

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ»**

Факультет прикладної математики

Кафедра прикладної математики

«На правах рукопису»
УДК 314.1:519.2

«До захисту допущено»

Завідувач кафедри
_____ О. Р. Чертов
(підпис)

« ____ » _____ 2015 р.

Магістерська дисертація

на здобуття ступеня магістра

зі спеціальності 8.04030101 «Прикладна математика»

на тему: ПРОГНОЗУВАННЯ ОЧІКУВАНОЇ ТРИВАЛОСТІ ЖИТТЯ ЗАЛЕЖНО
ВІД НАЙБІЛЬШ ВПЛИВОВИХ ФАКТОРІВ

Виконала: студентка 2 курсу, групи КМ-31М +

Савчук Олена Борисівна

_____ (підпис)

Науковий керівник

доцент, канд. техн. наук, доцент
Олефір О. С.

_____ (підпис)

Консультант із
нормоконтролю

старший викладач Мальчиков В. В.

_____ (підпис)

Рецензент

доцент, канд. фіз.-мат. наук, доцент
Шубенкова І.А.

_____ (підпис)

Засвідчую, що у цій магістерській
дисертації немає запозичень з праць
інших авторів без відповідних посилань.

Студент _____
(підпис)

Київ – 2015 року

**Національний технічний університет України
«Київський політехнічний інститут»**

Факультет прикладної математики

Кафедра прикладної математики

Рівень вищої освіти – другий (магістерський)

Спеціальність 8.04030101 «Прикладна математика»

ЗАТВЕРДЖУЮ

Завідувач кафедри

_____ О.Р. Чертов
(підпис)

«__» _____ 2015 р.

ЗАВДАННЯ

на магістерську дисертацію студентці

Савчук Олені Борисівні

1. Тема дисертації: «ПРОГНОЗУВАННЯ ОЧІКУВАНОЇ ТРИВАЛОСТІ ЖИТТЯ ЗАЛЕЖНО ВІД НАЙБІЛЬШ ВПЛИВОВИХ ФАКТОРІВ», науковий керівник дисертації Олефір Олександр Степанович, канд. техн. наук, доцент,
затверджені наказом по університету від «20» березня 2015 р. № 785-С
2. Термін подання студенткою дисертації «18» червня 2015 р.
3. Об'єкт дослідження: процес смертності населення України.
4. Предмет дослідження: система кількісних та якісних показників, що характеризують процес смертності населення в Україні залежно від найбільш впливових факторів.
5. Перелік завдань, які потрібно розробити:
 - провести аналіз існуючих моделей прогнозування очікуваної тривалості життя в залежності від найбільш впливових факторів в Україні та закордоном ;

- визначити найвагоміші фактори впливу на показник очікуваної тривалості життя населення України;
- розробити математичну модель прогнозування показників очікуваної тривалості життя населення України в залежності від визначених факторів на заданий період з урахуванням статевовікового розподілу;
- розробити програмну реалізацію побудованої моделі відтворення населення;
- отримати прогнозовані значення очікуваної тривалості життя населення України до 2020 року.

6. Орієнтовний перелік ілюстративного матеріалу:

- теоретичні засади методу;
- демонстраційні таблиці демографічних статистичних даних;
- таблиці з результатами експериментів;
- схема алгоритму запропонованого методу;
- структурна схема розробленої системи;
- демонстраційні приклади роботи методу.

7. Орієнтовний перелік публікацій

- VI наукова конференція магістрантів та аспірантів «Прикладна математика та комп'ютинг – ПМК'2015»;
- 17-та Міжнародна конференція «SAIT 2015».

8. Консультанти розділів дисертації

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Нормконтроль	Мальчиков В. В., старший викладач		

9. Дата видачі завдання «25» жовтня 2013 р.

10. Календарний план

№ з/п	Назва етапів виконання магістерської дисертації	Термін виконання етапів магістерської дисертації	Примітка
1	Огляд літератури та збір даних	01.01.2014	
2	Проведення порівняльного аналізу математичних методів демографічного прогнозування	01.03.2014	
3	Вибір, обґрунтування та опанування методів розв'язку задачі	01.05.2014	
4	Підготовка матеріалів першого розділу дисертації	01.09.2014	
5	Визначення найбільш впливових факторів на показник очікуваної тривалості життя	10.10.2014	
6	Розробка математичної моделі прогнозування очікуваної тривалості життя	15.11.2014	
7	Підготовка матеріалів другого розділу дисертації	09.01.2015	
8	Проектування архітектури розроблюваних програмних засобів	01.02.2015	
9	Розробка алгоритмів прогнозування	10.03.2015	
10	Програмна реалізація	10.04.2015	
11	Тестування програмного продукту та дослідження ефективності його роботи	12.05.2015	
12	Завершення оформлення магістерської дисертації	17.06.2015	

Студент

(підпис)

О. Б. Савчук

Науковий керівник дисертації

(підпис)

О. С. Олефір

РЕФЕРАТ

Актуальність теми. В сучасних умовах розвитку України все більше уваги приділяється повноті та достовірності даних державної статистики, оскільки реформування економіки та постійний пошук нових більш ефективних методів науково–обґрунтованої підтримки управлінських рішень вимагає наявності надійної статистичної інформації.

Демографічна ситуація в Україні викликає занепокоєння, для нашої країни характерне скорочення зниження тривалості життя. Показник очікуваної тривалості життя населення є одним з основних критеріїв, що характеризують якість життя та рівень здоров'я нації.

Статистика фіксує склад померлих за різними демографічними і соціальними ознаками, а це дає змогу проводити факторний аналіз показників смертності, оскільки в наш час гостро постає проблема визначення впливу того чи іншого фактору на нього. Особливо це стосується факторів смертності, яких можна уникнути або покращити (виліковні захворювання, алкогольна залежність і т.д.).

Метою дослідження є аналіз впливу природніх, соціальних і економічних факторів, розробка моделі залежності та прогнозування очікуваної тривалості життя на основі отриманої моделі.

Об'єкт дослідження – процес смертності населення України.

Предмет дослідження – система кількісних та якісних показників, що характеризують процес смертності населення в Україні залежно від найбільш впливових факторів.

Методи дослідження. В роботі використовуються методи математичного моделювання, множинного регресійного аналізу демографічного прогнозування.

Наукова новизна роботи полягає в наступному:

- Розроблено нову математичну модель прогнозування очікуваної тривалості життя населення України, яка враховує вплив різних соціально-економічних факторів.
- На базі розробленої програмної реалізації запропонованої моделі прогнозування очікуваної тривалості життя населення України та отримано прогнозні оцінки показника очікуваної тривалості життя на період до 2020 року.

Практична цінність результатів, отриманих під час роботи полягає в тому, що розроблені модель і ПЗ надає змогу зменшити затрати часу та покращити роботу спеціалістів-демографів, а також можуть бути основою для прийняття управлінських рішень для покращення демографічної та соціальної ситуації в Україні.

Наукова значимість роботи підтверджується тим, що результати впроваджено у відділі демографічного моделювання та прогнозування Інституту демографії та соціальних досліджень імені М. В. Птухи НАН України.

Апробація роботи. Основні положення і результати доповідались на VII науковій конференції магістрантів та аспірантів «Прикладна математика та комп'ютеринг» ПМК – 2015 (Київ, 15 – 17 квітня 2015 р.). Також результати роботи були опубліковані у збірнику матеріалів 17-ї Міжнародної науково-технічної конференції «Системний аналіз та інформаційні технології» SAIT 2015, Київ, 22-25 червня 2015 р.

Структура та обсяг роботи. Магістерська дисертація складається з вступу, п'яти розділів, висновків та додатків.

У вступі подано загальну характеристику роботи, виконано оцінку сучасного стану проблеми, обґрунтовано актуальність напрямку досліджень.

У першому розділі проведено огляд предметної області та аналіз існуючих програмних рішень. Також розглянуто математичні методи, які можуть використовуватись при прогнозуванні очікуваної тривалості

життя, проведено аналіз моделей на основі виділення їх основних характеристик.

У другому розділі наведено теоретичне обґрунтування та аналітичне представлення моделі, що використовується.

У третьому описано проектування розроблюваних програмних засобів, формати вхідних та вихідних даних, користувацького інтерфейсу та керівництво користувача.

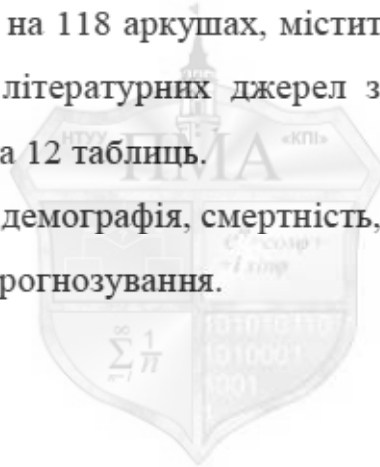
У четвертому розділі представлені опис алгоритму, програм, експериментальних даних та результати проведених експериментів.

У висновках проаналізовано отримані результати роботи.

У додатках наведено лістинг програми та екрані форми.

Робота виконана на 118 аркушах, містить 2 додатки та посилання на список використаних літературних джерел з 27 найменувань. У роботі наведено 21 рисунок та 12 таблиць.

Ключові слова: демографія, смертність, очікувана тривалість життя, математична модель, прогнозування.



РЕФЕРАТ

Актуальность темы. В современных условиях развития Украины все больше внимания уделяется полноте и достоверности данных государственной статистики, поскольку реформирование экономики и постоянный поиск новых более эффективных методов научно-обоснованной поддержки управленческих решений требует наличия надежной статистической информации.

Демографическая ситуация в Украине вызывает беспокойство, для нашей страны характерно снижение продолжительности жизни. Показатель ожидаемой продолжительности жизни населения является одним из основных критериев, характеризующих качество жизни и уровень здоровья нации.

Статистика фиксирует состав умерших по разным демографическим и социальным признакам, а это дает возможность проводить факторный анализ показателей смертности, поскольку в наше время остро стоит проблема определения влияния того или иного фактора на него. Особенно это касается факторов смертности, которых можно избежать или улучшить (излечимые заболевания, алкогольная зависимость и т.д.).

Целью исследования является анализ влияния природных, социальных и экономических факторов, разработка модели зависимости и прогнозирования ожидаемой продолжительности жизни на основе полученной модели.

Объект исследования – процесс смертности населения Украины.

Предмет исследования – система количественных и качественных показателей, характеризующих процесс смертности населения в Украине в зависимости от наиболее влиятельных факторов.

Методы исследования. В работе используются методы математического моделирования, множественного регрессионного анализа демографического прогнозирования.

Научная новизна работы заключается в следующем:

- Разработана новая математическая модель прогнозирования ожидаемой продолжительности жизни населения Украины, которая учитывает влияние различных социально-экономических факторов;
- На базе разработанной программной реализации предложенной модели прогнозирования ожидаемой продолжительности жизни населения Украины и получены прогнозные оценки показателя ожидаемой продолжительности жизни на период до 2025 года.

Практическая ценность результатов, полученных во время работы заключается в том, что разработанные модели и ПО дает возможность уменьшить затраты времени и улучшить работу специалистов-демографов, а также могут служить основой для принятия управленческих решений для улучшения демографической и социальной ситуации в Украине.

Научная значимость работы подтверждается тем, что результаты работы внедрены в отделе демографического моделирования и прогнозирования Института демографии и социальных исследований имени М. В. Птухи НАН Украины.

Апробация работы. Основные положения и результаты докладывались на VII научной конференции магистрантов и аспирантов «Прикладная математика и компьютеринг» ПМК - 2015 (Киев, 15 - 17 апреля 2015). Также результаты работы были опубликованы в сборнике материалов 17-й Международной научно-технической конференции «Системный анализ и информационные технологии» SAIT 2015, Киев, 22-25 июня 2015.

Структура и объем работы. Магистерская диссертация состоит из введения, пяти глав, заключения и приложений.

Во введении представлена общая характеристика работы, выполнена оценка современного состояния проблемы, обоснована актуальность направления исследований.

В первом разделе проведен обзор предметной области и анализ существующих программных решений. Также рассмотрены математические методы, которые могут быть использованы при прогнозировании ожидаемой продолжительности жизни населения, проведен анализ моделей на основе выделения их основных характеристик.

Во втором разделе приведено теоретическое обоснование и аналитическое представление модели, используется.

В третьем разделе описано проектирование разрабатываемых программных средств, в том числе его функциональное назначение и архитектуру, форматы входных и выходных данных, пользовательского интерфейса и руководство пользователя. Описаны требования к программным и техническим средствам.

В четвертом разделе представлены описание алгоритма, программ, экспериментальных данных и результаты проведенных экспериментов.

В выводах проанализированы полученные результаты работы.

В приложениях приведены листинг программы и экраны формы.

Работа выполнена на 118 листах, содержит 2 приложения и ссылки на список использованных литературных источников из 27 наименований. В работе приведены 21 рисунок и 12 таблиц.

Ключевые слова: демография, смертность, ожидаемая продолжительность жизни, математическая модель, прогнозирование.

ABSTRACT

Theme urgency. In modern conditions of Ukraine, more attention is paid to the completeness and reliability of government statistics, since the economic reform and constant search for new and more effective methods of scientific and reasonable support management decisions requires reliable statistical information.

The current demographic situation in Ukraine is of concern because, unlike the economically developed countries, our country is characterized declining population, reduced life expectancy and fertility. Life expectancy of the population is one of the main criteria that characterize the quality of living and the health of the nation.

Statistics captures the structure of deaths in different demographic and social characteristics, and it allows to carry out factor analysis of mortality because nowadays the acute problem of determining the effect of a factor on it. This is especially true factors of mortality, avoidable or improve (treatable disease, alcohol dependence, etc.).

Research objective is to analyze the impact of natural, social and economic factors, the development of models and forecasting depends life expectancy based on the received model.

Research objective is a process of mortality Ukraine.

The subject of study – system of quantitative and qualitative indicators of the process of mortality in Ukraine based on the most influential factors.

Research methods. In this paper, using methods of mathematical modeling, multiple regression analysis of demographic forecasting.

Scientific novelty

- A new mathematical model predicting life expectancy of the population of Ukraine, which includes the effect of various socio-economic factors;

- On the basis of the developed software implementation of the proposed model predicting life expectancy of the population of Ukraine and obtained estimates of life expectancy for the period until 2025.

The practical value of the results obtained during the work is developed model and software to predict life expectancy based on the most influential factors allows to reduce costs and improve time job specialists demographic research, and can be the basis for management decisions to improve the demographic and social situation in Ukraine.

The scientific importance of the work confirmed that the results have been implemented in the department of demographic modeling and forecasting of Ptoukha Institute for Demography and Social Studies of the National Academy of Sciences of Ukraine for forecasting the future and in retrospect the dynamics of population.

Approbation. The basic points and outcomes of the research have been presented and discussed at the VII scientific conference of graduate and post-graduate "Applied mathematics and computing" PMK - 2015 (Kyiv, 15 - 17 April 2015). Also, the results have been published in the in a collection of abstracts of the 17th International scientific-technical conference "System analysis and information technologies" SAIT 2015, Kyiv, 22-25 June 2015.

The structure and volume of work. Master's thesis consists of an introduction, five chapters, conclusions and applications.

The introduction presents the general characteristics of the work, the estimation of the current state of the problem, the urgency of research direction.

The first section of the survey of the subject area and analysis of existing software solutions. Also the mathematical methods that can be used in predicting life expectancy of the population, the analysis of models based on the allocation of their basic characteristics.

The second section provides a theoretical basis and analytical representation model used.

The third section describes the design developed software, including its functionality and architecture, formats of input and output data, the user interface and user manual. Described requirements for software and hardware.

The fourth section presents description of the algorithm, programs, experimental data and results of experiments.

The conclusions contain the results of analyzes.

The appendix provides program listings and screen forms.

Work performed at 118 pages, contains 2 applications and list of literature used 27 sources names. This paper gives 21 figures and 12 tables.

Keywords: demography, mortality, life expectancy, mathematical model forecasting.



ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, СКОРОЧЕНЬ І ТЕРМІНІВ	16
ВСТУП	17
ПОСТАНОВКА ЗАДАЧІ.....	20
1 ОГЛЯД І ПОРІВНЯЛЬНИЙ АНАЛІЗ ІСНУЮЧИХ МЕТОДІВ ПРОГНОЗУВАННЯ ОЧІКУВАНОЇ ТРИВАЛОСТІ ЖИТТЯ	21
1.1 Опис предметної області	21
1.2 Аналіз існуючих програмних рішень у даній галузі	26
1.3 Аналіз математичного апарату та існуючих математичних методів. 31	
1.3.1 Множинний регресійний аналіз	33
1.3.2 Модель Брасса	37
1.3.3 Модель Лі-Картера.....	39
1.3.4 Порівняння найпопулярніших моделей.....	42
Висновки до розділу 1	43
2 ОПИС МОДЕЛІ, ЩО ВИКОРИСТОВУЄТЬСЯ	44
2.1 Теоретичне обґрунтування.....	44
2.2 Аналітичне представлення моделі	47
2.2.1 Побудова таблиць дожиття	48
2.2.2 Модель Лі-Картера.....	54
2.2.3 Множинний регресійний аналіз	59
Висновки до розділу 2	63
3 ПРОЕКТУВАННЯ ПРОГРАМНИХ ЗАСОБІВ.....	65
3.1 Функціональне призначення програмного забезпечення	65

3.2 Архітектура програмного забезпечення	65
3.3 Формати вхідних та вихідних даних	67
3.4 Вимоги до програмних та технічних засобів	69
3.5 Опис роботи користувача з системою	69
Висновки до розділу 3	73
4 ПРОГРАМНА РЕАЛІЗАЦІЯ.....	74
4.1 Опис розробленого алгоритму.....	74
4.2 Опис експериментальних даних	75
4.3 Результати проведених експериментів	82
4.3.1 Оцінка якості моделі.....	82
4.3.2 Прогнозування очікуваної тривалості життя населення України	86
Висновки до розділу 4	93
ВИСНОВКИ.....	94
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	96
ДОДАТКИ.....	99
Додаток А. Лістинг програми	99
Додаток В. Ілюстративний матеріал	117

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, СКОРОЧЕНЬ І ТЕРМІНІВ

МНК – Метод найменших квадратів.

ПЗ – програмне забезпечення.

ПЕОМ – персональна електронно-обчислювальна машина.

СЛАР – система лінійних алгебраїчних рівнянь.



ВСТУП

Населення, його соціальне самопочуття та розвиток є визначальною ознакою суспільства, основою могутності держави. Тому демографічна сфера завжди є об'єктом пильної уваги політиків та громадськості. Розвиток суспільства, насамперед, має передбачати благополуччя населення, зокрема відтворення та поліпшення його базової властивості – здоров'я, що є запорукою створення матеріальних та духовних благ [1].

Сучасна демографічна ситуація в Україні викликає занепокоєння, адже, на відміну від економічно-розвинутих держав, для нашої країни характерне скорочення чисельності населення, зниження тривалості життя і народжуваності. Саме тому в нинішніх умовах все більше уваги приділяється достовірності даних державної статистики, оскільки пошук ефективних методів реформування соціальної політики та економіки вимагає наявності надійної статистичної інформації

Відповідно до Першого всеукраїнського перепису населення [2], станом на 5 грудня 2001р. в Україні проживало 48,5 млн. мешканців. Проте, нині їх стало ще менше. Станом на 1 січня 2011р. чисельність наявного населення України, за попередніми даними, становила вже 45,8 млн. осіб. Якщо порівнювати це число з 1993р., роком коли українців було найбільше за всю післявоєнну історію (52 мільйони), то стає очевидним, що чисельність населення України скоротилась на 6,2 мільйони. Головною причиною цього є депопуляція, тобто природне скорочення чисельності населення внаслідок щорічного перевищення кількості померлих над кількістю народжених.

Очікувана тривалість життя – це показник, який одним числом передає багато інформації. По суті, він показує, скільки років, ймовірно, проживе громадянин деякого віку певної країни. Але цей показник втілює значно більше. В соціальних та економічних дослідженнях очікувана

тривалість життя часто використовується як міра якості життя населення, суспільного здоров'я нації та розвитку науково-технічного прогресу. Саме тому, проблема оцінки та прогнозування очікуваної тривалості життя та показників смертності стоїть дуже гостро.

Хоча може здаватись, що на показник тривалості життя найбільший вплив мають біологічні чинники, проте, цей показник проявляє пряму безпосередню залежність від рівня соціально-економічного та культурного розвитку суспільства, які, в свою чергу, впливають на рівень добробуту, розвиток системи охорони здоров'я та медицини, тощо.

