). Малтусовият параметър тогава може да се пресметне като граница $\alpha = \lim_{t\to\infty}(ln(m_t))'$, след като намерим m_t по методологията в параграф 2.5. Малтусовият параметър може да бъде използван като мярка за демографското състояние на населението в различни държави или по различно време в историята на една държава, тъй като комбинира едновременно профила на раждаемостта и смъртността по възрасти. Страните с по-голяма стойност на Малтусовия параметър имат по-висока стойност на очакваното население в дългосрочен план. Ако Малтусовият параметър е по-малък от нула, тогава ОРП е „докритичен“, което значи, че очакваното население клони експоненциално към нула и вероятността за изчезване е 1 (виж [28, 52]). Ако Малтусовият параметър е по-голям от нула, тогава процесът е „надкритичен“ и неговото очакване нараства експоненциално към безкрайност, като вероятността за израждане е по-малка от 1. ОРП е „критичен“, ако Малтусовият параметър е 0, като тогава очакването му клони към положителна константа в безкрайността.

В параграф 3.4 са представени резултатите от прогнозата на населението на базата на ОРП, както и доверителни интервали за броя на населението и неговата възрастова структура. В този параграф се използва представената методология, базирана на ОРП, за да се изследва миналото и бъдещето на демографските процеси в България, Гърция, Испания, Италия и Португалия. Част от представените резултати са публикувани в [63, 62, 61]. Фигура 3.9 показва историята на Малтусовия параметър за тези държави. Интересно е, че стойностите на този параметър за представените държави са много близки помежду си и исторически са имали сходна динамика. Възможно обяснение на това е, че европейските държави споделят еднаква култура и история, еднакви икономически цикли и съществуват общи фактори, които им влияят едновременно. Забелязват се спадове в този параметър при периоди на икономическа криза, които могат да се използват за измерване на влиянието ѝ върху демографията на страната и критичността на разклоняващия се процес.

Очакваният брой на населението на България е представен на Фигура 3.12 заедно с доверителни интервали на прогнозата. Прогнозата показва намаляващо население на България за следващите 15 години. Симулациите показват, че най-вероятно спадът на населението ще е между 625,310 и 1,074,884, като вероятността да наблюдаваме по-голям спад е 5%. Едно от важните приложения на симулациите е, че ни позволяват да изследваме източниците на риск за прогнозата. Например, ако симулираме ОРП за населението на България за хоризонт от 15 години, използвайки прогнозите за $S(t)$ and $\mu(t)$ (техните очаквани стойности), тогава 90% доверителен интервал за броя на населението има широчина -0.03% до +0.03% от прогнозирания брой население. Този

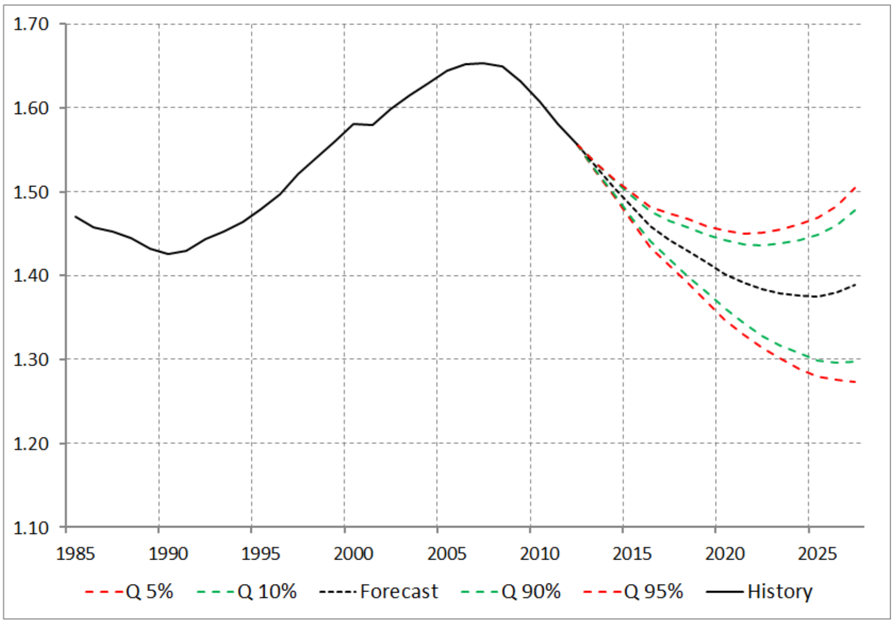


Фигура 3.9: История на Малтусовия параметър.



