

Міністерство освіти і науки України
Національний університет «Києво-Могилянська академія»
Факультет інформатики
Кафедра математики

Кваліфікаційна робота

освітній ступінь – бакалавр

на тему: «ПОРІВНЯЛЬНИЙ АНАЛІЗ ТРИВАЛОСТІ ЖИТТЯ ТА ВАРТОСТІ
СТРАХУВАННЯ ЖИТТЯ ДЛЯ РІЗНИХ ГРУП ІНДИВІДІВ»

Виконала: студентка 4-го року навчання,
Спеціальності

113 Прикладна математика

Яндола Наталія Олександрівна

Керівник Братик М. В., ст. викладач, к.н.

Рецензент _____

Кваліфікаційна робота захищена

з оцінкою _____

Секретар ЕК _____

«____» _____ 2024 р.

Київ – 2024

Міністерство освіти і науки України
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «КИЄВО-МОГИЛЯНСЬКА АКАДЕМІЯ»
Кафедра математики факультету інформатики

ЗАТВЕРДЖУЮ

Зав. кафедри математики,

Доцент, кандидат ф.-м. наук

_____ Р.К. Чорней

« ____ » _____ 2024 р.

ІНДИВІДУАЛЬНЕ ЗАВДАННЯ

на кваліфікаційну роботу

студентці 4-го курсу факультету інформатики Яндолі Наталії Олександрівні

Тема: Дослідження залежності тривалості життя та вартості страхування життя від різних груп індивідів.

Вихідні дані: Досліджено залежність тривалості життя та вартість страхування життя від різних груп індивідів.

Зміст ТЧ до кваліфікаційної роботи:

1. Анотація
2. Вступ
3. Розділ 1. Тривалість життя індивіда – основні поняття та дослідження впливу сторонніх факторів
4. Розділ 2. Страхування життя – види, специфіка та розрахунки
5. Розділ 3. Тривалість життя та страхування життя у воєнні та після воєнні часи – аналіз джерел та потенційний вплив на ситуацію в Україні
6. Висновки
7. Список літератури

Дата видачі: «22» жовтня 2023 р.

Керівник _____

Завдання отримано _____

Тема: Дослідження залежності тривалості життя та вартості страхування життя від різних груп індивідів.

Календарний план виконання роботи:

| № п/п | Назва етапу кваліфікаційної роботи | Термін виконання етапу | Примітка |
|-------|---|------------------------|----------|
| 1. | Отримання завдання на кваліфікаційну роботу. | Жовтень 2023 | |
| 2. | Ознайомлення з літературою за темою роботи | Листопад-грудень 2024 | |
| 3. | Написання теоретичної частини роботи | Березень-квітень 2024 | |
| 4. | Практична частина роботи – власне дослідження та опис результатів | Квітень-травень 2024 | |
| 5. | Підготовка до попереднього захисту кваліфікаційної роботи | Травень 2024 | |
| 6. | Дороблення практичної частини роботи та коригування роботи за результатами попереднього захисту | Травень 2024 | |
| 7. | Здача кваліфікаційної роботи | Червень 2024 | |

Студент: Яндола Н.О.

Керівник: Братик М.В.

«22» жовтня 2023 р.

Зміст

| | |
|---|----|
| Зміст | 4 |
| Анотація..... | 5 |
| Вступ..... | 6 |
| Розділ 1. Тривалість життя індивіда – основні поняття та дослідження впливу сторонніх факторів | 8 |
| 1.1 Основні теоретичні матеріали та позначення | 8 |
| 1.2 Таблиці тривалості життя | 13 |
| 1.3 Залежність тривалості життя від сторонніх факторів | 15 |
| 1.4 Метод Каплана-Мейєра для дослідження смертності населення | 37 |
| Розділ 2. Страхування життя – види, специфіка та розрахунки | 40 |
| 2.1 Загальна інформація про страхування життя та основні терміни..... | 40 |
| 2.2 Основні види страхування життя з виплатою страхової суми в кінці року смерті застрахованої особи | 41 |
| 2.3 Основні види страхування життя з виплатою в момент смерті застрахованої особи | 45 |
| 2.4 Розрахування нетто-премій на основі таблиць тривалості життя та аналіз результатів | 47 |
| Розділ 3. Тривалість життя та страхування життя у воєнні та після воєнні часи – аналіз джерел та потенційний вплив на ситуацію в Україні | 52 |
| Висновки..... | 54 |
| Список використаних джерел..... | 55 |

Анотація

У даній кваліфікаційній роботі здійснено порівняльний аналіз тривалості життя та вартості страхування життя для різних груп індивідів. В рамках дослідження було розглянуто ключові чинники, що впливають на тривалість життя, зокрема економічні, соціальні, та медичні аспекти. Використано методи статистичного аналізу та машинного навчання для дослідження даних про смертність населення. Продемонстровано зв'язок між впливом даних чинників на тривалість життя та вартістю страхування життя для індивіда. Висунуті припущення щодо впливу поточних подій в Україні на тривалість життя населення країни та вартість страхових послуг.