Тривалість життя населення визначається головним чином впливом ряду об'єктивних причин. Статистика фіксує склад померлих за різними демографічними і соціальними ознаками, а це дає змогу проводити факторний аналіз показників смертності, оскільки в наш час гостро постає проблема визначення впливу того чи іншого фактору на нього. Особливо це стосується факторів смертності, яких можна уникнути або покращити (виліковні захворювання, алкогольна залежність і т.д.).

На основі цього з точністю можна сказати, що повсякчас зростає необхідність вироблення адекватних заходів для покращення демографічної ситуації в країні. Разом з тим зростає роль наявності точної та актуальної демографічної інформації, а також її досліджень методами математичного моделювання.

В даний час важко переоцінити роль демографічних прогнозів в різних галузях економіки, освіти, охорони здоров'я тощо. Тому виникає необхідність побудови науково обґрунтованих демографічних прогнозів, як на найближче, так і на більш віддалене майбутнє.

Задачі прогнозування і удосконалення системи демографічної політики в області здоров'я населення тісно пов'язані між собою і не можуть розглядатись ізольовано. Варто враховувати, що при сучасному рівні розвитку демографічної науки прогноз очікуваної тривалості життя населення був і залишається науковою задачею, що потребує кожного разу

проведення спеціальних наукових досліджень. Він не можливий без глибокого осмислення закономірностей історичної еволюції цього процесу, соціальних факторів і передумов збільшення тривалості життя [1].

Протягом тисячоліть історії людства при майже повсюдно високому рівні народжуваності смертність змінювалася в широких межах та фактично визначала рівень природного приросту. Зовсім не випадково першим великим демографічним дослідженням у вітчизняній науці стала робота М.В. Птухи присвячена аналізу смертності серед основних етнічних груп європейської Росії у кінці XIX століття. Та хоча смертність – перший з демографічних процесів, який був підданий статистичному дослідженню, він і досі залишається не вивченим до кінця. Проблеми вивчення, моделювання та прогнозування смертності досліджувалися в роботах У. Брасса, Лі Картера, Р. Лі [3], Е.М. Андреева та багатьох інших. Вклад у дослідження смертності зробили і сучасні українські вчені на чолі з Е.М. Лібановою. Також потрібно відзначити, що прогнозування очікуваної тривалості життя, основане на аналізі динаміки і причин смертності досліджено в роботах Дж. Полларда, а також В.І. Стародубова і А.Е. Іванової.

Незважаючи на значущість і цінність здобутків вищезгаданих авторів, теоретико-методологічні й прикладні засади дослідження явищ смертності у демографічному вимірі є недостатньо розробленими. Проведений аналіз існуючих комп'ютерних систем в галузі прогнозування смертності та чисельності населення дав наступні результати: вітчизняні розробки в цій області практично відсутні, в основному, використовуються закордонні системи, які були «адаптовані» і впроваджені на території нашої держави.

Отже, створення нових, більш точних методів та методик прогнозування очікуваної тривалості життя населення України є важливою науково-практичною задачею.

ПОСТАНОВКА ЗАДАЧІ

Об'єкт дослідження – процес смертності населення України.

Предмет дослідження – система кількісних та якісних показників, що характеризують процес смертності населення в Україні залежно від найбільш впливових факторів.

Метою роботи є аналіз впливу природних, соціальних і економічних факторів на очікувану тривалість життя (смертність) населення України, визначення найбільш вагомих із них та прогнозування очікуваної тривалості життя населення України в залежності від обраних факторів на заданий період з урахування статево-вікового розподілу

Для досягнення даної цілі сформовано задачі:

- провести аналіз існуючих моделей прогнозування очікуваної тривалості життя в залежності від найбільш впливових факторів в Україні та закордоном ;
- визначити найвагоміші фактори впливу на показник очікуваної тривалості життя населення України;
- розробити математичну модель прогнозування показників очікуваної тривалості життя населення України в залежності від визначених факторів на заданий період з урахуванням статево-вікового розподілу;
- розробити програмну реалізацію побудованої моделі відтворення населення;
- отримати прогнозовані значення очікуваної тривалості життя населення України до 2020 року.

1 ОГЛЯД І ПОРІВНЯЛЬНИЙ АНАЛІЗ ІСНУЮЧИХ МЕТОДІВ ПРОГНОЗУВАННЯ ОЧІКУВАНОЇ ТРИВАЛОСТІ ЖИТТЯ

1.1 Опис предметної області

Демографія (грец. $\delta\eta\mu\omicron\varsigma$ — народ, грец. $\gamma\rho\acute{\alpha}\phi\omega$ — пишу) — наука про закономірності відтворення населення, про залежність його характеру від соціально-економічних, природних умов, міграції. Вивчає чисельність, територіальне розміщення і склад населення, їх зміни, причини і наслідки цих змін і дає рекомендації щодо їх поліпшення [3].

Дуже часто поняття «населення» розуміють як сукупність людей, що проживає на певній території. Такого розуміння, що виділяє в якості атрибутивних дві основні ознаки — кількісну (те, що населення — це сукупність) і територіальну (те, що населення пов'язано з певною територією), цілком достатньо для більшості наук, для яких населення — об'єкт дослідження.

Однак для демографії таке розуміння є недостатнім. Воно відображає важливий, але лише поверхневий шар визначення населення як поняття демографічної науки, те, з чого, власне, починається його вивчення, що є вихідним пунктом демографічного аналізу. А нас воно цікавить, як соціальний процес який є винятковим предметом демографії як науки, тобто, з точки зору його відтворення.

Основною атрибутивною ознакою є не територіальна (населення — сукупність людей, що живуть на тій чи іншій території), не кількісна (населення — сукупність) і не будь-яка інша, а здатність до постійного самооновлення через процес зміни поколінь, тобто через народжуваність і смертність — здатність до самовідтворення.

Відтворення населення — це історично і соціально економічно обумовлений процес постійного і безперервного поновлення людських поколінь. Це заміна старших, поступово вимираючих поколінь, новими.

Процес самовідтворення населення характеризується двома основними показниками: інтенсивністю вимирання старших поколінь (смертність) і інтенсивністю їх заміни новими (народжуваність). Перший — визначається смертністю для кожного віку, і характеризується показниками таблиць смертності (дожиття). Другий — інтенсивність заміни старших поколінь новими, що визначається рівнем народжуваності для кожного віку матері.

Смертність (Mortality) — масовий статистичний процес вимирання покоління або населення, складаний з багатьох одиничних смертей, що настають в різному віці і визначають у своїй сукупності порядок вимирання покоління [3].

Смерть є однією із двох складових процесу природного руху населення. Зменшення смертності має вплив на підвищення природного, як наслідок, загального приросту населення. Рівень смертності — це одна із характеристик економічного і соціального здоров'я країни [4].

На будь-який демографічний показник, впливає не один, а декілька факторів і показник очікуваної тривалості життя не є винятком. Причини смертності прийнято поділяти на дві основні групи: природні (ті, яких неможливо запобігти) та привнесені (ті, яких можна уникнути). До першої групи входять смерті від різноманітних спадкових аномалій та старіння організму. До другої групи прийнято зараховувати смерть від хвороб, різноманітні види насильницьких смертей (самогубства, вбивства, геноцид, війни тощо).

Конкретна дія тих чи інших безпосередніх причин смерті, їх взаємозв'язок та вплив на вікові і загальні показники смертності залежить від цілої низки чинників. Тут виділяються: стан здоров'я, освіта, умови праці, зайнятість та фах, прибутки, видатки та накопичення, умови

транспорту, житлові умови, розваги і відпочинок, соціальне страхування, свобода особистості, професійні захворювання, алкоголізм, наркоманія, охорона здоров'я, рівень урбанізації. Значний вплив на захворюваність, а через неї і на смертність справляють природно-екологічні чинники, що визначають умови життя людей та пов'язані з ними локальні хвороби, стихійні лиха тощо. Суттєве значення мають також і біологічні (фізіологічні та генетичні) фактори, що визначають особливості розвитку організму людини та його взаємодію з зовнішнім природним та соціальним середовищем.

Так як смертність підлягає статистичному аналізу, то доцільно дослідити зв'язки показників смертності і очікуваної тривалості життя в залежності від конкретних природних, економічних та соціальних чинників. З іншої сторони смертність, як статистичний процес підлягає не лише аналізу, а й моделюванню. Існує два основних підходи до розв'язку цієї задачі.

Один підхід полягає у відшуканні вигляду і параметрів математичної функції, що відбиває залежність тієї або іншої характеристики доживання від віку людини. Таким чином, з'являється можливість вивчення смертності на будь-яких, скільки завгодно коротких інтервалах часу, а також її прогнозування. Моделі доживання і смертності, реалізовані в рамках даного підходу, називають моделями з безперервним часом.

Другий підхід до моделювання законів смертності передбачає перерахування основних характеристик життя і смертності окремо в кожній віковій групі. Зібрані в єдиній таблиці ці характеристики утворюють так звану таблицю смертності населення (таблицю дожиття та середньої очікуваної тривалості життя).

Таблиці дожиття та середньої очікуваної тривалості життя є ймовірнісними моделями смертності, які використовуються для перспективних розрахунків чисельності та складу населення. Побудова

таблиць дожиття (смертності) здійснюється за даними поточного обліку та переписів населення країни .

У таблицях смертності немає жодного доживаючого або вмираючого. Взагалі – жодної людини. Є тільки ймовірності та частки. Послідовність змін чисел графічно являє собою лінію дожиття, що характеризує порядок вимирання покоління. Чим нижче рівень смертності, тим більша частка народжених (покоління) доживає до старшого [5].

Демографічний прогноз – це науково обґрунтоване передбачення основних параметрів руху населення та майбутньої демографічної ситуації: чисельності, віково-статевої та сімейної структур населення, народжуваності, смертності, міграції, якісних характеристик населення.

Тобто, демографічний прогноз – це розрахунок очікуваної чисельності та статево-вікової структури населення певної території на основі фактичної структури та існуючих або пропонованих рівнів народжуваності і смертності в різних віках, а також тенденцій міграції .

У демографічному прогнозі визначається також середня тривалість майбутнього життя, співвідношення між шлюбами і розлученнями, середній вік вступу в шлюб, розподіл населення на осіб у працездатному і непрацездатному віці, встановлюються терміни початку трудової діяльності та виходу на пенсію і так далі.

За тривалістю прогнозованого періоду прогнози поділяють на:

- короткострокові (5–10 років);
- середньострокові (15–45 років);
- довгострокові.

За цілями прогнозування демографічні прогнози поділяють на:

- аналітичні — вивчають реальну ситуацію, виявляють у ній проблемні моменти, збереження яких у майбутньому спричинить поглиблення існуючих та виникнення нових проблем;

- прогнози–попередження — вивчають та досліджують ті негативні наслідки, які у майбутньому можуть спричинити збереження та поглиблення вже існуючих проблемних;
- нормативні прогнози — розробляють конкретні рекомендації з метою досягнення певного бажаного стану демографічних процесів;
- функціональні прогнози — обґрунтовують прогнозовану інформацію щодо населення, яка потрібна для прийняття рішень в економічній, соціальній, політичній та інших сферах управління.

Порогом розрахунку називають момент часу, на який була визначена початкова вікова структура.

Кроком розрахунку слугує часовий інтервал, на який пересуваються окремі вікові групи (звичайно він буває однорічним чи п'ятирічним).

Горизонтом прогнозування називають рік, до якого здійснюється прогноз [1].

Демографічні прогнози є важливим елементом комплексного довгострокового соціально-економічного планування.

Задачі прогнозування смертності і удосконалення системи міри демографічної політики в області здоров'я населення тісно пов'язані між собою і не можуть розглядатись ізольовано. Побудова моделі з великою кількістю факторів, а також визначення впливу кожного з них окремо та в сукупності на модельований показник є важливою задачею.

Варто враховувати, що при сучасному рівні розвитку демографічної науки прогноз очікуваної тривалості життя був і залишається науковою задачею, що потребує кожного разу проведення спеціальних наукових досліджень. Він не можливий без глибокого осмислення закономірностей історичної еволюції цього процесу, соціальних факторів і передумов збільшення тривалості життя.

Без попереднього демографічного прогнозу неможливо уявити перспективи виробництва та споживання товарів і послуг, житлового

будівництва, розвитку соціальної інфраструктури, охорони здоров'я і освіти, пенсійної системи, рішення геополітичних проблем і т.д. [1].

1.2 Аналіз існуючих програмних рішень у даній галузі

Огляд існуючих систем демографічного прогнозування очікуваної тривалості життя показав, що на практиці застосовуються системи закордонного виробництва, які було акліматизовано до потреб нашої держави. Вітчизняні розробки практично відсутні і доволі застарілі.

Нижче наведено декілька програм, що застосовуються для демографічного прогнозування, зокрема прогнозування очікуваної тривалості життя:

- SAS/ETS and SAS High-Performance Forecasting [6];
- Статистичний пакет «STATISTICA» [7];
- GAUSS Mathematical & Statistical System [8];
- «Программа демографического прогнозирования» [9];
- Система моделювання «Спектрум» (розробка США) [10];
- Population Site: Software: FIVFIV [11].

В таблиці 1.1 представлений порівняльний аналіз перелічених методів:

Таблиця 1.1 – Порівняння існуючих програмних рішень

Характеристики	Країна - розробник	Тривалість прогнозованого періоду	Переваги	Недоліки
Система				
SAS/ETS	США	довгострокова	<ul style="list-style-type: none"> - проводить діагностику часового ряду; - створює список можливих моделей для прогнозування; - генерує прогноз для кожної моделі-кандидата. 	<ul style="list-style-type: none"> - Висока вартість. (Ціна однієї ліцензії для комерційних структур близько 6000 євро); - Складність у використанні.
Статистичний пакет «STATISTICA»	США	довгострокова	<ul style="list-style-type: none"> - процедури системи мають високу швидкість і точність обчислень; - реалізовано широкий набір методів опису, побудови моделей, декомпозиції та прогнозування часових рядів. 	<ul style="list-style-type: none"> - відсутність реалізації деяких важливих тестів часових рядів (зокрема - тестів на стаціонарність); - застосування пакета вимагає високої теоретичної підготовки в теорії ймовірності і математичній статистиці.
GAUSS Mathematical & Statistical System	США	довгострокова	<ul style="list-style-type: none"> - містить повний набір класичних методів аналізу даних (зокрема багатофакторного); 	<ul style="list-style-type: none"> - Висока вартість (ціна за пакет ≈10 тис грн.);
Программа демографического прогнозирования	Росія	середньострокова (15-40 років)	<ul style="list-style-type: none"> - графічне відображення динаміки основних вихідних демографічних показників для прогнозного населення 	<ul style="list-style-type: none"> - об'єм вхідних даних - обов'язкова наявність MS Excel 97 і вище
Система моделювання «Спектрум»	США	довгострокова	<ul style="list-style-type: none"> - період прогнозування; - можливість використання даних іншими модулями системи 	<ul style="list-style-type: none"> - складність вибору регіону прогнозуван.; - неможливо побудувати прогноз для маленької терит.; - непрацездатність при змін. початкових даних

Розглянемо більш детально деякі з систем:

SAS/ETS and SAS High-Performance Forecasting

Компанія S.A.S. (заснована 1976 р. в США), створює ПЗ для широкого застосування в бізнесі, науці та соціальній політиці. Зокрема, SAS/ETS and SAS High-Performance Forecasting найчастіше використовуються для проведення економічного аналізу, прогнозування, економічного та фінансового моделювання, аналізу часових рядів, фінансової звітності та перетвореннями даних часових рядів. ПЗ включає в себе наступні автоматичні процеси прогнозування:

- проводить діагностику часового ряду з використанням методів аналізу часових рядів;
- на основі аналізу, створює список можливих моделей (моделей-кандидатів);
- генерує прогноз для кожної моделі-кандидата;
- оцінює прогнози та рахує похибку для кожної моделі-кандидата;
- застосовуючи заданий критерій вибору моделі, вибирає найбільш підходящу;

Незважаючи на те, що система виконує великий діапазон аналітичних та статистичних процедур вона має два суттєві недоліки:

- Висока вартість. Ціна однієї ліцензії для комерційних структур близько 6000 євро;
- Складність у використанні. Потрібна наявність професійних навичок і високої кваліфікації, широкої первісної статистичної освіти, доступної літератури і консультаційних служб [6].

Статистичний пакет «STATISTICA»

STATISTICA - програмний пакет для статистичного аналізу, розроблений компанією StatSoft, який реалізує функції аналізу даних, управління даних, видобутку даних, візуалізації даних із залученням статистичних методів. STATISTICA - це інструмент розробки

користувальницьких додатків в бізнесі, економіці, фінансах, промисловості, медицині, страхуванні та інших областях. STATISTICA легка в освоєнні і використанні.

Всі аналітичні інструменти, наявні в системі, доступні користувачеві і можуть бути обрані за допомогою альтернативного користувача інтерфейсу. Користувач може всебічно автоматизувати свою роботу, починаючи з застосування простих макросів для автоматизації рутинних дій аж до поглиблених проектів, що включають в тому числі інтеграцію системи з іншими додатками або Інтернет. Технологія автоматизації дозволяє навіть недосвідченому користувачу налаштувати систему на свій проект.

Система володіє наступними загально визнаними достоїнствами:

- процедури системи STATISTICA мають високу швидкість і точність обчислень;
- гнучка і потужна технологія доступу до даних дозволяє ефективно працювати як з таблицями даних на локальному диску, так і з віддаленими сховищами даних;
- містить повний набір класичних методів аналізу даних: від основних методів статистики до просунутих методів, що дозволяє гнучко організувати аналіз;
- в модулі «Часові ряди» реалізований широкий набір методів опису, побудови моделей, декомпозиції та прогнозування часових рядів як в часовій, так і в частотній області;
- має найрізноманітніші можливості для перегляду і графічного представлення одновимірних і багатовимірних рядів;
- можна аналізувати дуже довгі ряди (більше 100 тис. спостережень для комп'ютера з 8 Mb оперативної пам'яті);
- з багатовимірними рядами можна працювати в активній робочій області, їх можна переглядати і зіставляти один з одним;

- програма автоматично відзначає всі етапи аналізу часового ряду і зберігає повну історію перетворень і отримані результати .

Але система також має певні недоліки:

- застосування пакета вимагає високої теоретичної підготовки в теорії ймовірності і математичній статистиці;
- відсутність реалізації деяких важливих тестів часових рядів (зокрема - тестів на стаціонарність) [7].

Система моделювання «Спектрум»

Система політичних моделей «Спектрум» - п'ятирічний проект «Поліс», фінансований Управлінням міжнародного розвитку США. Система моделювання "Спектрум" консолідує попередні моделі в інтегрований пакет, що містить компоненти, наведені нижче:

Демографічна модель, що входить в систему «Спектрум», відома під назвою «ДемПродж», являє собою комп'ютерну програму складання прогнозів в області народонаселення для окремих країн і регіонів. Ця програма вимагає інформації про чисельність груп населення за віком і статтю в базовий рік, а також дані за поточний рік і майбутні припущення про сумарний коефіцієнт народжуваності, віковий розподіл народжуваності, очікуваної тривалості життя при народженні представників обох статей, про найбільш придатною модельної таблиці смертності, а також про масштаби і характер міжнародної міграції. Ця інформація використовується для прогнозування чисельності майбутнього населення за віком і статтю на перспективу до 150 років.

Недоліки:

- користувач не знайомий з методикою розрахунків, не може повною мірою бути впевненим у правильності обраної моделі – перш ніж вибрати регіон географічного розташування території для якої робиться прогноз, необхідно перевірити первинне розподіл народжуваності в країні або регіоні і порівняти з модельними таблицями ООН вікового розподілу народжуваності;

- типові моделі побудовані для занадто великих за довжиною територій, не можуть повною мірою відобразити все різноманіття та особливості демографічних процесів, розташованих на них держав, з різними рівнями життя і економічним розвитком;
- неможливо побудувати прогноз для маленької території (наприклад для Київської області);
- непрацездатність програми при змінному наборі початкових даних [10].

1.3 Аналіз математичного апарату та існуючих математичних методів

Прогнозування як процес є важливою ланкою, що зв'язує існуючі і постійно вдосконалюємі теоретичні моделі досліджуваної регіональної соціально-економічної системи з практичними знаннями і фактами історії її розвитку [12]. Саме тому воно стає одним з головних способів рівня і без побудови та вдосконалення математичної моделі, не можливе ефективне планування та формування стратегії розвитку соціально-економічної системи.