Фигура 3.12: Общ брой население на България.

доверителен интервал отразява единствено стохастиката на ОРП. Ако добавим несигурността, породена от прогнозата на вероятностите за раждаемост, то доверителният интервал се разширява на -3.06% до +3.51% от прогнози-



Фигура 3.13: Отношение на хората в работоспособна възраст и извън нея.

раното население. Ако добавим към стохастиката на ОРП и несигурността, произтичаща едновременно от прогнозата на раждаемостта и смъртността по възрасти, то доверителният интервал е от -3.42% до +3.58% от прогнозираното население. Забелязва се, че най-голям принос към риска на прогнозата има раждаемостта, т.е. променливостта на функцията $\mu(t)$. Общият брой на населението е най-чувствителен към вероятностите за раждаемост на населението. Друг интересен резултат е, че стохастиката на ОРП представлява малък дял от общия риск на прогнозата, в сравнение с другите източници на риск (стохастиката на раждаемостта и смъртността). Това се дължи всъщност на големия брой хора в населението и закона за големите числа - дисперсията на процентната грешка е по-малка за по-голямо население. Въпреки това, ако се интересуваме от малки групи хора, като например конкретни възрастови групи или населението в малки населени места, тогава нараства значението на стохастиката на ОРП за риска на прогнозата.

На Фигура 3.13 можем да видим отношението между броя на хората в работоспособна възраст и хората извън нея (по-млади от 18 г., мъже над 64 г. и жени над 61 г.). Това отношение (още наричано демографско натоварване) намалява през последните 4 години главно поради възрастовата структура на населението и ще продължи да намалява до 2025 г., когато се очаква да се увеличи за кратко.

Процентът на населението в работоспособна възраст се очаква да намалее в България с 1-5% (процентни пункта), в Испания с 2-4%, в Гърция с 1-3%, в Италия с 3-5%, в Португалия с 1-3%. Резултатите показват, че постепенна пенсионна реформа е възможно да бъде наложителна за всяка една от тези държави, ако искаме да запазим процента на населението в работоспособна възраст на сегашното ниво. Наблюдаваните негативни тенденции са всъщност общи за повечето европейски държави, които преживяват сходни демографски процеси.

• В Приложение А са разгледани избрани твърдения от анализа на функционални данни. Тази съвременна теория се използва за моделиране на възрастовата структура на вероятностите за смъртност и раждаемост по възрасти (виж Глава 1), които от своя страна се използват в модела на ОРП (виж Глава 2).

• В Приложение Б са представени избрани части от програмния код, който осъществява моделирането и прогнозирането чрез ОРП.

Апробация

Резултатите в дисертационния труд са докладвани на Doctoral Conference in Mathematics, Informatics and Education (September 19-22, 2013, Sofia, Bulgaria), пролетната научна сесия на ФМИ (март 2014 г., 2015 г. и 2016 г.), 3rd Stochastic Modeling Techniques and Data Analysis International Conference (SMTDA 2014), XVI-th International Summer Conference on Probability and Statistics, Seminar on Statistical Data Analysis, Workshop on Branching Processes and Applications, Dedicated to the memory of B.A. Sevastyanov Pomorie (21-28 юни, 2014), юбилейната научна конференция по случай 125 години обучение по математика и природни науки в СУ „Св. Климент Охридски", Workshop on Branching Processes and Applications (WBPA-2015), 8th International Conference of the ERCIM WG on Computational and Methodological Statistics (CMStatistics 2015, 12-14 декември, 2015), Национален семинар по стохастика, 10 февруари, 2016 г. Списък от публикации по темата:

1. P. Trayanov. Crump–Mode–Jagers branching process: Modelling and application for human population. Pliska Stud. Math. Bulgar., 22:207–224, 2013.