Ключові слова: тривалість життя, страхування життя, порівняльний аналіз, регресійний аналіз, кореляція, таблиці тривалості життя, нетто-премії.

Вступ

Вибір теми даної роботи зумовлений необхідністю глибшого розуміння залежності тривалості життя та вартості страхування життя від різних факторів, що можуть мати негативний або позитивний вплив на різні групи населення. У сучасному світі актуальність цього питання стає все більш помітною у зв'язку з багатьма глобальними викликами, та особливо підкреслюється поточною геополітичною ситуацією у світі та зокрема в Україні. Пандемія значно вплинула на тривалість життя в різних країнах, особливо в умовах нестачі медичних ресурсів і нерівномірного доступу до вакцинації. Збройні конфлікти по всьому світу не тільки призводять до прямої втрати життя, але й мають довготривалий вплив на психічне та фізичне здоров'я населення, руйнують інфраструктуру охорони здоров'я, що в свою чергу ускладнює доступ до медичних послуг. Глобальні зміни клімату спричиняють екстремальні погодні явища, що призводить до збільшення захворюваності та смертності, особливо серед вразливих груп населення. Усі ці фактори взаємодіють між собою, створюючи складну мережу впливів, які необхідно детально досліджувати.

Об'єктом дослідження є індивіди різних соціально-економічних та регіональних груп та їх тривалість життя. Предметом дослідження є залежність вартості страхування життя від тривалості життя індивідів та факторів, що на неї впливають.

Метою даної роботи є визначення ключових факторів, що впливають на тривалість життя, та розрахування вартості страхування життя для різних груп індивідів. Для досягнення поставленої мети було визначено наступні завдання:

- Провести аналіз літератури та визначити важливість впливу очікуваної тривалості життя на вартість страхування життя індивіда;

- Зібрати та проаналізувати статистичні дані щодо тривалості життя населення в різних країнах та дослідити значущість впливу належності індивіда до різних груп на очікувану тривалість життя;
- Дослідити вагомість впливу різних чинників на тривалість життя за допомогою методів математичного навчання;
- Навести приклад розрахування вартостей страхування життя для різних груп населення;
- Додаткову увагу приділити аналізу джерел, що досліджують вплив збройних конфліктів на очікувану тривалість життя населення та на ситуацію у сфері страхування, і зробити припущення щодо впливу поточних подій в Україні на тривалість життя населення та страхування життя.

Розділ 1. Тривалість життя індивіда – основні поняття та дослідження впливу сторонніх факторів

1.1 Основні теоретичні матеріали та позначення

Аналіз теоретичного матеріалу за темою дослідження у розділах 1.1-1.2 та 2.2-2.3 виконано на основі підручників «Life insurance mathematics» Ганса Гербера [1] та «Basic Life Insurance Mathematics» Рагнара Норберга [2].

Для початку розглянемо тему тривалості майбутнього життя індивідів. Розглянемо індивіда віку x та позначимо тривалість його майбутнього життя як T або більш експліцитно – $T(x)$. Таким чином, сума $x + T$ відобразить вік, у якому цей індивід помер. Запишемо функцію розподілу випадкової величини T :

$$G(t) = P(T < t), t \geq 0$$

де, $G(t)$ – функція що визначає ймовірність смерті даного індивіда до закінчення t років. Припускається, що розподіл G випадкової величини T відомий і функція G неперервна та величина T має щільність розподілу $g(t) = G'(t)$. Тоді маємо:

$$g(t)dt = P(t < T < t + dt)$$

- ймовірність смерті індивіда у нескінченно малому проміжку часу між $x + t$ та $x + t + dt$.

Введемо деякі додаткові стандартизовані позначення:

Ймовірність, що особа віку x помре протягом t років:

$${}_tq_x := G(t)$$

Відповідно, ймовірність того що особа віку x проживе як мінімум t років є ймовірністю того що особа не помре протягом цього часу і позначається наступним чином:

$${}_tp_x := 1 - G(t)$$

Також визначимо ймовірність того, що особа віку x років проживе s років і після цього помре протягом t років:

$${}_s|tq_x := P(s < T < s + t) = G(s + t) - G(s) = {}_{s+t}q_x - {}_sq_x$$

Позначимо умовну ймовірність того, що індивід після досягнення віку $x + s$ проживе t років:

$${}_tp_{x+s} := P(T > s + t | T > s) = \frac{1 - G(s + t)}{1 - G(s)}$$

Аналогічно до цього визначення можемо вивести умовну ймовірність того, що індивід після досягнення віку $x + s$ помре протягом часу t :

$${}_tq_{x+s} := P(T \leq s + t | T > s) = \frac{G(s + t) - G(s)}{1 - G(s)}$$

З даних означень можна записати альтернативний варіант визначення ${}_s|tq_x$:

$$\begin{aligned} {}_s|tq_x &:= P(s < T < s + t) = G(s + t) - G(s) = [1 - G(s)] \frac{G(s + t) - G(s)}{1 - G(s)} \\ &= {}_sp_{xt}q_{x+s} \end{aligned}$$