Для демографічного прогнозування використовуються різні методи прогнозування. У 1927 році В. А. Базаров-Руднев запропонував 3 методи прогнозу: екстраполяція, аналітична модель, експертиза. В даний час існує близько трьох десятків різних методів прогнозування, але найчастіше на практиці використовуються не більше 10, серед них: фактографічні (екстраполяція, інтерполяція, тренд-аналіз), експертні (в т.ч. опитування, анкетування), публікаційного (у т. ч. патентні), цитатно-індексні, сценарні, матричні, моделювання, аналогій, побудова графів і т.д. [12]. Також існує велика кількість класифікацій. Одні базуються на певних нюансах

конкретного об'єкту прогнозування, інші – являють собою набір базових методів і відрізняються тільки послідовністю та кількістю методів, що розглядаються.

Для прогнозування соціально-економічних та демографічних показників можна використовувати моделі та методи, що ґрунтуються на різних алгоритмах: якісні й кількісні методи, методи аналізу часових рядів та прийняття рішень, а також класичні методи штучного інтелекту (зокрема, генетичні алгоритми), метод комплексування аналогів, метод моделей Байєса (мережі довіри Байєса), регресійні авторегресійні та нелінійні багатofакторні моделі [12].

На рисунку 1.1 відображено чотири групи методів: якісні (що базуються на оцінках та судженнях експертів), кількісні (що базуються на оптимізації та математичному моделюванні), аналізу часових рядів (що використовують передісторію протікання певного процесу для його прогнозування) та прийняття рішень (що засновані на багатоцільовому аналізі, евристичних підходах, створенні баз даних, логічних методах моделювання, експертних системах). Кожна група використовує спеціальні методи, для яких розроблені певні моделі та алгоритми їх застосування.



Рисунок 1.1 – Класифікація методів прогнозування

Зупинимось більш детально на найчастіше вживаних методах та моделях математичного моделювання смертності та очікуваної тривалості життя:

1.3.1 Множинний регресійний аналіз

Один із найпопулярніших методів, що використовується при аналізі причин та їхнього впливу на показник смертності, а також при прогнозуванні очікуваної тривалості життя. Основною ціллю є побудова моделі, що враховує дію певної кількості факторів, а також визначення

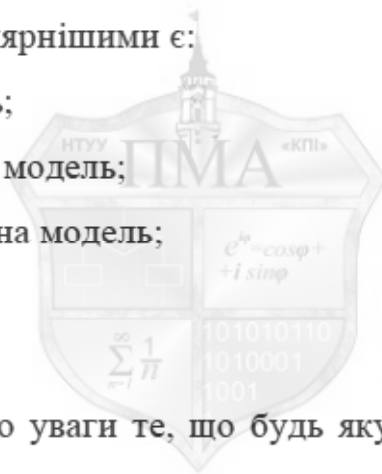
впливу кожного з них окремо та в сукупності на моделюємі показник. Так як множинний регресійний аналіз є розширенням парної регресії, то розглядається функція:

$$\hat{y} = f(x_1, x_2, \dots, x_p) \quad (1.1)$$

Математично завдання зводиться до знаходження аналітичного виразу, який якнайкраще відобразить зв'язок факторних ознак з результативною.

Фактично, форму зв'язку можна визначити шляхом перебору функцій різних типів, та це може суттєво збільшити кількість необхідних розрахунків. Найпопулярнішими є:

- лінійна модель;
- поліноміальна модель;
- експоненціальна модель;
- гіперболічна;
- логарифмічна.



Однак, беручи до уваги те, що будь яку функцію багатьох змінних шляхом логарифмування або заміни змінних можна звести до лінійного виду і з огляду на найбільш чітку інтерпретацію параметрів, найпопулярнішою із моделей множинної регресії є лінійна:

$$y = \alpha' + \beta_1' x_1 + \beta_2' x_2 + \dots + \beta_p' x_p + \varepsilon, \quad (1.2)$$

де: $\beta_1' = \frac{\partial y}{\partial x_1}, \beta_2' = \frac{\partial y}{\partial x_2}, \dots, \beta_p' = \frac{\partial y}{\partial x_p}$

Оцінка параметрів $\alpha', \beta_1', \beta_2', \dots, \beta_p'$ рівняння регресії, що представлено формулою (1.1) є одним із найважливіших завдань

множинного регресійного аналізу. Для його вирішення можна використовувати:

- метод найменших квадратів (МНК);
- матричний метод;
- скалярний метод;
- зведення до регресійної моделі в стандартизованому масштабі;
- використання частинних рівнянь регресії та інші.

Найпопулярнішим методом розв'язку цієї задачі є метод найменших квадратів (МНК).

На основі коефіцієнтів регресії не можна робити висновок, яка із факторних ознак найбільше впливає на результативну ознаку, тому що коефіцієнти регресії між собою не порівняльні, оскільки вони володіють різними одиницями вимірювання. З метою виявлення порівняльної сили впливу окремих факторів і резервів, які закладені в них, статистика вираховує часткові коефіцієнти еластичності ε_i , а також бета-коефіцієнти β_i за формулами:

$$\varepsilon_i = a_i \frac{\bar{x}_i}{\bar{y}}, \quad (1.3)$$

$$\beta_i = a_i \frac{\sigma_{xi}}{\sigma_y}, \quad (1.4)$$

де a_i – коефіцієнт регресії при i -му факторі;

\bar{x}_i – середнє значення i -го фактора;

\bar{y} – середнє значення результативної ознаки;

σ_{xi} – середнє квадратичне відхилення i -го фактора;

σ_y – середнє квадратичне відхилення результативної ознаки.

Часткові коефіцієнти еластичності показують, на скільки відсотків у середньому зміниться результативна ознака зі зміною на 1% кожного фактора за фіксованого положення інших факторів.

Для визначення факторів, у розвитку яких закладені найбільші резерви покращення досліджуваної ознаки, з урахуванням ступеня варіації факторів рівняння множинної регресії, враховують часткові коефіцієнти, які показують на яку частину середнього квадратичного відхилення змінюється результативна ознака зі зміною відповідної факторної ознаки на величину її середнього квадратичного відхилення.

Для характеристики ступеня тісноти зв'язку в множинній прямолінійній кореляції використовують множинний коефіцієнт кореляції:

$$R_{yx_1x_2\dots x_p} = \sqrt{1 - \frac{\sum(y_i - \bar{y})^2}{\sum(y_i - \bar{y})^2}}, \quad (1.5)$$

де $R_{yx_1x_2\dots x_p}$ – множинний коефіцієнт кореляції.

Множинний коефіцієнт кореляції показує, яку частину загальної кореляції складають коливання, під впливом факторів x_1, x_2, \dots, x_n , закладених у багатфакторну модель для дослідження. Він коливається в межах від 0 до ± 1 . При $R = 0$ зв'язок між досліджуваними ознаками відсутній, при $R = \pm 1$ – функціональний, причому, чим ближче він до ± 1 , тим тісніший зв'язок.

Безперечною перевагою множинного регресійного аналізу є його відносна простота. В наш час множинна регресія – один із найбільш популярних методів багатфакторного аналізу.

Проте, оскільки множинний регресійний аналіз є розвитком парного регресійного аналізу у випадках, коли результуюча(залежна) змінна пов'язана більш, ніж з одною незалежною змінною, з'являються два основних недоліки даного методу:

- перший недолік стосується впливу конкретної незалежної змінної на залежну змінну та розмежування її впливу на інші незалежні змінні;

- другим важливим недоліком є складність специфікації моделі, коли потрібно визначити, які фактори впливу потрібно включити в регресію, а які – ні [13].

1.3.2 Модель Брасса

Модель смертності Брасса – математична модель, що описує криву смертності. Розрахована англійським демографом У.Брассом в 1968 - 1971 рр. та широко відома в демографічній статистиці.

Область застосування моделі смертності Брасса традиційна:

- аналіз та прогноз смертності;
- побудова таблиць дожиття на основі неповної інформації;
- визначення якості обліку чисельності населення та числа смертей.

Модель має два параметра – вік та число доживаючих і основана на функції, що називається «логіт». Існує прямий та зворотній логіт. Перший має вигляд:

$$Y_x = \text{logit} (0p_x) = 0,5 \ln \left(\frac{0p_x}{0q_x} \right), \quad (1.6)$$

де $0p_x = \frac{l_x}{l_0}$ – ймовірність дожиття від 0 років до віку x ;

l_x – число доживаючих до віку x ;

l_0 – корінь таблиці смертності;

$$0q_x = 1 - 0p_x$$

Зворотній логіт $\text{logit} (0q_x)$, причому:

$$\text{logit} (0p_x) = - \text{logit} (0q_x) \quad (1.7)$$

У. Брас емпірично встановив, що логіт-трансформація шкал x і $0p_x$ приводить залежність між x і $0p_x$ до лінійної, тобто

$$Y_x = \alpha + \beta Y'_x, \quad (1.8)$$

де α та β – параметри;

Y_x – логіти досліджуємої таблиці смертності (трансформована вісь $0p_x$);

Y'_x – логіти стандартної таблиці дожиття (трансформована вісь віку).

Логіт – важливе поняття для побудови повних таблиць смертності. Для розрахунків за допомогою даної моделі потрібно мати щонайменше дві системи логітів: за той рік, для якого потрібно побудувати таблицю смертності, та систему логітів стандартної таблиці смертності.

В якості стандарту може бути використана будь-яка підходяща для цілей дослідження таблиця, проте, для широкого кола порівнянь зазвичай застосовується «загальний стандарт Брасса» або його модифікації. Наприклад, за стандартну таблицю може прийматися в тому разі, якщо різниці повікових логітів стандартної та першої лежать на прямій лінії або мало від неї відрізняються.

Параметр α рівняння (1.8) – один з індикаторів загального рівня смертності. Він показує, наскільки вище або нижче одна крива смертності знаходиться по відношенню до іншої. Параметр β – це нахил між кривими смертності. Він вказує на якісні відмінності між ними.

Існує ряд методів безпосередньої оцінки демографічних характеристик, розроблених на основі моделі Брасса. І хоча існує цілий ряд інших методів, простота моделі приваблює і її часто використовують при аналізі смертності [14].

1.3.3 Модель Лі-Картера

Часто поштовхом до удосконалення демографічних методів та моделей виступають практичні потреби суспільства, уряду, управлінської організації. Зокрема підвищена цікавість до прогнозу смертності супроводжувалась розвитком нових і більш складних методів. Наприклад в США наприкінці 1980-х – 1990-х помітно підвищилась смертність молодих чоловіків (віку 15 – 44 рр.) у зв'язку із епідемією СНІДу. Саме в цей період, а точніше 1992р., Р.Лі і Л.Р.Картер розробили та опублікували нову модель довгострокових прогнозів рівня та вікової структури смертності, заснований на комбінації статистичних методів часових рядів та простого підходу до розподілу померлих за віком. Головною ціллю якої було адекватно відобразити в прогнозі даний підйом смертності [15].

Модель описує побудову часових рядів вікових коефіцієнтів смертності як суму вікових коефіцієнтів компонент, які не залежать від часу та інших компонентів, що є результатом зміни параметру у часі, та відображає загальний рівень смертності і показує на скільки швидко змінюється рівень смертності для кожної вікової категорії.

Результуюча оцінка змінних у часі параметрів моделюється і прогнозується як стохастичні динамічні (часові) ряди з використанням стандартних методів. Із спрогнозованого загального рівня смертності, обраховуються фактичні повікові показники. Прогнози різних показників таблиць дожиття мають розподіл ймовірності, тому ймовірнісні інтервали можуть бути розраховані для кожного показника та для зведених показників, як наприклад, очікувана тривалість життя [16].

Переважає більшість сучасних демографічних прогнозів смертності розраховані методом екстраполяції. Гіпотеза про те, що майбутня зміна смертності є продовженням змін, які відбувалися і відбуваються, здається доволі природною, особливо, якщо врахувати те, що смертність населення

більшості країн світу стійко знижується. Але і в найбільш успішних країнах часом виникають зворотні тенденції, які ставлять під сумнів можливості звичайного екстраполяційного прогнозування. Метод Лі-Картера є суттєвим відходом від існуючих методів. Зокрема, це пов'язано з двома факторами (вік і час) та використанням матричного розкладу для того, щоб отримати єдиний змінний за часом коефіцієнт рівня смертності. Метод був розроблений для довгострокового прогнозу, ґрунтованого на тривалому часовому ряді історичних даних.

Використовуючи модель Лі-Картера припускають, що центральне значення смертей всіх вікових груп залежить від віку когорти (група осіб, об'єднаних будь-якою загальною статистичною ознакою, в даному випадку - віком) і календарного року.

Використовувана система позначень моделі Лі-Картера:

- $D_{x,j}$ – число смертей між віком x і $x + 1$ в рік t ;
- $E_{x,t}$ – число схильних до ризику від x до $x + 1$ в рік t ;
- $q_{x,t} = \frac{D_{x,j}}{E_{x,t}}$ – центральне значення смертності для віку x в рік t .
- $m_{x,t}$ – логарифмічно перетворений вікової центральний рівень смертності:

Модель, що використовується для представлення смертності :

$$\ln(m_{x,t}) = a_x + b_1 k_t + \varepsilon_{x,t} \quad (1.9)$$

a_x з рівняння (1.9) – повіковий параметр, рівний середній величині $m_{x,t}$ – в період t і мінімізує суму квадратів залишкового члена $S = \sum_{x,t} \varepsilon_{x,t}^2$.

$$a_x = \frac{1}{T} \sum_x \ln m_{x,t}, \quad \sum_x b_x^2 = 1, \quad \sum_t k_t = 0 \quad (1.10)$$

де T – повне число календарних років.

Використання звичайного метода найменших квадратів неможливе. Аби подолати цю перешкоду використовується двох-етапна процедура оцінки, яка дає точні рішення. У першій стадії до матриці $\{\ln(m_{x,t}) - a_x\}$ застосовується сингулярний розклад, аби отримати оцінки основних компонент b_x і k_t . На другій стадії часовий ряд k_t повторно оцінюється, за допомогою вирішення

$$D_t = \sum \{\exp(a_x + b_x k_t) N_{x,t}\}, \quad (1.11)$$

де D_t - загальна кількість смертельних випадків в часі t ,

N_x є популяційний вік x в часі t .

Для оцінки параметрів моделі найчастіше використовуються:

- сингулярний розклад;
- зважений метод найменших квадратів;
- оцінка максимальної правдоподібності.

На наступному кроці потрібно змодельовати k як стохастичний процес часового ряду. Найчастіше використовується стандартна одновимірна модель часового ряду ARIMA (0,1,0).

$$k_t = \theta + k_{t-1} + \varepsilon_t, \quad (1.12)$$

$$\theta = \frac{k_t - k_{t-1}}{T-1}, \quad (1.13)$$

де θ параметр дрейфу.

На основі моделі (1.12) можна визначити яким чином значення параметра за минулі роки буде впливати на майбутнє. Саме це і дозволить в подальшому скласти прогноз.

Модель Лі-Картера дозволяє отримати прогнози різного роду:

- очікуваної тривалості життя;
- ймовірності смерті;
- кількості померлих та народжених [16].

1.3.4 Порівняння найпопулярніших моделей

Для порівняння найпоширеніших моделей демографічного прогнозування, виділимо їх основні характеристики і об'єднаємо у таблицю (табл. 1.2):

Таблиця 1.2 – Порівняння найпопулярніших моделей

Характеристики	Багато-факторне прогнозування	Переваги	Недоліки
Метод/модель			
Множинний регресійний аналіз	+	- простота; - гнучкість; - велика кількість різновидів моделей. - простота інтерпретування параметрів;	- складність підбору функціональної залежності; - трудомісткість визначення параметрів моделі.
Модель Брасса	-	- простота моделі; - можливість побудови прогнозу та таблиць смертності на основі неповної інформації.	- не застосовується для багатофакторного прогнозування;
Модель Лі-Картера	-	- простота інтерпретування параметрів; - добре підходить для довгострокових прогнозів;	- трудомісткість визначення параметрів моделі;
Нейронні мережі	+	- нелінійність, здатність встановлювати нелінійні залежності між майбутніми та фактичними значеннями процесів; - швидкість обчислень;	- відсутність прозорості моделювання; - високі вимоги до несуперечності навчальної вибірки; - складність вибору алгоритму навчання та ресурсомісткість процесу їх навчання
Моделі і методи на базі ланцюгів Маркова	+	- простота моделювання; - одноманітність аналізу і проектування.	- неможливість моделювання процесів з довгою пам'яттю; - вузька застосовність моделей.

Висновки до розділу 1

Демографічні прогнози і демографічні моделі є важливими елементами в управлінні суспільними процесами. Вони дозволяють цілеспрямовано впливати на розвиток соціально-економічних явищ, коригувати їх у необхідну для країни сторону.

Традиційні підходи до моделювання смертності полягають у побудові таблиць дожиття, які являють собою модель процесу вимирання когорти народжених. Проте існує велика кількість моделей та методів для аналізу та прогнозування процесів смертності. Проведений аналіз множини моделей показує, що математичний апарат визначення залежності демографічних показників від соціально-економічних факторів залишається не достатньо розвинутих в наш час.

Саме тому у розділі було розглянуто найпоширеніші моделі, що використовуються в демографічному прогнозуванні процесів смертності та зроблено їх порівняльний аналіз, виділено переваги і недоліки.

Також був проведений аналіз існуючих систем прогнозування. Результати якого показали, що вітчизняні розробки в галузі аналізу та прогнозування демографічних показників практично відсутні. А існуючі системи, які мають добре розвинений апарат моделювання та прогнозування коштують занадто дорого.

Тому, реалізація математичного та програмного забезпечення для визначення факторів впливу на показники смертності та прогнозування очікуваної тривалості життя в залежності цих факторів є необхідною передумовою розвитку соціальної політики країни.

Для реалізації поставленої задачі вибрано використання двох методів: модифікованого Лі-Картера та множинного регресійного аналізу

2 ОПИС МОДЕЛІ, ЩО ВИКОРИСТОВУЄТЬСЯ

2.1 Теоретичне обґрунтування

Метод прогнозування являє собою послідовність дій, які слід виконати для отримання моделі прогнозування. За аналогією з кулінарією метод є послідовність дій, згідно з якою готується страва – тобто зробиться прогноз.

Модель прогнозування – це функціональне уявлення, що адекватно описує досліджуваний процес і що є основою для отримання його майбутніх значень. У тій же кулінарній аналогії модель є список інгредієнтів і їх співвідношення, необхідний для нашого блюда – прогнозу [17].

Математичний апарат даної роботи включає основи роботи із динамічними (часовими) рядами, регресійним аналізом та моделлю Лі-Картера. Для прогнозування очікуваної тривалості життя використовуються таблиці смертності – ймовірнісні моделі з дискретним часом, дані яких можна розглядати, як динамічні (часові) ряди.

Основний підхід при моделюванні законів смертності передбачає перерахування базових характеристик життя і смертності окремо в кожній віковій групі. Зібрані в єдину таблицю, ці характеристики утворюють так звану таблицю смертності населення яку іноді називають таблицею дожиття населення (оскільки подією, протилежною до смерті, є факт доживання людини до певного віку). Також таблиці смертності називають моделями з дискретним часом, бо смертність в них змінюється дискретно [17].

До переваг таблиць дожиття можна віднести:

- високу точність відомостей, що містяться в них;
- відносну простоту використання;
- наочність.

Недоліки таблиць смертності:

- дискретність, тобто, відсутність відомостей про розподіл смертності усередині виділеної а таблиці вікової групи;
- принципову неможливість формалізації законів смертності у формі математичної функції, що знижує можливості її аналізу [18].

Таблиці смертності відтворюють співвідношення сили смертності в різних вікових групах на протязі життя, а показники середньої тривалості життя характеризують рівень здоров'я населення. Вони можуть бути визначені як збірка статистичних даних про смертність і доживання населення.

Розрізняють повні таблиці смертності і короткі. Повні будуються для однорічних вікових груп. Але вікові групи від 85 років об'єднуються в одну групу – особи 85 років і старше. Короткі таблиці будуються п'ятирічним віковим групам. Ці таблиці в першу чергу мають велику точність при отриманні вікових коефіцієнтів смертності по збільшеним інтервалам, ніж по малим, у зв'язку з відносно більшою дисперсією в даних в останньому випадку.

Таблиці смертності містять наступні показники: вік, число доживаючих до віку x , число померлих у віці x , число людино-років майбутнього життя для осіб, що досягли віку x [19].

Таблиці смертності будуються на основі спостережень над фактичною смертністю населення різноманітних вікових груп. Дані представляються у вигляді динамічних рядів. Традиційний підхід заключається в тому, щоб включати в одну таблицю смертності дані про смертність різних поколінь. Цю особливість потрібно враховувати при використанні таблиць смертності у практиці [17].