2. P. Trayanov and M. Slavtchova-Bojkova. Crump–Mode–Jagers branching processes: Application in population projections. Advanced research in mathematics and computer science, Proceedings of the doctoral conference in mathematics, informatics and education, Sofia, Bulgaria, 19-20.09.2013, 1-–8, September 2013.

3. P. Trayanov and M. Slavtchova-Bojkova. Branching Processes: Forecasting Human Population. SMTDA2014 Conference Proceedings, Lisbon, Portugal, 11-17.06.2014, 141–149, 2014.

4. P. Trayanov and M. Slavtchova-Bojkova. Estimating The Effect of Economic Crisis with Crump–Mode–Jagers Branching Process. Pliska Stud. Math. Bulgar. 24:99–110, 2015.

5. P. Trayanov. Crump–Mode–Jagers branching process: A numerical approach. Branching Processes and their Applications, Springer, 2016 (in press).

Авторска справка

Настоящият дисертационен труд е посветен на приложението на ОРП за моделиране и прогнозиране на население. При изпълнение на поставената задача са изведени нови резултати в теорията на разклоняващите се процеси, които преодоляват някои от съществуващите ограничения на класическия ОРП и позволяват чрез него да се моделира население. Основните научни и научно-приложни приноси могат да бъдат обобщени по следния начин:

• Представено е интегрално уравнение за очакваното население, започващо от жена на произволна възраст към момента 0, както и обобщение за население, започващо от много на брой жени на различни възрасти към момента 0.

• За да намерим очакваната бъдеща структура на населението е необходимо да решим голям брой уравнения на възстановяване. Представен е числен метод за проектиране на възрастовата структура на населението, базиран на теорията на ОРП, който също би могъл да се използва и за решаване на уравнения на възстановяване, удоволетворяващи определени условия за гладкост. Представеният числен метод включва единствено изчислителни операции за вдигане на степен на матрица, което води до голяма изчислителна ефективност и скорост на алгоритъма.

• Доказано е, че проекцията на населението с матрица на Лесли, широко използвана в демографията, е всъщност частен случай на разгледания числен метод. Прилагайки теорията на ОРП в непрекъснато време ни предоставя оценка на грешката при използването на матрица на Лесли, която произлиза от дискретизация на времето.

• Представена е методология за моделиране и прогнозиране на население с Общ разклоняващ се процес (ОРП), която е базирана на демографските данни с които разполагаме. Направени са симулации на ОРП, които отразяват несигурността в прогнозата на стохастичните закони на раждаемостта и смъртността, както и несигурността произтичаща от самия ОРП, в който те участват. Анализиран е приносът на отделните източници на несигурност към риска на прогнозата за възрастовата структура на населението.

• Като пример за ползите от използването на ОРП в демографията е представена прогноза за възрастовата структура на населението на България за хоризонт от 15 години, доверителни интервали за направената прогноза, както и сравнение на демографското състояние на различни европейски страни през различни години посредством Малтусовия параметър, изчислен с помощта на ОРП.

Благодарности

Искам да изкажа своята благодарност на научния си ръководител проф. Марусия Божкова за подкрепата и ползотворните дискусии и насоки по време на докторантурата, както и за създадената благоприятна атмосфера за развитие и работа! Благодаря още и на проф. Косто Митов, чиито лекции по Процеси на възстановяване ми бяха изключително полезни при подготовката на дисертационния труд. Благодарности на всички колеги от катедра Вероятности, операционни изследвания и статистика на ФМИ–СУ за подкрепата и създадената благоприятна творческа атмосфера.

Библиография

[1] Н. Обрешков. *Теория на вероятностите.* Наука и Изкуство, 1963.

[2] А. Обретенов. *Теория на вероятностите.* Наука и Изкуство, 1974.

[3] J. Alho and B. Spencer. *Statistical Demography and Forecasting.* Springer, first edition, July 2005.

[4] D. Atanasov, V. Stoimenova, and N. Yanev. Estimators in branching processes with immigration. *Pliska Stud. Math. Bulgar.*, 18:19–40, 2007.

[5] D. Bartholomew. An approximate solution to the integral equation of renewal theory. *J. Royal Statist. Soc.*, B 25(2):432–441, October 1963.

[6] H. Bernardelli. Population waves. *Journal of Burma Research Society*, 31:1–18, 1941.

[7] R. Boiko and V. Ryazanov. Stochastic model of a nuclear power reactor. *Atomic Energy*, 93(2):625–634, 2002.

[8] G. Box, G. Jenkins, and G. Reinsel. *Time Series Analysis: Forecasting and Control.* Prentice-Hall, Inc., third edition, February 1994.

[9] C. Chiang. *Introduction to Stochastic Processes in Biostatistics.* Wiley, John & Sons, Incorporated, 1968.

[10] C. Chiang. *Life Table and Its Applications.* Krieger Pub Co, December 1984.

[11] K. Crump and C. Mode. A general age-dependent branching process I. *Journal of Mathematical Analysis and Applications*, 24(3):494–508, 1968.

[12] K. Crump and C. Mode. A general age-dependent branching process II. *Journal of Mathematical Analysis and Applications*, 25(1):8–17, 1969.

[13] J. Cryer and K. Chan. *Time Series Analysis With Applications in R.* Springer, second edition, April 2008.

[14] N. Daskalova. EM estimation of the offspring distribution in multitype branching processes – a model in cell kinetics. *Pliska Stud. Math. Bulgar.*, 20:463–474, 2010.

[15] N. Daskalova. Using inside-outside algorithm for estimation of the offspring distribution in multitype branching processes. *Serdica Journal of Computing*, 4(4):45–52, 2011.

[16] N. Daskalova. Maximum likelihood estimation in multitype branching processes with terminal types. *Compt. rend. Acad. bulg. Sci.*, 65(5):575–580, 2012.

[17] C. de Boor. *A Practical Guide to Splines.* Springer, November 2001.

[18] R. Durrett. *Branching Process Models of Cancer.* Springer, 2015.

[19] Heyde C.C. (Editor). *Branching Processes. Proceedings of the First World Congress.* Springer, 1995.

[20] Eurostat. Eurostat database. `http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database`.

[21] S. From. Some new approximations for the renewal function. *Commun. Statist.-Simula.*, 30(1):113–128, 2001.

[22] M. Gonzalez, R. Martinez, and M. Slavtchova-Bojkova. Time to extinction of infectious diseases through age-dependent branching models. *Ch.17 in: Workshop on Branching Processes and Their Application, Lecture Notes in Statistics – Proceedings, M. Gonzales, M. Molina, I. del Puerto, M. Mota, R. Martinez, A. Ramos (Editors), Springer*, pages 241–256, 2010.

[23] P. Green. Modelling yeast cell growth using stochastic branching processes. *J. Appl. Prob.*, 18(4):799–808, 1981.

[24] S. Grishechkin. Crump–Mode–Jagers branching processes as a method of investigating M/G/1 systems with processor sharing. *Theory Probab. Appl.*, 36(1):16–33, 1991.

[25] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Springer, second edition, February 2009.

[26] O. Hyrien and N. Yanev. Asymptotic behavior of cell populations described by two-type reducible age-dependent branching processes with non-homogeneous immigration. *Mathematical Population Studies*, 19:164—-176, 2012.

[27] P. Jagers. A general stochastic model for population development. *Skand. Aktuarietidskr.*, 52:84–103, 1969.

[28] P. Jagers. *Branching Processes with Biological Applications.* John Wiley & Sons Ltd, first edition, January 1975.

[29] N. Keyfitz and H. Caswell. *Applied Mathematical Demography.* Springer, third edition, January 2005.

[30] M. Kimmel and D. Axelrod. *Branching Processes in Biology.* Springer, 2002.