Ряд динаміки (часовий ряд) – це впорядкована (за часом) послідовність значень деякої довільної змінної величини. Кожне окреме значення даної змінної називається відліком тимчасового ряду. Тим самим, часовий ряд істотно відрізняється від простої вибірки даних. Часовий ряд може бути записаний у вигляді: $Y_t = \{y_1, y_2, \dots, y_n\}$, де індекс t вказує на момент часу, в який зафіксовано значення або номер спостереження.

Ряди динаміки бувають одномірні й багатовимірні. Одномірні ряди містять спостереження за зміною тільки одного параметра досліджуваного процесу або об'єкта, а багатовимірні – за двома або більше параметрами. Елементами часових рядів є значення спостережуваного показника, що зветься рівнем ряду, а також моменти та інтервали часу, до яких відносяться рівні.

Якщо в динамічному ряді проявляється тривала («вікова») закономірність зміни рівнів, то говорять, що має місце тренд. Таким чином тренд визначає загальний напрямок розвитку економічного процесу.

Існують дві основні цілі аналізу часових рядів:

- визначення природи ряду;
- прогнозування (пророкування) майбутніх значень часового ряду по теперішнім і минулим значенням.

Обидві ці цілі вимагають, щоб модель ряду була ідентифікована і, більш-менш формально описана. Як тільки модель визначена, можна з її допомогою інтерпретувати представлені дані (наприклад, використовувати у для розуміння сезонної зміни цін на товари, якщо досліджувати економіку). Не звертаючи уваги на глибину розуміння і справедливості теорії, можна екстраполювати ряд на основі знайденої моделі, ґрунтуючись на відомих подіях минулого, тобто передбачити його майбутні значення [20].

Смертність – найбільш інерційний із демографічних показників. Тренди коефіцієнтів смертності за однорічними віковими групами можуть не мати чітко вираженої динаміки. Натомість такі показники за

п'ятирічними віковими інтервалами характеризуються більш чіткою спрямованістю. Вони часто виражають у часі досить чіткі тенденції змін смертності, зумовлені переважно соціальними й економічними умовами як найбільш кон'юнктурними у коротких проміжках часу. Саме тому прогноз будується по 5-річним віковим категоріям.

Оскільки для показників смертності притаманний лінійний тренд, доцільно використовувати модель лінійної множинної регресії для прогнозування рівня смертності в залежності від впливу тих чи інших причин [21]. Проте як було доведено, ця модель дає не достатньо добрі результати для віку 0-24 років та для категорії 70+.

Під час досліджень було виявлено, що для цих вікових інтервалів краще використовувати модель Лі-Картера а модель множинної лінійної регресії дає гарні результати у вікових категоріях 25-70

2.2 Аналітичне представлення моделі

Побудова таблиць смертності та середньої очікуваної тривалості життя – один із найбільш традиційних методів дослідження даних про смертність. Будуються з метою отримати можливість динамічного співставлення демографічних показників та подальшої роботи з ними [13].

В якості математичної моделі для прогнозування очікуваної тривалості життя від найбільш впливових факторів розглядається комбінація трьох відомих моделей, з їх удосконаленням.

Для побудови прогнозу потрібно зробити наступні кроки:

- Використовуючи дані про населення, кількість народжених та померлих, побудувати таблиці смертності та очікуваної тривалості життя;

- Виконати прогноз очікуваної тривалості життя для вікових категорій 0-24 та 70+ використовуючи модифікований метод Лі-Картера;
- Маючи інформацію про розподіл померлих за причинами смерті, обрахувати вплив цих факторів на повіковий коефіцієнт смертності та спрогнозувати очікувану тривалість життя для вікової категорії 25-70pp, використовуючи модель множинного регресійного аналізу.;

2.2.1 Побудова таблиць дожиття

Для побудови таблиць смертності та середньої очікуваної тривалості життя для населення України, використаємо методику, розроблену науковцями ІДСД ім. М.В. Птухи та Держкомстату України [19].

Цей алгоритм передбачає побудову скорочених таблиць смертності та середньої очікуваної тривалості життя для України в цілому (у т. ч. за статтю).

Для побудови таблиць смертності та середньої очікуваної тривалості життя використовуються такі дані:

- Розподіл померлих за однорічними віковими групами з урахуванням померлих невідомого віку;
- Розподіл постійного населення за однорічними віковими групами на початок року.
- Розподіл народжених за статтю.
- Розподіл померлих немовлят за роком народження з урахуванням померлих невідомого року народження.

Спочатку будується повна таблиця смертності, для однорічних вікових інтервалів:

Імовірність не дожити до наступного віку q_x розраховується на базі повікового коефіцієнта смертності, який обчислюється наступним чином:

m_x - коефіцієнт смертності населення x вікової групи:

$$m_x = \frac{M_x}{S_x} \quad (2.1)$$

Спочатку визначається чисельник формули - кількість померлих осіб (M) - на основі даних статистичної форми розподіленою кількістю померлих осіб невідомого віку

Для вікових груп 0, 1, 2 роки:

$$M_x = M_x^{n-1} + M_x^n, \quad (2.2)$$

де n – рік, для якого проводиться побудова таблиць.

Для інших вікових груп:

$$M_x = M_{x-1}^{n-1} + M_x^n \quad (2.3)$$

Відповідно до розрахунку кількості померлих (M_x) визначається і чисельність населення у вікових групах (S_x)

Для вікових груп 0, 1, 2 роки:

$$S_x = \overline{S_x^{n-1}} + \overline{S_x^n} \quad (2.4)$$

Для інших вікових груп:

$$S_x = \overline{S_{x-1}^{n-1}} + \overline{S_x^n} \quad (2.5)$$

Імовірність померти q_0 для народжених (від народження до досягнення 1 року) обчислюється за формулою (2.1) :

$$q_0 = 1 - \frac{N^t - {}^tM_0^t}{N^t} \cdot \frac{N^{t-1} - {}^{t-1}M_0^{t-1} - {}^{t-1}M_0^t}{N^{t-1} - {}^{t-1}M_0^{t-1}}, \quad (2.6)$$

де N^t – число народжених у році t ;

${}^tM_0^t$ – число померлих у віці 0 років у році t з числа народжених у році t .

Визначення коефіцієнта смертності населення вікової групи. Для вікових груп 0, 1, 2 роки коефіцієнт смертності x вікової групи буде становити:

$$m_x = \frac{M_x^{n-1} + M_x^n}{S_x^{n-1} + S_x^n}, \quad (2.7)$$

де m_x – коефіцієнт смертності населення x вікової групи у році n ;

M_x^n – кількість померлих у віці x років в n -му році;

S_x^n - середньорічна чисельність населення у віці x років в n -му році.

Для проведення подальших розрахунків необхідно здійснити перехід від коефіцієнта смертності населення x вікової групи до ймовірності не дожити за наступною формулою

$$q_x = \frac{m_x}{1 + 0,5 \cdot m_x} \quad (2.8)$$

Значення показників імовірності не дожити у віковій групі 50-100 років після розрахунку підлягають згладжуванню методом плинної середньої третього порядку з подвійною вагою групи, для якої проводиться розрахунок. Заміна значень показників проводиться за допомогою середньої геометричної з суміжних значень показників:

Для віку 50-99 згладжування відбуватиметься за формулою:

$$\overline{q_x} = \sqrt[4]{q_{x-1}(q_x)^2 \cdot q_{x+1}} \quad (2.9)$$

де x – вік;

q_x – ймовірність не дожити до наступного року;

$\overline{q_x}$ – середнє згладжене значення показника ймовірності не дожити до наступного віку.

Для віку 100 років:

$$\overline{q_{100}} = \sqrt[4]{q_{99}(q_{100})^2} \quad (2.10)$$

Розрахунок решти показників таблиці смертності і середньої очікуваної тривалості майбутнього життя відбуватиметься на основі згладжених значень показника імовірності не дожити.

p_x – ймовірність дожити до наступного віку ($x+1$) для осіб, яким виповнилося x років. Показник є похідним від q_x і розраховується відповідно:

$$p_x = 1 - q_x, \quad (2.11)$$

де l - число осіб, які доживають до даного віку (x).

Для новонароджених l береться як одиниця з п'ятьма нулями (число нулів забезпечують певний рівень точності). Для інших вікових груп l визначаємо за формулою:

$$l_x = l_{x-1}p_{x-1} \quad (2.12)$$

d_x – число осіб, які вмирають у віці x років:

$$d_x = l_x q_x \quad (2.13)$$

L_x – число осіб, число осіб, які живуть у віці від x до $x+1$ років. На відміну від l_x , що є числом моментним, L_x – це середня чисельність теоретичного населення в кожній x віковій групі та визначається (окрім останнього значення):

$$L_x = \frac{l_x + l_{x+1}}{2} \quad (2.14)$$

T_x – число людино-років майбутнього життя для людей, що живуть у віці x і старше, тобто T_x – це кумулятивне число осіб, які живуть у віці від x до $x+1$.

T_x має ще один зміст – це сумарна чисельність стаціонарного населення (населення, яке не зазнає змін в чисельності та в складі) в таблиці смертності і середньої очікуваної тривалості майбутнього життя.

Якщо T_x – це потенціал, то його значення в групі новонароджених буде становити:

$$T_0 = \sum_{k=0}^{100} L_x \quad (2.15)$$

e_x^0 - середня очікувана тривалість майбутнього життя для людей, яким виповнилось x років. e_x^0 - середнє число років життя, яке може прожити покоління при народженні, за умови збереження сучасної інтенсивності вмирання. Середня очікувана тривалість майбутнього життя розраховується:

$$e_x^0 = \frac{T_x}{l_x} \quad (2.16)$$

Виконавши всі обчислення отримаємо повну таблицю смертності. Для забезпечення тотожності скорочена таблиця отримується на основі даних повної таблиці.

З початку з повної таблиці вибираються числа осіб, які доживають до віку x років (l_x): 0, 1, 5, 10 і далі, для вікових груп, кратних 5, до 85 років.

Величину ймовірність померти у віці від народження (0 років) до 1 року необхідно взяти з повної таблиці. Решта ймовірностей померти у віці від x до $x + n$ до вікового інтервалу 100+ включно, обчислюється за формулою:

$${}_nq_x = 1 - \frac{l_{x+n}}{l_x}, \quad (2.17)$$

де n – ширина вікового інтервалу.

Ймовірність ${}_np_x$ дожити до наступного віку $x + n$ обчислюється за формулою (2.11) до вікового інтервалу 100+ включно.

${}_nd_x$ – число осіб, які вмирають у віці від x до $x + n$ для віку $1 \leq x < 100$ розраховується за формулою:

$${}_nd_x = l_x - l_{x+n} \quad (2.18)$$

для $x = 100$ значення переносяться з повної таблиці смертності.

${}_nL_x$ – число осіб, які живуть у віці від x до $x + n$, для $x = 0$ переносяться з повної таблиці. Для решти вікових інтервалів до інтервалу 100+ включно, значення обчислюється за формулою:

$${}_nL_x = \sum_x^n L_x, \quad (2.19)$$

де L_x отримуються з повної таблиці.

Кількість людино-років майбутнього життя при досягненні віку x років T_x переноситься з повної таблиці.

Середня очікувана тривалість життя при досягненні віку x років переноситься з повної таблиці [19].

2.2.2 Модель Лі-Картера

По суті модель Лі-Картера описується логарифмічно перетвореною центральною нормою смертності ($m_{x,t}$ – показник смертності для віку x в час t), як сума вікового компонента, який незалежний від часу (a_x), і продукт змінного часом параметра (k_t , також відомий як індекс смертності), який підсумовує загальний рівень смертності і додаткового вікового компонента (b_x), який показує, як швидко або повільно змінюється смертність в кожному віці, коли індекс смертності змінюється. Було доведено, що модель такого роду дає найбільш задовільний результат з великого класу простих моделей [16].

Модель має вигляд:

$$\ln(m_{x,t}) = a_x + b_x k_t + \varepsilon_{x,t} \quad (2.20)$$

Останній компонент $\varepsilon_{x,t}$ є залишковим членом, який відображає вікові впливи, що не захоплені моделлю.

Щоб отримати унікальне рішення, ми беремо a_x як середнє арифметичне $\ln(m_{x,t})$ протягом довгого часу, тоді як суми b_x і k_t нормалізовані і дорівнюють одиниці та нулю відповідно:

$$a_x = \frac{1}{T} \sum_x \ln m_{x,t}, \quad \sum_x b_x^2 = 1, \quad \sum_t k_t = 0 \quad (2.21)$$

Використання звичайного метода найменших квадратів неможливе. Аби подолати цю перешкоду використовується двох-етапна процедура оцінки, яка дає точні рішення. У першій стадії до матриці $\{\log(m_{x,t}) - a_x\}$ застосовується сингулярний розклад, аби отримати оцінки основних компонент b_x і k_t . На другій стадії часовий ряд k_t повторно оцінюється, за допомогою :

$$D_t = \sum \{ \exp(a_x + b_x k_t) N_{x,t} \}, \quad (2.22)$$

де D_t - загальна кількість смертельних випадків в часі t ,

N_x є популяційний вік x в часі t .

Є декілька переваг в тому, щоб здійснювати таким чином другу оцінку параметра k . Таким чином, емпіричний часовий ряд k може бути розширений, щоб включити роки, на протязі яких вікові дані про смертність не доступні.

Дослідники Марі Клер Коїст, Арнольд Шапіро та Горан Хогнас провели порівняльний аналіз застосування найпопулярніших методів для оцінки параметрів моделі [16]. Їх результати показали, що сингулярний розклад – це найкраща альтернатива для пошуку індексу смертності k , особливо якщо використовуються дані для Європейських країн. Саме тому його було обрано для роботи.

Розглянемо метод сингулярного розкладу (Singular Value Decomposition – SVD), що використовується.

З рівняння (2.20), параметр a_x можна легко обчислити, як середнє за часом від логарифма загальної смертності. Потім застосовуємо сингулярний розклад матриці

$$Z_{x,t} = \ln(m_{x,t}) - \hat{a}_x \quad (2.23)$$

і отримаємо матриці:

$$ULV' = \text{SVD}(Z_{x,t}) = L_1 U_{x1} V_{t1} + \dots + L_x U_{xx} V_{tx}, \quad (2.24)$$

де U – вікова компонента;

L – сингулярні значення;

V – компонента часу.

Наближення першого доданку дає оцінку :

$$\widehat{b}_x = U_{x1} \quad (2.25)$$

$$\widehat{k}_t = L_1 V_{t1} \quad (2.26)$$

В загальному, застосування методу сингулярного розкладу (SVD) є відносно простим і складається із шести кроків:

Крок 1:

$$\widehat{a}_x = \frac{1}{T} \sum_x \ln(m_{x,t}) \quad (2.27)$$

Крок 2. Створюємо матрицю $Z_{x,t}$ для оцінки b_x і k_t , де

$$Z_{x,t} = \ln(m_{x,t}) - \widehat{a}_x = b_x k_t \quad (2.28)$$

Крок 3. Застосовуємо сингулярний розклад для матриці $Z_{x,t}$. В результаті отримуємо декомпозицію на три матриці, що представлені формулою (2.24).

Крок 4. Вибір сингулярного розкладу. \widehat{k}_t походить від першого вектора часу компонент матриці і першого сингулярного значення

($\hat{k}_t = L_1 V_{t1}$), а \hat{b}_x походить від першого вектора вікової компоненти матриці ($\hat{b}_x = U_{x1}$).

Крок 5. (Лі-Картер) Апроксимація нової матриці $\hat{Z}_{x,t}$ за допомогою результатів оцінювання параметрів \hat{b}_x і \hat{k}_t та отримання $\hat{Z}_{x1,t1} = \hat{b}_{x1} \hat{k}_{t1}$:

$$\hat{Z}_{x1,t1} = \begin{bmatrix} \hat{z}_{x_1 t_1} & \hat{z}_{x_1 t_2} & \dots & \hat{z}_{x_1 t_n} \\ \hat{z}_{x_2 t_1} & \hat{z}_{x_2 t_2} & \dots & \hat{z}_{x_2 t_n} \\ \dots & \dots & \dots & \dots \\ \hat{z}_{x_A t_1} & \hat{z}_{x_A t_2} & \dots & \hat{z}_{x_A t_n} \end{bmatrix} \hat{b}_x \hat{k}_t \quad (2.29)$$

Крок 6. Оцінка логарифму показника смертності

$$\hat{\ln}(m_{x,t}) = \hat{a}_x + \hat{Z}_{x,t} = \hat{a}_x + \hat{b}_x \hat{k}_t \quad (2.30)$$

Після оцінки параметрів, приступаємо, безпосередньо, до прогнозування. Однією з переваг методу Лі-Картера є те, що після того, як дані визначені для моделі і значення векторів \hat{a}_x , \hat{b}_x та \hat{k}_t знайдені, прогнозувати потрібно тільки індекс смертності \hat{k}_t .

На наступному кроці потрібно змодельювати k як стохастичний процес часового ряду. Найчастіше використовується стандартна одновимірна модель часового ряду ARIMA (0,1,0).

$$\hat{k}_t = \hat{k}_{t-1} + \theta + \varepsilon_t, \quad (2.31)$$

$$\hat{\theta} = \frac{\hat{k}_T - \hat{k}_1}{T-1}, \quad (2.32)$$

де θ параметр зсуву.

Це означає, що θ залежить лише від перших і останніх оцінок k_t , тоді як ε_t — залишковий член. Потім, щоб передбачити на два періоди вперед, ми лиш замінюємо k_{t1} :

$$\hat{k}_t = \hat{k}_{t-1} + \hat{\theta} + \varepsilon_t = (\hat{k}_{t-2} + \hat{\theta} + \varepsilon_{t-1}) + \hat{\theta} + \varepsilon_t = \hat{k}_{t-2} + 2\hat{\theta} + (\varepsilon_{t-1} + \varepsilon_t) \quad (2.33)$$

Щоб передбачити \hat{k}_t на час $T+(\Delta t)$ з даними відомими для періоду T , ми слідуємо за тією ж самою процедурою і отримуємо:

$$\hat{k}_{T+(\Delta t)} = \hat{k}_T + (\Delta t)\hat{\theta} + \sum_n^{(\Delta t)} \varepsilon_{T+n-1} = \hat{k}_T + (\Delta t)\hat{\theta} + \sqrt{(\Delta t)}\varepsilon_t \quad (2.34)$$

Якщо проігнорувати залишковий член, ми можемо отримати пряму, яка є функцією від (Δt) , з нахилом θ :

$$\hat{k}_{T+(\Delta t)} = \hat{k}_T + (\Delta t)\hat{\theta} = \hat{k}_T + (\Delta t) \frac{\hat{k}_T - \hat{k}_1}{T-1} \quad (2.35)$$

Прогнозування \hat{k}_t в цьому випадку не складне: екстраполяція прямої, проведеної через першу точку \hat{k}_1 і останню точку \hat{k}_T і через решту точок \hat{k}_t ігнорується.

Недоліком методу є те, що він пропонує визначений спосіб зміни смертності у розподілі за віком, таким чином, що норми зниження у різному віці завжди підтримують одне і те ж відношення на протязі довгого часу. Але на практиці, відносна швидкість зниження в різному віці може змінюватись. Наприклад у Швеції смертність в старості мала тенденцію зменшуватись більш повільно, ніж в іншому віці. Метод не може врахувати такі зміни [22].

Модифікацію цього методу полягає у тому, що замість оцінки коефіцієнт a_x як середнє арифметичне $\ln(m_{x,t})$, можна встановити a_x рівним логарифму нових показників смертності. Це забезпечить, що перший рік прогнозу близько співпаде з показниками смертності, які спостерігались востаннє. В якості початкової оцінки потрібно взяти останні показники. Тоді прогноз отримаємо за допомогою формули:

$$\ln(m_{x,t+s}) = \ln(m_{x,t}) - b_x(k_{t+s} - k_t) \quad (2.36)$$

Це забезпечує, що початкова оцінка точно відповідає віковим показникам смертності та тривалості життя, що спостерігаються [16].

2.2.3 Множинний регресійний аналіз

Розглянемо алгоритм застосування множинного регресійного аналізу для побудови моделі залежності коефіцієнта смертності залежно від факторів впливу для вікових категорій 25-70 років :

Оскільки показник смертності та очікуваної тривалості життя доволі інертний, лінійна апроксимація цього показника на майбутнє показує доволі точні результати [23] та з огляду на найбільш чітку інтерпретацію параметрів, будемо використовувати лінійну модель множинної регресії:

$$y = \alpha' + \beta'_1 x_1 + \beta'_2 x_2 + \dots + \beta'_p x_p + \varepsilon, \quad (2.37)$$

$$\text{де } \beta'_1 = \frac{\partial y}{\partial x_1}, \beta'_2 = \frac{\partial y}{\partial x_2}, \dots, \beta'_p = \frac{\partial y}{\partial x_p}.$$

Параметр α називається вільним членом і визначає значення y в тому випадку, коли всі вагові коефіцієнти рівні 0. Але, як і в випадку парної регресії, фактори за своїм смислом не можуть приймати нульових значень, саме тому значення вільного члена не має економічного смислу.

Значення кожного коефіцієнта β'_1 дорівнює середній зміні y при збільшенні величини x_j на одиницю але тільки при умові, що решта факторів залишилися незмінними.

Величина ε являє собою випадкову похибку регресійної залежності.

Для оцінки впливу параметрів $x_1 \dots x_p$ потрібно знайти значення коефіцієнтів $\alpha', \beta'_1, \beta'_2, \dots, \beta'_p$ рівняння регресії (2.37). Але, оскільки вони є випадковими величинами, то коректно визначити їх значення використовуючи дані з вибірки неможливо. Саме тому, замість теоретичного рівняння регресії (2.37) оцінюється емпіричне рівняння регресії, яке представлено у вигляді:

$$y = a + b_1x_1 + b_2x_2 + \dots + b_px_p + e, \quad (2.38)$$

де a, b_1, b_2, \dots, b_p – емпіричні коефіцієнти регресії (оцінки

теоретичних значень $\alpha', \beta'_1, \beta'_2, \dots, \beta'_p$;

e – оцінка відхилення ε .

Тоді вираз для розрахунків має вигляд:

$$\hat{y} = a + b_1x_1 + b_2x_2 + \dots + b_px_p \quad (2.39)$$

Нехай ε n спостережень, що пояснюють змінні та відповідні значення результативного признака:

$$(x_{i1}, x_{i2}, \dots, x_{ip}, y_i), \quad i = \overline{1, n} \quad (2.40)$$

Для однозначного визначення значень параметрів рівняння (2.39), об'єм вибірки n має бути не меншим, ніж кількість параметрів, тобто $n \geq p + 1$. Інакше значення параметрів не може бути визначено однозначно. На практиці досліджено, що об'єм вибірки повинен перевищувати кількість параметрів при x_j в 6-7 раз.

Для проведення аналізу в рамках лінійної моделі множинної регресії потрібно виконання ряду передумов МНК. В основному це ті ж

передумови, що і для парної регресії, однак тут потрібно додати припущення, специфічні для множинної регресії:

- Специфікація моделі має вигляд (2.37);
- Відсутність мультиколінеарності: між факторними змінними відсутня строга лінійна залежність, що грає важливу роль у відборі факторів при вирішенні проблеми специфікації моделі;
- Похибки $\varepsilon_i, i = \overline{1, n}$ мають нормальний розподіл ($\varepsilon_i \sim N(0, \sigma)$).

Виконання цих умов потрібно для перевірки статистичних гіпотез і побудови інтервальних оцінок. Якщо всі ці умови задовольняються, має місце багатовимірний аналог теореми Гауса - Маркова: оцінки a, b_1, b_2, \dots, b_p , отримані за МНК, є найбільш ефективними (в смислі найменшої дисперсії) в класі лінійних незміщених оцінок [13].

Розрахунок параметрів регресії виконувався за допомогою МНК, що застосовувався до стандартизованої моделі регресійного рівняння:

$$t_y = \gamma_1 t_{x_1} + \gamma_2 t_{x_2} + \dots + \gamma_p t_{x_p} + \varepsilon, \quad (2.41)$$

де $t_y, t_{x_1}, t_{x_2}, \dots, t_{x_p}$ – стандартизовані змінні.

$$t_y = \frac{y - \bar{y}}{\sigma_y}; \quad t_{x_j} = \frac{x_j - \bar{x}_j}{\sigma_{x_j}}, \quad j = \overline{1, n} \quad (2.42)$$

де $\bar{t}_y = \bar{t}_{x_1} = \bar{t}_{x_2} = \dots = \bar{t}_{x_p} = 0$ – середнє значення;

$\sigma_y = \sigma_{t_{x_j}} = 1, j = \overline{1, n}$ – середнє квадратичне відхилення;

γ_j – стандартизовані коефіцієнти рівняння регресії.

Застосовуючи метод найменших квадратів до (2.41), після відповідних перетворень отримують систему нормальних рівнянь :

Включення в модель факторів з високою взаємною кореляцією може призвести до небажаних наслідків – система нормальних рівнянь може виявитися погано обумовленою і спричинити нестійкість і ненадійність оцінок коефіцієнтів регресії. Якщо між факторами існує висока кореляція, то не можна визначити їх ізольований вплив на результативний показник і параметри рівняння регресії виявляються такими що неможливо інтерпретувати.

Відбір факторів проводиться на основі якісного теоретико - економічного аналізу. Теоретичний аналіз часто не дозволяє однозначно відповісти на питання про кількісну взаємозв'язку розглянутих ознак і доцільності включення фактора в модель. Тому відбір факторів зазвичай здійснюється в дві стадії: на першій підбираються фактори, виходячи із сутності проблеми; на другий - аналізується матриця показників кореляції і встановлюється, які з факторів найбільш тісно пов'язані з результатом, а які - між собою.

Таким чином, хоча теоретично регресійна модель дозволяє врахувати будь-яке число факторів, практично в цьому немає необхідності.

Висновки до розділу 2

Однією із фундаментальних властивостей народонаселення є його постійне самооновлення в процесі послідовної зміни поколінь. Розуміння об'єктивного характеру закономірностей відтворення населення та його складових, а також знання їх тенденцій дозволяють дослідникам аналізувати та прогнозувати демографічні процеси, а також розробляти рекомендації щодо загальних напрямків і конкретних заходів соціальної та демографічної політики.

Проте, прогнозування тільки загального показника очікуваної тривалості життя має досить обмежене застосування та мало інформативний. Саме тому в даному розділі було розглянуто методи визначення міри впливу соціально-економічних процесів на показники смертності, а також моделі за допомогою яких можна виконати прогноз.

Для вирішення поставленої задачі було розроблено математичний апарат, що включає основи роботи із динамічними (часовими) рядами, множинний регресійний аналіз та модель Лі-Картера.

Дана багатокомпонентна модель в дозволяє провести якісний та кількісний аналіз основних тенденцій і змін показника очікуваної тривалості життя населення України.

Важливою характеристикою демографічних прогнозів є їх достовірність, тобто відповідність характеристик демографічних прогнозів тому, якими вони будуть в дійсності. Достовірність демографічного прогнозу визначається точністю вихідної інформації, обґрунтованістю прийнятих гіпотез, тривалістю прогнозного періоду [5]. Для отримання найбільш достовірної та практично-цінної інформації було обрано прогнозувати на середньостроковий період.

3 ПРОЕКТУВАННЯ ПРОГРАМНИХ ЗАСОБІВ

3.1 Функціональне призначення програмного забезпечення

Розроблене ПЗ призначене для визначення найвагоміших факторів впливу на процеси смертності та прогнозування очікуваної тривалості життя населення України на середньо-строковий період, тобто 10-15 років.

Такі дані можуть використовуватися під час прийняття управлінських рішень стосовно соціально-економічної і демографічної політики, наприклад, фінансування програм розвитку нових ліків та підвищення рівня медицини.

Результати роботи впроваджено у відділі демографічного моделювання та прогнозування Інституту демографії та соціальних досліджень імені М. В. Птухи НАН України для прогнозів на майбутнього рівня очікуваної тривалості життя населення.

3.2 Архітектура програмного забезпечення

Архітектура розроблених програмних засобів (схема взаємодії компонент програми) зображена на рисунку 3.1:



Рисунок 3.1 – Схема взаємодії програми

Розглянемо детальніше складові модулі програми:

- Головний модуль програми, через який здійснюється взаємодія інших компонентів. Містить юзабіліті-інтерфейс управління програмою;
- Введення даних. Модуль, що відповідає за введення даних користувачем для побудови аналізу та прогнозу;
- Перевірка даних на коректність. Модуль, в якому здійснюється перевірка введених даних на коректність. Критерій коректності має відповідати умові задачі.

- Редагування даних. Модуль, який надає можливість редагувати введені/завантажені початкові дані. Відредаговані дані перевіряються на коректність;
- Модуль визначення міри впливу факторів. Виконується оцінка впливу обраних факторів на показник очікуваної тривалості життя на основі введених даних;
- Модуль прогнозування очікуваної тривалості життя. Здійснюється прогнозування очікуваної тривалості життя, відповідно до обраних факторів на заданий термін на основі заданого періоду;
- Виведення результатів. Модуль виведення результатів роботи програми на дисплей або збереження в файл;
- Графічне відображення прогнозу. Модуль, що забезпечує побудову графіків, що ілюструє динаміку показників очікуваної тривалості життя;
- Графічне відображення прогнозу. Модуль, за допомогою якого можна побудувати графік, який ілюструє динаміку чисельності населення;
- Модуль виходу з програми.

3.3 Формати вхідних та вихідних даних

Так як програма створюється для роботи із таблицями даних, то в цілях економії часу, вхідні дані доцільно завантажувати з файлів (зокрема файлів Excel), а не вводити вручну. З тієї ж причини результати роботи програми теж зручно зберігати у вигляді табличних даних Excel.

З програмної точки зору вхідні та вихідні дані представляють собою масиви. Вхідні дані, що використовуються:

- кількість померлих за однорічними віковими групами з урахуванням померлих невідомого віку – це масив чисел, кожне з яких може приймати значення від 0 до сотень тисяч (array[1..60,1..60] of double);
- розподіл постійного населення за однорічними віковими групами на початок року – це масив чисел, кожне з яких може приймати значення від 0 до сотень тисяч (array[1..60,1..60] of double);
- кількість народжених за статтю для року – масив чисел, кожне з яких може приймати значення від 0 до сотень тисяч (array[1..60,1..60] of double);
- розподіл померлих немовлят за роком народження з урахуванням померлих невідомого року народження – масив чисел, кожне з яких може приймати значення від 0 до сотень тисяч (array[1..60,1..60] of double);
- період за яким прогнозуємо – кількість часових одиниць (років) за які беруться дані для побудови прогнозу(integer);
- період на який прогнозуємо – це кількість років для яких будується прогноз (integer);
- кількісні показники факторів впливу – масив чисел, кожне з яких може приймати значення від 0 до сотень тисяч (array[1..60,1..60] of double).

Вихідні дані розробленої програми:

- виведені на екран значення прогнозованого показника очікуваної тривалості життя, розподілу вікової структури чоловічого та жіночого населення (array[1..60,1..60] of double);
- файл що містить прогнозовані значення чисельності та статевовікового розподілу у форматі Excel (xls-файл).

3.4 Вимоги до програмних та технічних засобів

Програма реалізована в середовищі Delphi 7 з використанням стандартних компонентів візуалізації на платформі операційної системи Windows 7.

Мінімальні вимоги для коректної роботи розробленого ПЗ:

- процесор: Intel® Pentium 4 і вище або аналоги;
- оперативна пам'ять: 512 Мб;
- відеопам'ять: 128 Мб;
- операційна система Windows XP і вище;
- Microsoft Office Excel.



3.5 Опис роботи користувача з системою

Для використання програми необхідно скопіювати папку Forecasting на жорсткий диск та запустити файл «Prognoz.exe».

Початкове вікно введення даних (рисунок 3.2) з'являється відразу після запуску програми. Користувачу потрібно задати період наявних даних та горизонт прогнозування, а також ввести або завантажити з Excel кількість померлих за однорічними віковими групами, розподіл постійного населення за однорічними віковими групами, кількість народжених за статтю, розподіл померлих немовлят за роком народження.

Рисунок 3.2 – Вікно введення початкових даних

Помилково введені або завантажені дані можна очистити за допомогою кнопок [Очистити], що розміщені безпосередньо під кожною таблицею.

Після того, як всі початкові дані було завантажено в таблицю, потрібно натиснути кнопку [Далі >>]. Для того, щоб вийти, необхідно натиснути кнопку [ВИХІД].

Якщо було завантажено некоректні дані та натиснуто кнопку [Далі >>], на екран виводиться інформуюче повідомлення (рисунок 3.3):

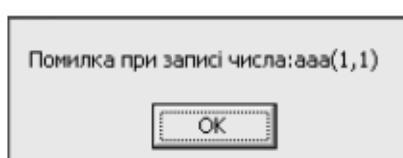


Рисунок 3.3 – Повідомлення про помилку

Якщо ж початкові дані введено коректно, з'являється нове вікно, в якому потрібно ввести або завантажити з Excel значення показників, що

відповідають обраним факторам для розрахункового та прогнозованого періоду (рисунок 3.4).

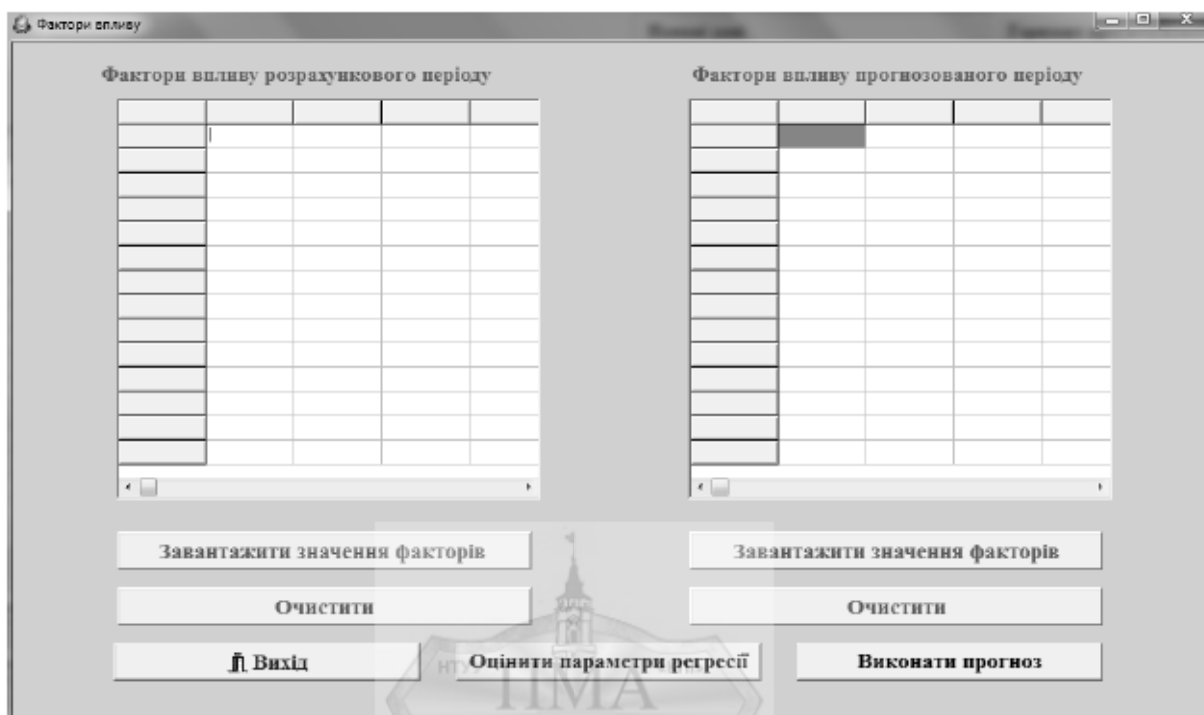


Рисунок 3.4 – Вікно введення значень факторів впливу

Далі можна безпосередньо виконати прогноз або оцінити параметри регресії. Відповідні кнопки розміщені внизу вікна.

Якщо була натиснута кнопка [Оцінити параметри регресії], у вікні виводяться коефіцієнти стандартизованого рівняння регресії (рисунок 3.5). Щоб повернутись до попереднього вікна потрібно натиснути кнопку [Вихід].

Прогнозовані коефіцієнти смертності та показники очікуваної тривалості життя виводяться у вікні результату (рисунок 3.6), якщо було натиснуто кнопку [Виконати прогноз].

Отримані результати можна зберегти, натиснувши кнопку [ЗБЕРЕГТИ В EXCEL]. Щоб повернутись у початкове вікно, потрібно натиснути кнопку [ВИХІД].

Коефіцієнти стандартизованого рівняння логістичної регресії

	x1	x2	x3	x4
20-24				
25-29				
30-34				
35-39				
40-44				
45-49				
50-54				
55-59				
60-64				
65-69				

Вихід

Рисунок 3.5 – Вікно виведення стандартизованих коефіцієнтів регресії

Результат та похибка

	Спрогнозовані коефіцієнти смертності m^x						Середня очікувана тривалість життя e^x					
	2006	2007	2008	2009	2010	2011						
0												
1-4												
5-9												
10-14												
15-19												
20-24												
25-29												
30-34												
35-39												
40-44												
45-49												
50-54												
55-59												
60-64												

Середня очікувана тривалість життя

Побудувати графік коефіцієнтів смертності

Побудувати графік тривалості життя

Відносна похибка:

Середня квадратична похибка:

Рисунок 3.6 – Вікно виведення результатів прогнозування

Висновки до розділу 3

На основі обраної моделі було спроектовано програмне забезпечення для прогнозування очікуваної тривалості життя, показано основні блоки програми та їх взаємодія. Створено інструкцію користувача та описано вимоги до програмних і технічних засобів, що мають бути встановлені на ЕОМ для коректної роботи програми.

Розроблений програмний продукт має зручний користувацький інтерфейс, дані виводяться в чіткому, зрозумілому вигляді.