[31] A. Kolmogorov and N. Dmitriev. Branching random processes. *Dokl. Acad. Nauk SSSR*, 56(1):7—10 (See, English transl., Selected works of A. N. Kolmogorov, vol. II: Probability theory and Mathematical Statistics, Kluwer, Dordrecht, 1992), 1947.

[32] A. Kolmogorov and B. Sevastyanov. The calculation of final probabilities for the branching random processes. *Dokl. Acad. Nauk SSSR*, 56(8):783–786 (See, English transl., Selected works of A. N. Kolmogorov, vol. II: Probability theory and Mathematical Statistics, Kluwer, Dordrecht, 1992), 1947.

[33] P. Leslie. On the use of matrices in certain population mathematics. *Biometrika*, 33(3):183–212, 1945.

[34] P. Leslie. Some further notes on the use of matrices in population mathematics. *Biometrika*, 35:213–245, 1948.

[35] E. Lewis. On the generation and growth of a population. *Sankhya, Indian Journal of Statistics*, 6(1):93–96, 1942.

[36] P. Mayster. Stationary distributions of the alternating branching processes. *Ch.4 in: Workshop on Branching Processes and Their Application, Lecture Notes in Statistics – Proceedings, M. Gonzales, M. Molina, I. del Puerto, M. Mota, R. Martinez, A. Ramos (Editors), Springer*, pages 53–67, 2010.

[37] G. Mitov and K. Mitov. Randomly indexed Galton–Watson branching processes. *Proceedings of the 35th Spring Conference of the UBM: Mathematics and Education in Math.*, pages 275–281, 2006.

[38] G. Mitov, S. Rachev, Y. Kim, and F. Fabozzi. Barrier option pricing by branching processes. *Int. J. Theor. Appl. Finan.*, 12:1055–1073, 2009.

[39] K. Mitov, G. Mitov, and N. Yanev. Limit theorems for critical randomly indexed branching processes. *Ch.7 in: Workshop on Branching Processes and Their Application, Lecture Notes in Statistics – Proceedings, M. Gonzales, M. Molina, I. del Puerto, M. Mota, R. Martinez, A. Ramos (Editors), Springer*, pages 95–108, 2010.

[40] K. Mitov and E. Omey. *Renewal Processes.* Springer, 2013.

[41] K. Mitov and E. Omey. Intuitive approximations for the renewal function. *Statistics and Probability Letters*, 84:72–80, 2014.

[42] K. Mitov and N. Yanev. Regenerative branching processes. *Ch.3 in: Records and Branching processes, M. Ahsanullah, G.P. Yanev (Editors), Nova Science Publishers, Inc.*, pages 37–62, 2008.

[43] K. Mitov and N. Yanev. Branching stochastic processes: Regulation, regeneration, estimation, application. *Pliska Stud. Math. Bulgar.*, 19(1):5–58, 2009.

[44] C. Mode. *Stohastic Processes In Demography and Their Computer Implementation.* Springer, October 1985.

[45] P. Oloffson, O. Schwalb, R. Chakraborty, and M. Kimmel. An application of a general branching process in the study of the genetics of aging. *J. Theor. Biol.*, 213(4):547–557, 2001.

[46] J. Pollard. On the use of the direct matrix product in analysing certain stochasic population models. *Biometrika*, 53:397–415, 1966.

[47] S. Preston, P. Heuveline, and M. Guillot. *Demography: Measuring and Modeling Population Processes*. Wiley-Blackwell Publishing, 2000.

[48] J. Ramsay and B. Silverman. *Functional Data Analysis*. Springer, second edition, 2005.

[49] V. Shkolnikov. Methodology note on the human life-table database (hld). Technical report, 2007.

[50] M. Slavtchova-Bojkova, P. Becker-Kern, and K. Mitov. Total progeny in a subcritical branching processes with state-dependent immigration. *Pliska Stud. Math. Bulgar.*, 16:229–245, 2004.

[51] M. Slavtchova-Bojkova, M. Gonzalez, and R. Martinez. Age-dependent branching processes for surveillance of vaccine-preventable diseases with incubation period. *Front. Psychiatry 1:127, DOI 10.3389/FPSYT.2010.00127.*, 2010.