4 ПРОГРАМНА РЕАЛІЗАЦІЯ

4.1 Опис розробленого алгоритму

На рисунку 4.1 наведено основну блок-схему алгоритму роботи програми

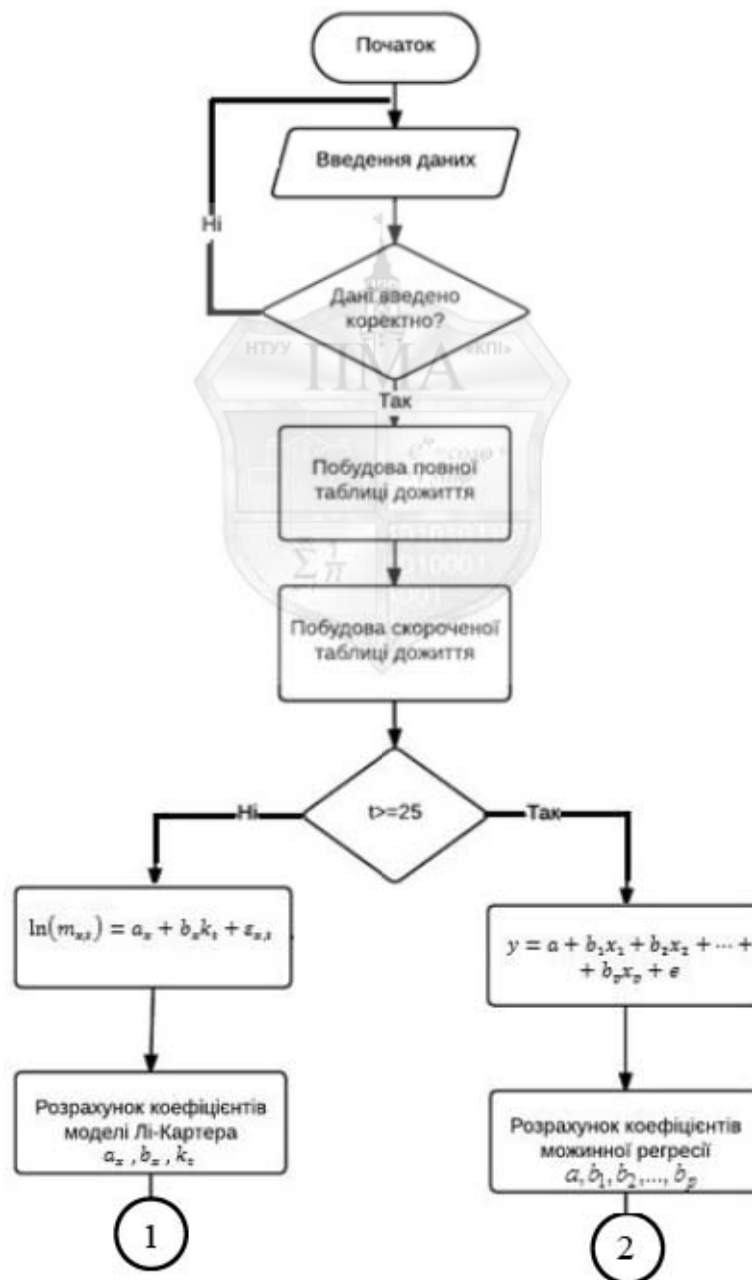


Рисунок 4.1 Блок-схема алгоритму програми

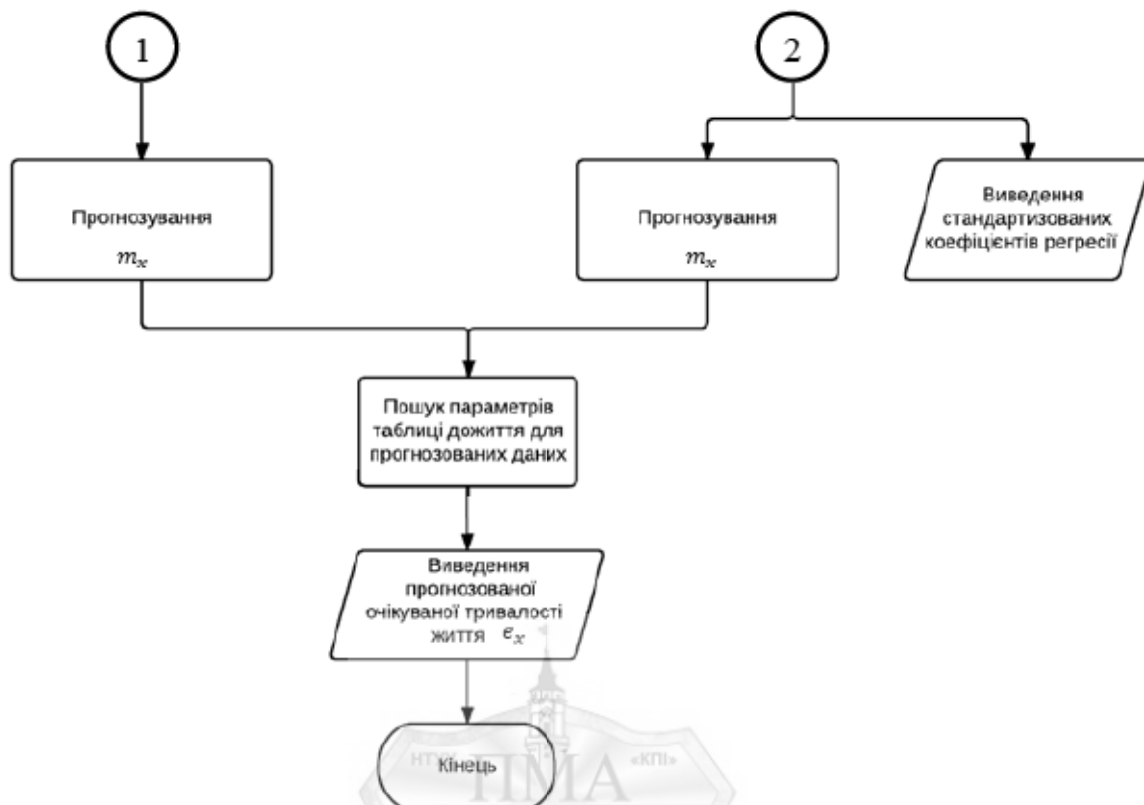


Рисунок 4.1, Аркуш 2

4.2 Опис експериментальних даних

Вхідними даними є дані Державної служби статистики України:

- кількість померлих за однорічними віковими групами з урахуванням померлих невідомого віку (табл.4.1);
- розподіл постійного населення за однорічними віковими групами на початок року (табл.4.2);
- кількість народжених за статтю (табл.4.3);
- розподіл померлих немовлят за роком народження з урахуванням померлих невідомого року народження (табл.4.4) [25].

Виконаємо прогноз очікуваної тривалості життя для жіночого населення України на основі статистичних даних за період 1991-2005 рр. до 2011 року, для демонстрації коректності роботи програми. Результати прогнозу порівняємо з реальними даними за період 2006-2011 рр.

Таблиця 4.1 – Кількість померлих за однорічними віковими групами, жінки

	1991	1992	1993	1994	1995	...	2004	2005
0	317 211	306 206	289 309	271 320	250 881	...	197 084	205 946
1	336 108	317 853	307 991	288 476	269 050	...	186 965	196 618
2	362 501	336 926	320 814	308 220	286 640	...	179 999	186 759
3	364 403	363 292	339 812	321 312	306 621	...	185 434	179 860
4	384 603	365 966	365 980	340 226	320 094	...	188 220	185 319
5	375 319	386 492	369 634	365 947	338 489	...	200 882	188 080
6	383 697	376 972	390 279	370 667	363 951	...	208 905	200 738
7	392 458	385 380	380 530	391 539	369 783	...	225 572	208 777
8	361 647	393 959	389 006	381 709	390 814	...	234 412	225 469
9	351 469	362 668	397 189	390 219	381 012	...	243 634	234 310
10	361 698	352 761	365 332	398 116	389 577	...	258 466	243 510
11	353 101	362 966	355 482	365 915	397 363	...	278 266	258 333
12	353 271	353 449	365 521	356 304	364 926	...	298 295	278 185
13	347 474	353 629	355 224	366 153	355 837	...	313 882	298 185
14	363 743	347 551	355 517	355 060	365 485	...	329 456	313 761
15	361 873	365 206	350 736	356 128	353 928	...	351 498	329 292
16	360 509	361 197	367 608	350 530	354 636	...	363 789	351 294
17	359 332	359 998	366 294	365 725	347 275	...	387 068	363 635
18	370 949	357 441	362 271	364 649	363 613	...	376 976	386 930
19	353 121	370 050	356 784	359 763	361 573	...	385 886	376 696
20	344 178	353 665	369 548	354 307	356 396	...	392 490	385 586
21	320 938	345 047	354 731	366 861	350 406	...	357 234	392 087
22	329 505	321 929	344 673	352 686	361 746	...	350 130	356 759
23	327 443	330 376	322 407	343 939	349 189	...	358 352	349 597
24	336 697	328 389	331 868	322 201	341 636	...	339 803	357 844
25	332 455	337 935	330 453	331 402	320 005	...	339 938	339 184
26	351 914	334 073	340 440	330 192	329 028	...	330 936	339 290
27	377 976	353 376	336 926	340 503	327 839	...	348 360	330 255
28	391 793	379 356	356 339	337 397	338 452	...	343 982	347 712
29	400 022	393 096	382 508	356 848	335 578	...	336 003	343 351
30	416 588	402 092	396 821	382 963	354 842	...	327 864	335 314
31	411 682	418 652	405 776	397 323	381 030	...	340 227	327 312
32	405 736	412 778	422 398	406 593	395 382	...	336 367	339 533
33	399 299	407 149	415 598	423 441	405 244	...	334 306	335 752
34	394 509	400 239	410 205	415 760	422 063	...	310 714	333 634

Продовження таблиці 4.1

35	377 784	395 893	402 646	410 803	413 698	...	317 028	310 030
36	391 859	378 668	398 637	402 624	409 367	...	315 267	316 245
37	366 061	392 693	380 663	399 060	400 655	...	326 776	314 416
38	385 506	366 357	394 754	380 590	397 657	...	326 123	325 926
39	379 271	385 950	367 953	394 690	378 760	...	342 868	325 274
40	382 165	380 153	387 656	367 480	392 701	...	368 636	341 862
41	405 966	382 769	382 215	387 091	365 359	...	380 227	367 546
42	344 012	406 294	384 295	382 236	384 802	...	394 693	379 005
43	310 646	343 736	407 537	384 000	380 596	...	410 632	393 345
44	301 801	310 586	344 436	406 697	382 193	...	396 841	409 171
45	201 725	301 790	311 054	343 329	404 425	...	395 422	395 211
46	241 270	201 717	302 404	310 159	341 017	...	382 533	393 710
47	219 183	240 828	202 032	301 662	307 905	...	383 480	380 692
48	289 080	218 674	240 586	201 549	299 820	...	361 572	381 668
49	374 326	288 270	218 277	239 632	200 335	...	374 926	359 613
50	392 104	374 041	287 906	217 315	237 874	...	345 168	372 637
51	439 450	391 603	374 031	286 557	215 933	...	363 020	343 030
52	470 134	438 301	391 443	372 712	284 336	...	362 341	360 546
53	486 483	468 282	437 339	389 889	370 264	...	361 069	359 669
54	397 967	484 357	466 550	435 027	387 022	...	379 585	358 264
55	336 935	396 204	482 534	463 545	431 039	...	317 018	376 274
56	257 476	335 125	394 410	479 300	458 768	...	285 329	314 123
57	232 738	255 931	333 335	391 222	474 187	...	278 933	282 509
58	305 978	230 564	254 111	330 451	386 677	...	186 112	275 987
59	318 779	303 124	228 070	251 561	326 221	...	218 937	183 998
60	378 497	315 467	299 757	225 036	248 014	...	197 881	216 084
61	378 214	375 918	311 696	295 789	221 408	...	258 055	195 373
62	404 468	373 128	372 345	307 126	290 765	...	338 412	254 425
63	381 250	398 838	367 469	368 097	301 594	...	350 689	333 273
64	384 081	375 550	392 463	360 725	362 546	...	384 751	344 745
65	366 123	377 590	368 766	384 930	352 670	...	403 429	377 850
66	361 951	359 213	370 156	361 153	375 850	...	413 535	395 363
67	319 453	353 847	351 539	361 355	352 061	...	333 646	404 580
68	294 555	311 860	344 906	342 626	351 172	...	276 796	325 692
69	251 871	286 628	303 075	334 709	332 133	...	206 733	269 299
70	213 251	244 720	277 678	293 419	322 754	...	178 571	200 596
71	219 843	206 069	236 840	267 749	282 282	...	232 299	172 621
72	187 088	211 005	198 123	227 900	256 215	...	236 598	224 009
73	146 227	180 936	201 636	189 488	217 843	...	288 470	227 460
74	150 659	138 758	173 909	191 200	179 803	...	260 090	276 373
75	177 005	143 086	131 008	165 732	179 504	...	270 842	247 399
76	236 197	167 408	134 712	122 656	156 439	...	246 746	256 208
77	197 261	222 325	156 912	125 665	113 371	...	234 761	232 157

Продовження таблиці 4.1

78	223 593	184 061	206 922	145 292	115 971	...	213 403	219 293
79	173 689	207 501	170 097	190 419	132 742	...	194 286	197 753
80	156 635	159 028	190 087	154 964	172 183	...	160 780	178 143
81	139 095	142 262	143 785	171 472	138 878	...	135 504	145 871
82	118 094	124 953	126 972	128 330	152 081	...	110 573	121 590
83	104 507	104 795	110 537	111 387	112 094	...	83 420	98 291
84	90 432	91 548	91 002	95 549	95 618	...	71 386	73 567
85	74 721	78 309	78 453	77 389	81 128	...	66 303	61 310
86	59 186	63 699	66 348	65 847	64 372	...	36 239	56 357
87	50 287	49 516	53 056	54 654	53 872	...	35 073	30 362
88	40 356	41 355	40 617	42 887	43 953	...	35 749	28 677
89	22 967	32 665	33 043	32 024	33 931	...	40 542	28 517
90	27 573	17 593	25 547	25 521	24 426	...	27 844	31 979
91	8 073	20 381	13 499	19 324	19 392	...	28 327	21 164
92	8 247	4 987	15 406	10 367	14 198	...	14 573	21 958
93	4 922	5 072	3 256	11 843	7 971	...	12 851	10 343
94	4 500	2 790	3 332	2 249	9 210	...	6 410	9 226
95	3 753	2 702	1 730	2 329	1 641	...	4 582	4 094
96	3 028	2 579	1 783	1 168	1 717	...	3 321	2 874
97	2 880	2 292	1 877	1 276	853	...	2 158	2 122
98	2 794	2 549	1 846	1 440	966	...	1 398	1 397
99	2 200	2 744	2 311	1 537	1 141	...	824	909
100 +	3 549	3 039	2 766	2 132	1 313	...	1 489	1 507

Таблиця 4.2 – Розподіл постійного населення на початок року, жінки

	1991	1992	1993	1994	1995	...	2004	2005
0	4 100	3 824	3 695	3 485	3 355	...	1 848	1 890
1	498	471	490	458	470	...	222	213
2	288	263	285	253	232	...	112	119
3	233	235	199	204	204	...	74	76
4	221	201	194	199	180	...	78	75
5	167	175	161	176	157	...	71	68
6	163	131	150	153	140	...	65	74
7	150	126	129	149	137	...	70	50
8	123	128	113	104	111	...	59	55
9	119	116	124	104	111	...	54	54
10	113	98	98	101	99	...	69	50
11	95	102	95	115	110	...	49	63
12	110	104	110	106	119	...	49	48
13	121	101	103	129	114	...	91	79
14	137	141	136	145	135	...	90	80
15	142	185	158	169	203	...	107	109
16	181	214	206	196	206	...	135	130

Продовження таблиці 4.2

17	188	204	205	233	227	...	184	174
18	232	218	238	257	261	...	181	177
19	231	228	209	244	237	...	193	181
20	207	226	228	269	283	...	223	196
21	200	246	248	245	279	...	225	231
22	227	268	225	285	272	...	219	243
23	218	242	257	294	290	...	277	266
24	218	211	246	275	308	...	299	281
25	259	244	260	277	351	...	355	365
26	221	252	318	294	351	...	415	388
27	259	247	305	321	338	...	428	408
28	277	286	316	338	373	...	437	448
29	338	346	343	367	398	...	451	539
30	364	378	421	407	425	...	486	561
31	371	388	423	455	486	...	533	572
32	437	501	451	499	515	...	574	590
33	445	488	484	545	620	...	596	637
34	406	481	521	608	621	...	646	664
35	508	569	624	654	788	...	674	777
36	552	567	594	718	786	...	756	698
37	568	618	644	737	822	...	785	807
38	597	682	693	774	926	...	806	861
39	759	734	764	890	929	...	898	854
40	715	866	892	950	1 075	...	1 011	1 025
41	850	893	923	927	1 110	...	1 076	1 109
42	953	986	1 016	1 139	1 180	...	1 226	1 238
43	741	1 143	1 134	1 169	1 335	...	1 350	1 390
44	914	863	1 297	1 299	1 386	...	1 448	1 470
45	751	1 146	1 060	1 478	1 677	...	1 592	1 710
46	774	879	1 287	1 155	1 741	...	1 718	1 695
47	879	887	973	1 497	1 462	...	1 788	1 781
48	1 022	1 059	997	1 153	1 775	...	1 793	1 891
49	1 459	1 080	1 259	1 144	1 388	...	1 997	1 968
50	1 782	1 675	1 330	1 447	1 460	...	2 240	2 277
51	2 140	1 989	1 915	1 405	1 615	...	2 243	2 291
52	2 554	2 447	2 297	2 123	1 620	...	2 494	2 385
53	2 873	2 795	2 757	2 608	2 538	...	2 730	2 738
54	3 063	3 147	3 390	3 120	3 022	...	2 904	2 828
55	2 749	3 408	3 744	3 776	3 655	...	3 340	3 474
56	2 530	3 048	3 964	4 276	4 328	...	2 500	3 773
57	1 995	2 791	3 511	4 439	4 751	...	3 043	2 964
58	2 424	2 294	3 321	3 982	5 029	...	2 458	3 461
59	3 111	2 769	2 709	3 761	4 432	...	2 336	2 572
60	3 881	3 610	3 235	2 906	4 190	...	2 796	2 667

Продовження таблиці 4.2

61	4 648	4 328	4 029	3 508	3 278	...	2 832	2 999
62	5 248	5 191	4 947	4 538	3 846	...	4 419	3 234
63	5 707	5 663	6 006	5 531	5 026	...	5 406	4 967
64	6 236	6 591	6 580	6 547	6 210	...	6 442	5 702
65	6 531	7 009	7 466	7 482	7 314	...	7 495	6 925
66	7 280	7 266	8 142	8 150	8 252	...	8 415	8 084
67	7 868	7 922	8 637	9 021	8 842	...	8 568	8 990
68	7 174	8 519	9 119	9 317	9 650	...	7 670	9 509
69	7 420	8 121	9 846	10 133	10 228	...	6 953	8 304
70	7 032	8 281	9 172	11 076	11 031	...	5 618	7 670
71	7 255	7 716	9 324	9 985	11 803	...	6 878	6 109
72	8 160	7 696	8 664	10 269	10 844	...	8 722	7 464
73	7 035	8 773	8 727	9 365	10 866	...	10 533	9 329
74	6 483	7 420	10 052	9 540	9 915	...	12 377	11 482
75	8 331	7 053	8 482	10 764	9 985	...	13 604	13 292
76	11 724	8 832	7 611	9 050	11 082	...	14 595	14 594
77	13 686	12 635	9 998	8 296	9 336	...	15 181	15 469
78	14 753	14 483	13 886	10 415	8 437	...	15 305	15 983
79	15 241	15 264	15 849	14 626	10 751	...	15 753	16 026
80	14 699	15 641	16 570	16 720	14 641	...	15 980	16 390
81	14 711	14 963	17 013	16 814	16 352	...	13 742	16 367
82	13 145	14 847	15 449	17 115	16 490	...	13 247	14 261
83	13 067	13 227	15 704	15 561	16 539	...	11 068	12 989
84	12 808	13 150	13 749	15 079	14 686	...	10 092	10 686
85	11 263	12 225	13 094	12 932	14 317	...	10 144	9 719
86	10 317	10 978	12 289	12 253	11 922	...	7 643	9 773
87	9 306	9 507	10 698	11 206	10 870	...	6 116	7 029
88	8 217	8 401	9 345	9 413	9 992	...	6 717	5 508
89	6 527	7 415	8 022	8 078	8 232	...	8 039	5 997
90	5 689	5 846	6 958	6 743	6 926	...	7 692	7 135
91	5 114	4 896	5 156	5 290	5 444	...	6 521	6 583
92	2 993	4 206	4 271	4 165	4 319	...	5 374	5 378
93	2 762	2 356	3 688	3 179	3 161	...	3 826	4 252
94	2 117	2 014	1 914	2 771	2 530	...	3 012	3 036
95	1 648	1 562	1 659	1 415	2 214	...	2 023	2 302
96	1 277	1 284	1 222	1 252	957	...	1 407	1 378
97	966	928	985	871	766	...	1 035	1 015
98	744	721	728	689	601	...	603	648
99	596	537	536	491	381	...	390	430
100 +	1 508	1 487	1 525	1 238	1 047	...	658	753

Таблиця 4.3 – Кількість народжених за статтю, жінки

1991	307239
1992	290377
1993	270920
1994	253163
1995	239203
1996	226228
1997	215137
1998	203990
1999	188540
2000	185922
2001	182541
2002	188955
2003	198594
2004	207506
2005	206568

Таблиця 4.4 – Розподіл померлих немовлят за роком народження, жінки

	0	-1
1991	3311	789
1992	3065	761
1993	2905	791
1994	2713	772
1995	2613	740
1996	2386	697
1997	2192	607
1998	1906	494
1999	1784	485
2000	1644	393
2001	1598	386
2002	1393	348
2003	1450	308
2004	1515	333
2005	1563	327

В таблиці 4.3 в стовпці «0» – кількість немовлят, що померли в рік народження. В стовпці «-1» – кількість немовлят, що померли у віці 0 років але рік народження і смерті різні (наприклад рік народження – 1991, а рік смерті – 1992).

4.3 Результати проведених експериментів

4.3.1 Оцінка якості моделі

Для оцінки якості роботи моделі було проведено ретроспективні розрахунки очікуваної тривалості життя населення України, розподіленої за статтю. Якості базових, було використано статистичні дані, розподілені за статтю за період 1991 – 2005 (рисунок 4.2, рисунок 4.3). Від базових років здійснювався прогноз до 2011 року (рисунок 4.4, рисунок 4.5).

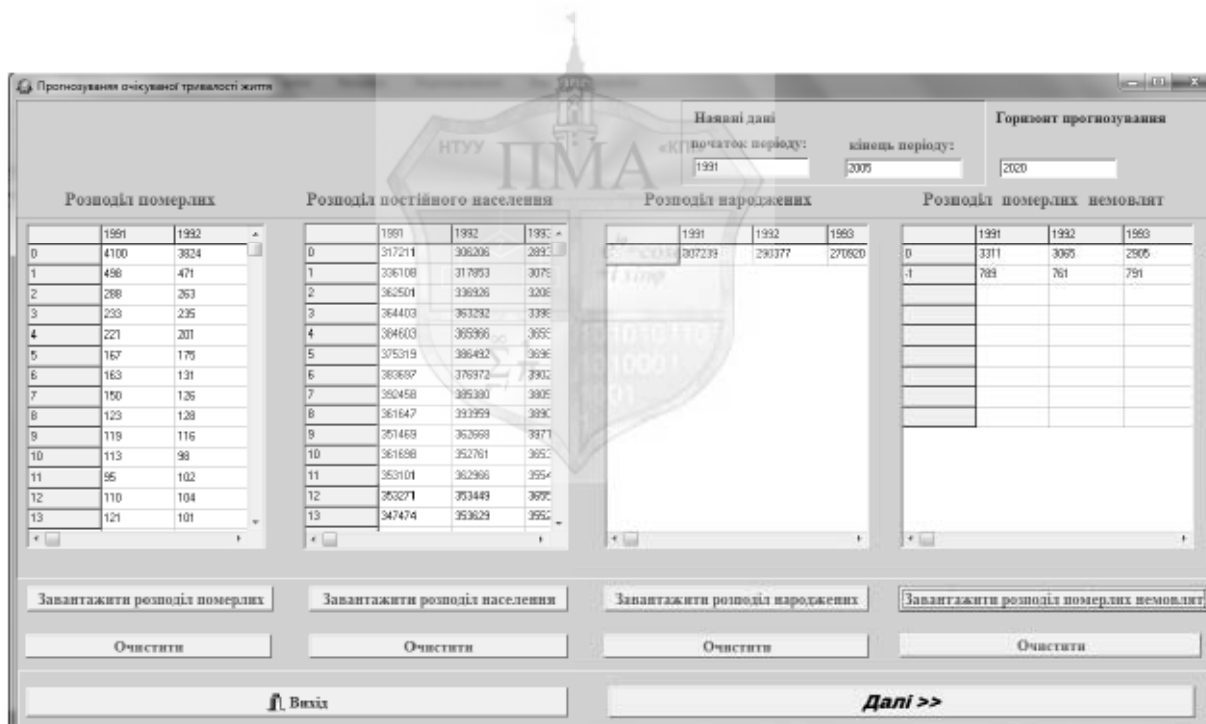


Рисунок 4.2 – Вікно з завантаженими статистичними даними для базового періоду, жінки.

Прогнозування очікуваної тривалості життя

Назви даних: початок періоду: 1991, кінець періоду: 2005, Горизонт прогнозування: 2020

	1991	1992
0	5600	5439
1	646	691
2	359	337
3	317	310
4	302	292
5	269	292
6	246	252
7	269	217
8	244	209
9	239	201
10	193	190
11	182	176
12	200	180
13	198	211

	1991	1992	1993
0	333779	323446	30461
1	352070	334189	3247
2	360915	353795	3378
3	360973	361403	3562
4	398856	382792	3837
5	388243	401783	3962
6	366552	390194	4054
7	406417	401494	3935
8	376274	408323	4045
9	363290	377838	4115
10	375371	364522	3908
11	368555	376444	3671
12	364744	366726	3796
13	357746	364807	3675

	1991	1992	1993
	304443	307241	287306

	1991	1992	1993
0	4671	4414	4488
-1	930	1024	1007

Завантажити розподіл померлих, Завантажити розподіл населення, Завантажити розподіл народжених, Завантажити розподіл померлих немовлят

Очистити, Очищення, Очищення, Очищення, Очищення

Вихід, Далі >>

Рисунок 4.3 – Вікно з завантаженими статистичними даними для базового періоду, чоловіки.

Результат та похибка

	Спрогнозовані коефіцієнти смертності m_x						Середня очікувана тривалість життя e^x					
	2006	2007	2008	2009	2010	2011	2006	2007	2008	2009	2010	2011
0	0.0089	0.0086	0.0083	0.008	0.0078	0.007	73.07	73.19	73.48	73.93	73.99	73.99
1-4	0.0006	0.0006	0.0005	0.0005	0.0004	0.000	72.72	72.82	73.09	73.53	73.57	73.57
5-9	0.0003	0.0003	0.0002	0.0002	0.0002	0.000	68.9	68.99	69.25	69.67	69.69	69.69
10-14	0.0002	0.0002	0.0002	0.0002	0.0002	0.000	63.99	64.07	64.33	64.74	64.76	64.76
15-19	0.0005	0.0004	0.0004	0.0004	0.0004	0.000	59.06	59.14	59.39	59.81	59.83	59.83
20-24	0.0007	0.0007	0.0007	0.0007	0.0007	0.000	54.19	54.27	54.52	54.93	54.95	54.95
25-29	0.0012	0.0012	0.0011	0.0009	0.0009	0.000	49.38	49.46	49.71	50.13	50.14	50.14
30-34	0.0017	0.0017	0.0017	0.0015	0.0015	0.001	44.66	44.73	44.97	45.36	45.36	45.36
35-39	0.0024	0.0024	0.0023	0.002	0.002	0.001	40.02	40.1	40.32	40.67	40.67	40.67
40-44	0.0031	0.0031	0.003	0.0027	0.0027	0.002	35.47	35.55	35.76	36.06	36.06	36.06
45-49	0.0044	0.0044	0.0042	0.0038	0.0039	0.003	30.99	31.07	31.26	31.52	31.51	31.51
50-54	0.0067	0.0066	0.0062	0.0058	0.0059	0.005	26.63	26.7	26.86	27.09	27.09	27.09
55-59	0.0105	0.0104	0.0101	0.0095	0.0095	0.009	22.44	22.51	22.63	22.8	22.81	22.81
60-64	0.0146	0.0144	0.0138	0.0132	0.0135	0.013	18.51	18.57	18.68	18.79	18.8	18.8

Середня очікувана тривалість життя

Побудувати графік коефіцієнтів смертності, Побудувати графік тривалості життя

Вихід, Зберегти m_x в Excel, Зберегти e^x в Excel

Рисунок 4.4 – Результати прогнозу на період 2006 – 2011 рр., жінки.

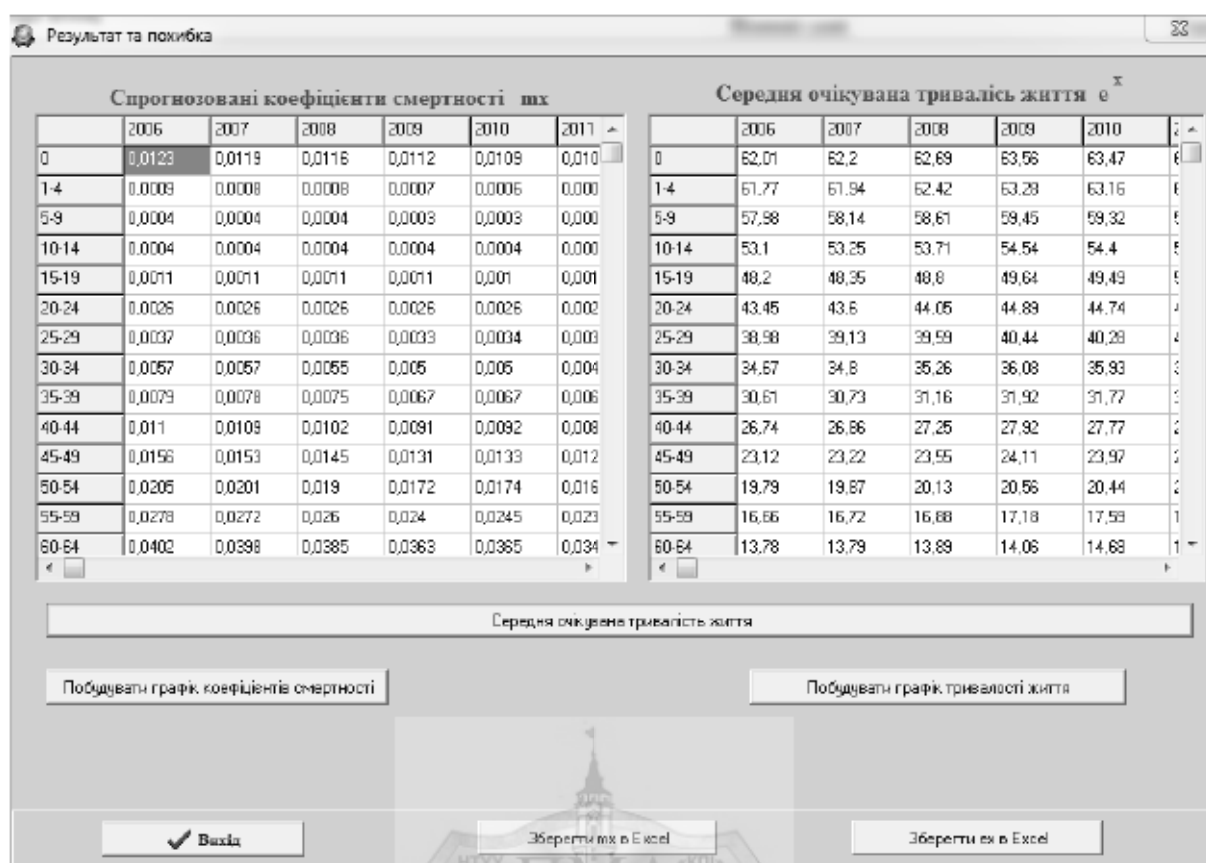


Рисунок 4.5 – Результати прогнозу на період 2006– 2011 рр., чоловіки.

Для порівняння, на рисунку 4.6 наведено динаміку прогнозованих та статистичних даних щодо очікуваної тривалості життя жінок у віці 0 за період 2006 – 2011 рр. На рисунку 4.7 – динаміку прогнозованих та статистичних даних для жінок у віці 25-29рр.

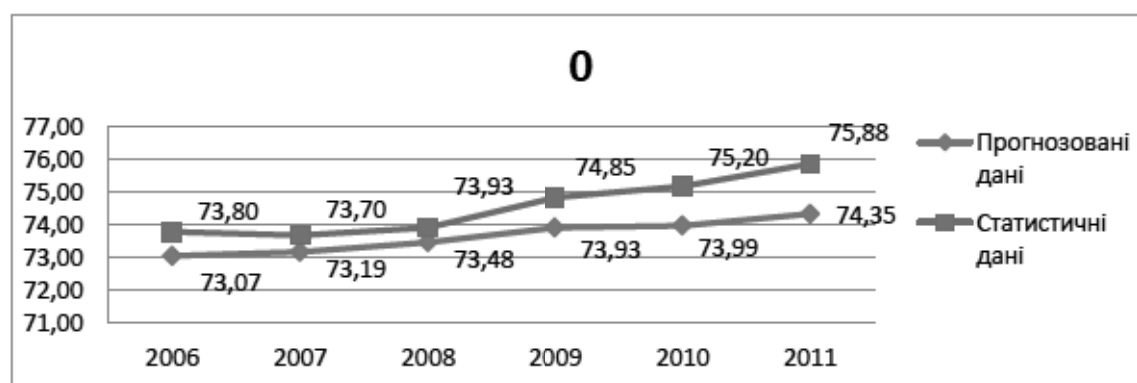


Рисунок 4.6 – Динаміка очікуваної тривалості життя жінок у віці 0 років за результатами прогнозу та статистичних даних

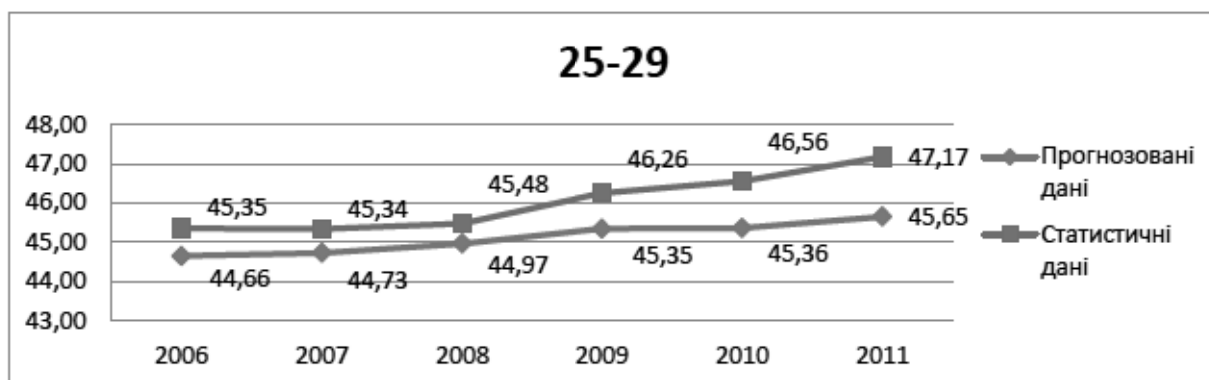


Рисунок 4.7 – Динаміка очікуваної тривалості життя жінок у віці 25-29 рр. за результатами прогнозу та статистичних даних

Для визначення похибки для кожної вікової групи, прогнозовані показники очікуваної тривалості життя порівнювались з статистичними даними. Результати наведено у табл. 4.5 та табл.4.6:

Таблиця 4.5 – Відносна похибка розрахунку на період 2006 – 2011 рр., жінки.

	2006	2007	2008	2009	2010	2011
0	1,0%	0,7%	0,6%	1,2%	1,6%	2,0%
1-4	1,0%	0,8%	0,7%	1,3%	1,6%	2,1%
5-9	1,0%	0,8%	0,7%	1,3%	1,7%	2,2%
10-14	1,1%	0,9%	0,8%	1,4%	1,9%	2,4%
15-19	1,2%	1,0%	0,8%	1,5%	2,0%	2,6%
20-24	1,3%	1,1%	0,9%	1,6%	2,2%	2,8%
25-29	1,4%	1,2%	0,9%	1,7%	2,3%	2,9%
30-34	1,5%	1,3%	1,1%	2,0%	2,6%	3,2%
35-39	1,8%	1,7%	1,5%	2,3%	2,9%	3,6%
40-44	2,1%	2,0%	1,9%	2,7%	3,3%	4,2%
45-49	2,4%	2,5%	2,4%	3,3%	3,9%	4,8%
50-54	2,8%	3,1%	3,1%	4,0%	4,4%	2,7%
55-59	3,3%	3,8%	3,9%	4,8%	5,0%	3,6%
60-64	4,0%	4,5%	4,7%	3,6%	3,6%	4,4%
65-69	2,4%	2,9%	2,9%	2,5%	3,4%	2,6%
70-74	4,3%	4,8%	4,8%	4,3%	4,8%	4,2%
75-79	2,2%	3,1%	3,4%	2,8%	3,4%	1,9%
80-84	3,8%	2,6%	2,6%	3,9%	2,7%	2,2%
85+	1,6%	2,6%	2,6%	2,3%	1,1%	3,9%

Таблиця 4.6 – Відносна похибка розрахунку на період 2006 – 2011 рр., чоловіки

	2006	2007	2008	2009	2010	2011
0	0,4%	-0,7%	-0,7%	1,2%	2,7%	2,9%
1-4	0,4%	-0,6%	-0,7%	1,2%	2,7%	2,9%
5-9	0,4%	-0,7%	-0,8%	1,2%	2,8%	3,1%
10-14	0,4%	-0,7%	-0,9%	1,3%	3,1%	3,3%
15-19	0,4%	-0,8%	-1,0%	1,4%	3,3%	3,6%
20-24	0,4%	-0,9%	-1,2%	1,4%	3,6%	3,9%
25-29	0,3%	-1,1%	-1,6%	1,2%	3,6%	3,9%
30-34	0,4%	-1,0%	-1,6%	1,3%	3,7%	4,0%
35-39	0,6%	-0,7%	-1,3%	1,7%	4,2%	4,4%
40-44	0,7%	-0,5%	-1,0%	2,0%	4,6%	5,0%
45-49	1,0%	0,1%	-0,4%	2,4%	4,9%	2,7%
50-54	1,0%	0,5%	0,1%	2,6%	4,8%	3,1%
55-59	1,8%	1,5%	1,5%	3,5%	2,5%	4,1%
60-64	2,5%	2,6%	3,4%	4,9%	1,5%	3,9%
65-69	3,0%	3,6%	4,6%	4,4%	4,5%	4,9%
70-74	2,1%	1,5%	3,0%	2,9%	3,5%	4,8%
75-79	1,6%	2,6%	4,1%	3,6%	2,9%	4,4%
80-84	1,6%	-1,4%	-1,0%	1,6%	-1,1%	4,5%
85+	1,1%	2,5%	-1,4%	-2,8%	-3,0%	4,6%

На основі табл.4.5 та табл.4.6, можна зробити висновок, що розроблена модель дала хороші результати при розрахунку очікуваної тривалості життя (відносна похибка не перевищує 5%). Тобто можна вважати, що її можна використовувати для прогнозування очікуваної тривалості життя населення України.

4.3.2 Прогнозування очікуваної тривалості життя населення України

Для обчислення перспективної очікуваної тривалості життя населення використовується інформація та прогнози Державної служби

статистики[25], Всесвітньої Організації Здоров'я[26] та Світового Банку[27] за період 1991 – 2011 рр. Горизонтом прогнозу буде 2020 рік, тобто отримуємо середньостроковий прогноз. Розробимо два типи прогнозів.

Перший враховує значення факторів для майбутніх років, базуючись на прогнозах Державної служби статистики [25], Всесвітньої Організації Здоров'я [26] та Світового Банку [27] до 2020 року (рисунки 4.8):

- ВВП скорочується на 2% щорічно до 2017 і на 3% щорічно до 2020р;
- кількість серцево-судинних захворювань зростає на 10% до 2020р;
- кількість захворювань органів дихання зменшиться на 5% до 2020р;
- легальний обіг сигарет (виробництво + імпорту - експорт) залишається на рівні 2011р.

Фактори впливу розрахункового періоду					Фактори впливу прогнозованого періоду				
	к1	к2	к3	к4		к1	к2	к3	к4
1997	3.44	1497	14129	50.47	2012	8.32	2318	13220	83.8
1998	3.43	1690	13877	51.37	2013	8.51	2256	13259	73.6
1999	3.43	1950	14495	53.15	2014	8.34	2256	13198.05	73.6
2000	3.5	2338	14639	53.6	2015	8.17	2256.01	13103.1	73.6
2001	3.82	2384	14213	69.9	2016	8.01	2256.01	13008.15	73.6
2002	4.31	2370	13372	80	2017	7.77	2256.01	12913.2	73.6
2003	4.85	2386	13895	95.7	2018	7.53	2256.02	12818.25	73.6
2004	5.25	2498	13511	105.5	2019	7.31	2256.02	12723.3	73.6
2005	5.07	2430	13894	116.2	2020	7.09	2481.6	12628.35	73.6
2006	5.479	2431	13309	116.6					
2007	7.219	2437	13946	123.5					
2008	8.046	2478	13871	124.7					
2009	8.44	2423	14528	110					
2010	7.284	2390	14995	95.8					
2011	7.713	2346	14149	87.8					

Рисунок 4.8 – Очікуване значення факторів на основі даних, Держкомстату, WHO, Світового Банку

Другий – можливий позитивний варіант розвитку країни, якщо(рисунок 4.9):

- ВВП зростає на 2% кожного року;
- кількість серцево-судинних захворювань зменшиться на 30% до 2020р.;
- кількість захворювань органів дихання зменшиться в 2 рази до 2020р.;
- легальний обіг сигарет (виробництво + імпорт - експорт) зменшиться на 30% до 2020р.

Фактори впливу розрахункового періоду					Фактори впливу прогнозованого періоду				
	x1	x2	x3	x4		x1	x2	x3	x4
1992	3,44	1497	14129	50,47	2012	8,32	2318	13220	83,8
1998	3,43	1650	13877	51,37	2013	8,51	2256	13293	73,6
1999	3,43	1950	14485	53,15	2014	8,88	2199,31	12343,5	71,5
2000	3,5	2338	14639	53,6	2015	9,05	2062,63	11394	69,39
2001	3,82	2384	14213	68,9	2016	9,03	1965,94	10444,5	67,29
2002	4,31	2370	13372	80	2017	9,21	1969,26	9496	65,19
2003	4,65	2366	13835	95,7	2018	9,39	1772,57	8545,5	63,09
2004	5,23	2498	13511	105,5	2019	9,59	1675,89	7596	60,99
2005	6,07	2430	13834	116,2	2020	9,77	1579,2	6646,5	58,88
2006	6,479	2431	13308	116,5					
2007	7,219	2437	13946	125,5					
2008	8,046	2478	13671	124,7					
2009	8,44	2423	14529	110					
2010	7,284	2380	14535	95,8					
2011	7,713	2346	14148	87,8					

Рисунок 4.9 – Можливе значення факторів, у випадку позитивного сценарію розвитку країни

Результати наведено на рисунках 4.10 – 4.11 для жінок, та на рисунках 4.12– 4.13 для чоловіків.