[52] M. Slavtchova-Bojkova and N. Yanev. *Branching Stochastic Processes.* University Publ. "St. Kl.Ohridski", 2007.

[53] V. Stoimenova. Robust parametric estimation of branching processes with a random number of ancestors. *Serdica - Bulg. Math. Journal*, 31(3):243–262, 2005.

[54] V. Stoimenova, D. Atanasov, and N. Yanev. Simulation and robust modifications of estimates in branching processes. *Pliska Stud. Math. Bulgar.*, 16:259–271, 2004.

[55] V. Stoimenova, D. Atanasov, and N. Yanev. Robust estimation and simulation of branching processes. *C.R. Acad. Bulg. Sci.*, 57(5):19–22, 2004a.

[56] V. Stoimenova, D. Atanasov, and N. Yanev. Algorithms for generation and robust estimation of branching processes with random number of ancestors. *Proceedings of the 34th Spring Conference of the UBM: Mathematics and Education in Math.*, pages 196–201, 2005.

[57] R. Thatcher, V. Kannisto, and K. Andreev. The survivor ratio method for estimating numbers at high ages. *Demographic Research*, 6:1–18, 2002.

[58] R. Thatcher, V. Kannisto, and J. Vaupel. *The Force of Mortality at Ages 80 to 120.* Odense University Press, 1998.

[59] P. Trayanov. Crump–Mode–Jagers branching process: Modelling and application for human population. *Pliska Stud. Math. Bulgar.*, 22:207–224, 2013.

[60] P. Trayanov. Crump–Mode–Jagers branching process: A numerical approach. *Springer*, 2016 (in press).

[61] P. Trayanov and M. Slavtchova-Bojkova. Crump–Mode–Jagers branching processes: Application in population projections. *Advanced research in mathematics and computer science, Proceedings of the doctoral conference in mathematics, informatics and education, Sofia, Bulgaria, 19-20.09.2013*, pages 1–8, September 2013.

[62] P. Trayanov and M. Slavtchova-Bojkova. Branching processes: Forecasting human population. *SMTDA2014 Conference Proceedings, Lisbon, Portugal, 11-14.06.2014*, pages 141–149, 2014.

[63] P. Trayanov and M. Slavtchova-Bojkova. Estimating the effect of economic crisis with Crump–Mode–Jagers branching process. *Pliska Stud. Math. Bulgar.*, 24:99–110, 2015.

[64] H. Watson and F. Galton. On the probability of the extinction of families. *Journal of the Anthropological Institute of Great Britain*, 4:138–144, 1875.

[65] J. Wilmoth, K. Andreev, D. Jdanov, and D. A. Glei. Methods protocol for the human mortality database. `http://www.mortality.org`, May 2008.

[66] Rob J Hyndman with contributions from Heather Booth, Leonie Tickle, and John Maindonald. *Demography: Forecasting mortality, fertility, migration and population data*, 2014. R package version 1.18.

[67] M. Xie. On the solution of renewal-type integral equations. *Communications in Statistics - Simulation and Computation*, 18(1):281–293, 1989.

[68] A. Yakovlev and N. Yanev. *Transient Processes in Cell Proliferation Kinetics.* Lecture Notes in Biomathematics, v.82, Springer, 1989.

[69] A. Yakovlev and N. Yanev. Branching stochastic processes with immigration in analysis of renewing cell populations. *Math. Biosci.*, 203(1):37–63, 2006.

[70] A. Yakovlev and N. Yanev. Branching populations of cells bearing a continuous label. *Pliska Stud. Math. Bulg.*, 18:387–400, 2007.

[71] N. Yanev. Branching processes in cell proliferation kinetics. *Ch.12 in: Workshop on Branching Processes and Their Application, Lecture Notes in Statistics – Proceedings, M. Gonzales, M. Molina, I. del Puerto, M. Mota, R. Martinez, A. Ramos (Editors), Springer*, pages 159–178, 2010.

[72] N. Yanev. In memory of N. A. Dmitriev. *Pliska Stud. Math. Bulgar.*, 24:13–20, 2015.