Результат та похибка

Спрогнозовані коефіцієнти смертності m_x						Середня очікувана тривалість життя e^x					
	2016	2017	2018	2019	2020		2016	2017	2018	2019	2020
0	0,0064	0,0061	0,0059	0,0057	0,0056	0	74,5	75,31	75,38	75,45	75,67
1-4	0,0001	0,0001	0,0001	0,0001	0,0001	1-4	73,98	74,77	74,83	74,89	75,1
5-9	0,0001	0,0001	0,0001	0,0001	0,0001	5-9	70,06	70,81	70,95	70,9	71,11
10-14	0,0002	0,0002	0,0002	0,0002	0,0002	10-14	65,14	65,83	65,87	65,92	66,12
15-19	0,0004	0,0004	0,0004	0,0004	0,0003	15-19	60,25	60,69	60,93	60,97	61,17
20-24	0,0007	0,0007	0,0007	0,0007	0,0007	20-24	55,41	55,99	56,03	56,08	56,27
25-29	0,0007	0,0007	0,0007	0,0008	0,0007	25-29	50,65	51,19	51,23	51,27	51,47
30-34	0,0012	0,0012	0,0013	0,0013	0,0013	30-34	45,86	46,36	46,41	46,46	46,65
35-39	0,0015	0,0016	0,0016	0,0016	0,0016	35-39	41,18	41,63	41,68	41,74	41,94
40-44	0,0023	0,0024	0,0024	0,0024	0,0024	40-44	36,53	36,94	37	37,06	37,26
45-49	0,0033	0,0034	0,0034	0,0035	0,0034	45-49	32,3	32,35	32,41	32,46	32,67
50-54	0,005	0,0051	0,0052	0,0053	0,0052	50-54	27,8	27,86	27,93	28,01	28,19
55-59	0,0068	0,0069	0,009	0,0091	0,0089	55-59	23,44	23,52	23,6	23,7	23,87
60-64	0,0124	0,0126	0,0127	0,0129	0,0127	60-64	19,39	19,48	19,57	19,68	19,84

Середня очікувана тривалість життя

Побудувати графік коефіцієнтів смертності

Побудувати графік тривалості життя

Вихід

Зберегти як в Excel

Зберегти як в Excel

Рисунок 4.10 – Прогноз що базується на даних Держкомстату, WHO, Світового Банку, жінки

Результат та похибка

Спрогнозовані коефіцієнти смертності m_x						Середня очікувана тривалість життя e^x					
	2016	2017	2018	2019	2020		2016	2017	2018	2019	2020
0	0,0064	0,0061	0,0059	0,0057	0,0056	0	74,75	75,68	75,95	76,05	76,24
1-4	0,0001	0,0001	0,0001	0,0001	0,0001	1-4	74,98	75,15	75,31	75,48	75,66
5-9	0,0001	0,0001	0,0001	0,0001	0,0001	5-9	71,03	71,18	71,34	71,5	71,67
10-14	0,0002	0,0002	0,0002	0,0002	0,0002	10-14	66,06	66,2	66,35	66,52	66,69
15-19	0,0004	0,0004	0,0004	0,0004	0,0003	15-19	61,12	61,26	61,41	61,57	61,74
20-24	0,0007	0,0007	0,0007	0,0007	0,0007	20-24	56,23	56,37	56,52	56,67	56,84
25-29	0,0006	0,0006	0,0005	0,0005	0,0005	25-29	51,43	51,57	51,72	51,87	52,04
30-34	0,0011	0,0011	0,001	0,001	0,001	30-34	46,57	46,7	46,85	47	47,16
35-39	0,0014	0,0013	0,0013	0,0013	0,0012	35-39	41,81	41,94	42,08	42,23	42,38
40-44	0,0022	0,0021	0,0021	0,0021	0,0021	40-44	37,08	37,2	37,34	37,48	37,63
45-49	0,0031	0,003	0,003	0,0029	0,0029	45-49	32,46	32,58	32,71	32,85	32,99
50-54	0,0046	0,0045	0,0044	0,0043	0,0042	50-54	27,92	28,04	28,16	28,3	28,44
55-59	0,0068	0,0068	0,0068	0,0069	0,0069	55-59	23,51	23,61	23,73	23,85	23,98
60-64	0,0119	0,0118	0,0117	0,0116	0,0115	60-64	19,45	19,56	19,69	19,82	19,96

Середня очікувана тривалість життя

Побудувати графік коефіцієнтів смертності

Побудувати графік тривалості життя

Вихід

Зберегти як в Excel

Зберегти як в Excel

Рисунок 4.11 – Прогноз за умови розвитку країни, чоловіки

Результат та помилка

Спрогнозовані коефіцієнти смертності m^x						Середня очікувана тривалість життя e^x					
	2016	2017	2018	2019	2020		2016	2017	2018	2019	2020
0	0,0091	0,0089	0,0095	0,0093	0,0091	0	64,74	64,52	64,29	64,05	63,95
1-4	0,0002	0,0002	0,0001	0,0001	0,0001	1-4	64,33	64,09	63,84	63,58	63,48
5-9	0,0001	0,0001	0,0001	0,0001	0,0001	5-9	60,38	60,13	59,88	59,61	59,5
10-14	0,0003	0,0003	0,0003	0,0003	0,0003	10-14	55,41	55,15	54,89	54,62	54,5
15-19	0,0009	0,0009	0,0009	0,0009	0,0009	15-19	50,49	50,23	49,97	49,69	49,57
20-24	0,0026	0,0026	0,0026	0,0025	0,0025	20-24	45,72	45,46	45,19	44,9	44,78
25-29	0,0032	0,0033	0,0033	0,0033	0,0034	25-29	41,27	41	40,73	40,45	40,32
30-34	0,0042	0,0043	0,0044	0,0044	0,0045	30-34	36,9	36,64	36,36	36,08	35,97
35-39	0,0056	0,0057	0,0058	0,0059	0,0059	35-39	32,63	32,38	32,11	31,84	31,72
40-44	0,0073	0,0075	0,0077	0,0079	0,0078	40-44	28,49	28,24	27,98	27,72	27,61
45-49	0,011	0,0113	0,0115	0,0118	0,0117	45-49	24,45	24,23	23,99	23,73	23,61
50-54	0,0145	0,0148	0,0151	0,0154	0,0152	50-54	20,7	20,49	20,25	20,02	19,89
55-59	0,0219	0,0222	0,0226	0,023	0,0228	55-59	17,06	16,87	16,66	16,43	16,25
60-64	0,0329	0,0333	0,0336	0,034	0,0339	60-64	13,75	13,56	13,35	13,12	12,92

Середня очікувана тривалість життя

Побудувати графік коефіцієнтів смертності

Побудувати графік тривалості життя

Вихід Зберегти як в Excel Зберегти як в Excel

Рисунок 4.12 – Прогноз що базується на даних Держкомстату, WHO, Світового Банку, чоловіки

Результат та помилка

Спрогнозовані коефіцієнти смертності m^x						Середня очікувана тривалість життя e^x					
	2016	2017	2018	2019	2020		2016	2017	2018	2019	2020
0	0,0091	0,0088	0,0086	0,0083	0,0081	0	65,43	65,49	65,54	65,56	65,55
1-4	0,0002	0,0002	0,0001	0,0001	0,0001	1-4	65,08	65,07	65,1	65,1	65,08
5-9	0,0001	0,0001	0,0001	0,0001	0,0001	5-9	61,08	61,12	61,13	61,13	61,1
10-14	0,0003	0,0003	0,0003	0,0003	0,0003	10-14	58,11	58,13	58,14	58,13	58,11
15-19	0,0009	0,0009	0,0009	0,0009	0,0009	15-19	51,19	51,22	51,22	51,21	51,18
20-24	0,0026	0,0026	0,0026	0,0025	0,0025	20-24	46,42	46,44	46,44	46,43	46,4
25-29	0,0031	0,0031	0,0031	0,0031	0,0031	25-29	41,99	42,01	42,01	41,99	41,96
30-34	0,0038	0,0038	0,0037	0,0036	0,0035	30-34	37,61	37,63	37,63	37,61	37,58
35-39	0,005	0,0049	0,0048	0,0047	0,0046	35-39	33,29	33,3	33,29	33,26	33,2
40-44	0,0062	0,0059	0,0057	0,0054	0,0052	40-44	29,07	29,06	29,03	28,98	28,91
45-49	0,0097	0,0095	0,0092	0,0089	0,0087	45-49	24,91	24,88	24,8	24,71	24,6
50-54	0,0131	0,0128	0,0125	0,0123	0,012	50-54	21,03	20,95	20,85	20,73	20,59
55-59	0,0207	0,0206	0,0205	0,0204	0,0202	55-59	17,28	17,17	17,04	16,88	16,7
60-64	0,031	0,0306	0,0302	0,0298	0,0294	60-64	13,89	13,76	13,61	13,43	13,22

Середня очікувана тривалість життя

Побудувати графік коефіцієнтів смертності

Побудувати графік тривалості життя

Вихід Зберегти як в Excel Зберегти як в Excel

Рисунок 4.13 – Прогноз за умови розвитку країни, чоловіки

Прогнози, отримані на основі побудованої моделі порівнюємо з прогнозованими значеннями, які були отримані ІДСД НАН України [1] (табл.4.7). Було виявлено, що Прогноз 1, побудований на основі даних щодо очікуваного значення факторів впливу Держкомстатом, WHO, Світовим Банком відповідає песимістичному варіанту прогнозу очікуваної тривалості життя ІДСД НАН України [1]. Прогноз 2, що будувався за умови розвитку країни і покращення значень факторів впливу, відповідає середньому варіанту прогнозу очікуваної тривалості життя ІДСД НАН України[1].

Таблиця 4.7 – Відносна похибка Прогнозу 1 у порівнянні з середнім варіантом прогнозу ІДСД, жінки

Песимістичний		Прогноз 1		Похибка	
2015	2020	2015	2020	2015	2020
72,9	73,3	73,67	75,67	1,0%	3,1%
58,9	59,3	59,60	61,17	1,2%	3,1%
31,2	31,4	31,59	32,67	1,2%	3,9%
19	19,2	18,92	19,84	-0,4%	3,2%

Таблиця 4.8 – Відносна похибка Прогнозу 1 у порівнянні з середнім варіантом прогнозу ІДСД, чоловіки

Песимістичний		Прогноз 1		Похибка	
2015	2020	2015	2020	2015	2020
61,2	61,7	63,89	63,96	4%	4%
47,3	47,8	49,69	49,57	5%	4%
22,5	22,8	23,53	23,61	4%	3%
13,8	13,9	13,91	13,13	1%	-6%

Таблиця 4.9 – Відносна похибка Прогнозу 2 у порівнянні з середнім варіантом прогнозу ІДСД, жінки

Середній ІДСД		Прогноз 2		Похибка	
2015	2020	2015	2020	2015	2020
75	76	75,34	76,24	0,5%	0,3%
60,9	61,9	60,99	61,74	0,1%	-0,3%
32,5	33,4	31,70	32,99	-2,5%	-1,2%
19,8	20,5	18,95	19,96	-4,5%	-2,7%

Таблиця 4.10 – Відносна похибка Прогнозу 2 у порівнянні з середнім варіантом прогнозу ІДСД, чоловіки

Середній ІДСД		Прогноз 2		Похибка	
2015	2020	2015	2020	2015	2020
64	65,9	65,35	65,55	2%	-1%
50	51,9	51,16	51,18	2%	-1%
24,4	25,6	24,94	24,60	2%	-4%
15	15,7	14,23	14,88	-5%	-6%

На основі отриманого прогнозу можна зробити висновок, що очікувана тривалість життя населення скоротиться до 2020 року. Це можна пояснити значною втратою демографічного потенціалу, зокрема низькою тривалістю життя та високим рівнем смертності. Проте, це можна змінити. Потрібно покращувати соціальну та економічну політику в державі, розвивати медицину та науку. Зокрема:

- припинити воєнні дії на території нашої держави;
- підвищити тривалість життя населення за рахунок реалізації резервів зниження запобіжної смертності від неінфекційної хронічної патології (хвороби системи кровообігу, дихання, новоутворення та інші), в першу чергу, в молодому віці;
- збільшити тривалість здорового (активного) життя шляхом профілактики неінфекційної патології у підлітків та молоді;
- поліпшити якість життя хронічно хворих та інвалідів шляхом надання їм умов для реалізації наявного (залишкового) потенціалу здоров'я;
- збільшити народжуваність та створити більш сприятливу соціальну обстановку по відношенню до сімей з дітьми, поліпшити якість життя батьків, щоб було легше поєднувати виконання сімейних та професійних обов'язків;
- збільшити доходи населення, зокрема шляхом розвитку української економіки та створенням нових робочих місць.

Висновки до розділу 4

З метою оцінки якості та коректності роботи моделі було виконано ретроспективні розрахунки очікуваної тривалості життя залежно від найбільш впливових факторів. Запропонована модель дає хороші результати при розрахунку у всіх вікових групах, для кожної статі окремо (відносна похибка не перевищує 5%).

Також було створено два варіанти прогнозу очікуваної тривалості життя до 2020р. Результати порівнювалися із прогнозом ІДСД НАН України [1].

Отримані результати підкреслюють складність демографічної ситуації в Україні та необхідність проведення чіткої державної демографічної політики на різних рівнях. Саме тому було сформовано список рекомендацій щодо соціально-економічної політики держави, спрямовані на покращення в майбутньому очікуваної тривалості життя населення України.

ВИСНОВКИ

Дана магістерська дисертація присвячена прогнозуванню очікуваної тривалості життя населення України залежно від найбільш впливових факторів. Наразі, ця задача дуже актуальна не тільки в демографії, адже знання про очікувану тривалість життя лежать в основі багатьох спеціальностей, що займаються підрахунком ризиків, зокрема в страхуванні, яке останні декілька років надзвичайно активно розвивається в Україні. До того ж Україна знаходиться в стані демографічної кризи, для подолання якої потрібно проводити правильну соціально-економічну політику, зокрема правильно визначати видатки держави на фінансування соціальних виплат, освітніх та медичних закладів, пенсій, тощо.

У роботі розглянуто найпоширеніші моделі, що використовуються в демографічному прогнозуванні процесів смертності та зроблено їх порівняльний аналіз. Запропоновано нову модель прогнозування очікуваної тривалості життя населення України, яка враховує вплив соціально-економічних факторів на показник дожиття.

На основі цих досліджень спроектовано та реалізовано програмне забезпечення за допомогою якого на основі ретроспективних розрахунків очікуваної тривалості життя населення України, розподіленої за статтю та віком, було підтверджено адекватність моделі.

За допомогою цього ПЗ отримано два варіанти середньострокового прогнозу очікуваної тривалості життя населення України до 2020 року. За допомогою яких було виявлено, що без впровадження заходів щодо покращення рівня економіки, соціальної політики та медицини, сценарій для України є дуже несприятливим: тривалість життя Українців буде постійно зменшуватись і становитиме при народженні 64 роки для чоловіків та 75 років для жінок. Проте, це менше, ніж в розвинених країнах Європи та лідером по тривалості життя населення – Японії (за даними ООН).

На основі прогнозних оцінок розроблено рекомендації щодо стабілізації демографічної ситуації в Україні. Які, в першу чергу, стосуються охорони здоров'я, соціальної захищеності людей та зайнятості, тобто факторів, що безпосередньо впливають на демографічну ситуацію.

Практична значимість роботи підтверджується тим, що результати роботи впроваджено у відділі демографічного моделювання та прогнозування Інституту демографії та соціальних досліджень імені М. В. Птухи НАН України для прогнозів на майбутнє і для ретроспективи динаміки відтворення населення.



СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Комплексний демографічний прогноз України на період до 2050 р. (колектив авторів)/ за ред. чл.-кор. НАНУ, д.е.н., проф. Е.М. Лібанової. – К.: Український центр соціальних реформ, 2006. – 138 с.
2. Банк даних Першого всеукраїнського перепису населення. [Електронний ресурс]. – Режим доступу: http://database.ukrcensus.gov.ua/MULT/Database/Census/databasetre_e_uk.asp
3. Медков В.М. Демографія: Учебное пособие. Серия "Учебники и учебные пособия". - Ростов-на-Дону: "Феникс", 2002. – 448 с.
4. Стеценко С. Г. Демографічна статистика: Підручник. – К.: Вища шк., 2005. – 415 с.
5. Пузиков О. С. Социально-экономическое прогнозирование. Курс лекций. – Ростовский государственный строительный университет, 2000. – 124 с.
6. SAS/ETS(R) 9.2 User's Guide. [Електронний ресурс]. – Режим доступу: https://support.sas.com/documentation/cdl/en/etsug/60372/HTML/default/viewer.htm#etsug_intro_sect039.htm
7. Программа статистического анализа Statistica. Обзор. [Електронний ресурс]. – Режим доступу: <http://bourabai.kz/tpoi/statistica/>
8. Aptech products. The GAUSS™ Mathematical and Statistical System. [Електронний ресурс]. – Режим доступу: <http://www.aptech.com/products/gauss-mathematical-and-statistical-system/>

9. Электронный журнал Демоскоп [Электронный ресурс]. – 2010. Режим доступа: <http://demoscope.ru>
10. Про комп'ютерну програму «Система політичних моделей «Спектрум» [Электронный ресурс]. – 2007. Режим доступа: http://data.unaids.org/pub/Manual/2007/demproj_2007_ru.pdf
11. HPN Technologies Inc. [Электронный ресурс]. – Режим доступа: <http://www.hpntech.com/fivfiv/>
12. Громова Н.М., Громова Н.И. Основы экономического прогнозирования. – М.: Акад. естествознания, 2007. – 112 с.
13. Костромин А. В. Конспект лекций по курсу «Эконометрика». Часть 2. Казанский государственный финансово – экономический институт. – Казань 2004. – 47 с.
14. Модель смертности Брасса. Материалы конференции «Демографическое развитие Дальнего Востока» [Электронный ресурс]. – Режим доступа: <http://www.demographia.ru/node/333>
15. Rob J. Hyndman, Leonie Tickle, Booth H. Demographic research: Article 9, Rostock, GERMANY, 2006. – P.289 – 310.
16. Jenny Zheng Wang. Fitting and Forecasting Mortality for Sweden: Applying the Lee-Carter Model // Mathematical Statistics Stockholm University Examensarbete – 2007 – 54p.
17. Четыркин Е. М. Статистические методы прогнозирования. Изд. 2-е. перераб. и доп. М. «Статистика». 1977. – 200 с
18. Моделювання законів смертності [Электронный ресурс]. – Режим доступа: <http://edu2.tsu.ru/res/1668/text/page53.html>
19. Методика побудови таблиць смертності та середньої очікуваної тривалості життя для населення України та регіонів за статтю та типом поселення / Відповідальний за випуск Тимошенко Г.М. 2011 – 32с.
20. Методы и модели анализа временных рядов : методические указания к лабораторным работам / сост. С.И. Татаренко. –

Тамбов : Изд-во Тамбовского Государственного Технического университета, 2008. – 32с.

21. Олефір О. С., Савчук О. Б. Визначення найбільш впливових факторів для прогнозування очікуваної тривалості життя населення України // Прикладна математика та комп'ютинг ПМК – 2015: зб. тез доп. / редкол.: І. А. Дичка [та ін.]. — К. : Просвіта, 2015. — С. 217 — 221.
22. Zheng Wang J. Fitting and forecasting mortality for Sweden: applying the Lee-Carter model – Stockholm, 2007. – P.8-47.
23. Шевчук П.Є. Застосування формальних методів до прогнозування смертності населення // Науково-економічний та суспільно-політичний журнал «Демографія та соціальна економіка» – К. Інститут демографії та соціальних досліджень імені М.В. Птухи НАН України, № 2(16)2011. – С.73 - 82
24. Олефір О.С., Савчук О.Б. Порівняння впливу різних факторів на очікувану тривалість життя населення України // Системний аналіз та інформаційні технології: матеріали Міжнародної науково-технічної конференції SAIT 2015, Київ, 22 – 25 червня 2015 р. / ННК «ІПСА» НТУУ «КПІ». — К. : ННК «ІПСА» НТУУ «КПІ», 2015. — С. 175.
25. Населення України за 2012 рік. Демографічний щорічник. — К. : Державна служба статистики України, 2013. — 442 с.
26. Global Health Observatory (GHO) data [Електронний ресурс]. – Режим доступу: <http://www.who.int/gho/en/>
27. World Bank Open Data [Електронний ресурс]. – Режим доступу: <http://data.worldbank.org/country/ukraine>