А також через співвідношення записати ймовірність того, що особа віку x проживе $s + t$ років:

$${}_{t+s}p_x := 1 - G(s + t) = [1 - G(s)] \frac{1 - G(s + t)}{1 - G(s)} = {}_sp_{xt}p_{x+s}$$

Для визначення очікуваної тривалості залишку життя спочатку введемо означення математичного сподівання та дисперсії випадкової величини:

Математичне сподівання – це числова характеристика, що характеризує середнє значення випадкової величини. Якщо деяка випадкова величина X є неперервною, приймає значення на проміжку $[a, b]$ та має щільність розподілу ймовірностей позначену як $f(x)$, то її математичне сподівання розраховується за формулою:

$$E(X) = \int_a^b xf(x)dx$$

Дисперсія — це міра розсіювання значень випадкової величини відносно її середнього значення.

Очікувана тривалість залишку життя для особи віку x є математичним сподіванням величини T та у міжнародному актуарному товаристві позначається як:

$$E(T) = e_x^\circ = \int_0^\infty t g(t) dt$$

або

$$e_x^\circ = \int_0^\infty (1 - G(t)) dt = \int_0^\infty {}_t p_x dt$$

Значення $D(T)$ обраховується наступним чином:

$$D(T) = E(T^2) = \int_0^\infty t^2 {}_t p_x \mu_{x+t} dt = \int_0^\infty 2t {}_t p_x dt,$$

де, μ_{x+t} – сила смертності.

В актуарній науці сила смертності є ефективним інструментом для індикації рівня смертності населення який є миттєвою мірою ймовірності смерті в конкретний момент часу за умови дожиття до цього моменту [4]. Сила смертності для індивіда у віці x визначається як миттєва ймовірність смерті у віці x . Для індивіда віку $x+t$ сила смертності визначається за даними формулою:

$$\mu_x(t) := \mu_{x+t} := \frac{g(t)}{1 - G(t)} = -\frac{d}{dt} \ln(1 - G(t))$$

$$\mu_x(t) := \mu_{x+t} := -\frac{d}{dt} \ln {}_t p_x$$

Тепер можна записати альтернативний варіант розрахунку ймовірності смерті індивіда у проміжок часу від t до $t + dt$:

$$g(t) dt = P(t < T < t + dt) = {}_t p_x \mu_{x+t} dt$$

та очікуваної тривалості життя:

$$e_x^\circ = \int_0^\infty t {}_t p_x dt$$

Після інтегрування $\mu_x(t)$ отримаємо нове визначення для ${}_t p_x$:

$${}_t p_x = e^{-\int_0^t \mu_{x+s} ds}$$

Функція G може вважатись аналітичним законом розподілу за умов що вона виражена простою формулою. Впродовж років було запропоновано різні варіанти аналітичних розподілів для функції, наприклад:

- Розподіл Де Муавра

Припускається існування деякого максимального віку людини ω і стверджується про рівномірність розподілу T у проміжку значень від 0 до $\omega - x$, отже:

$$g(t) = \frac{1}{\omega - x}$$

$$\mu_x(t) = \frac{1}{\omega - x - t}$$

- Розподіл Гомпертца

Припускається що сила смертності зростає експоненційно у вигляді $\mu_{x+t} = Bc^{x+t}$, $t > 0$, де B і c – сталі значення, і $B > 0$, $c \geq 1$.

Даний підхід краще відображає закон старіння, ніж закон Де Муавра враховуючи специфіку різниці у віці індивідів та позбавляє необхідності робити припущення щодо максимального віку осіб ω .

- Розподіл Гомпертца – Мейхема (узагальнений Мейхемом закон Гомпертца)

Стверджується, що рівень смертності людини є сумою вікової складової (функція Гомперца), яка експоненціально зростає з віком, і незалежної від віку складової (термін Мейхема). У захищеному середовищі, де зовнішні причини смерті є рідкісними (лабораторні умови, країни з низькою смертністю тощо), незалежна від віку складова смертності часто є незначущою. У цьому випадку формула спрощується до закону смертності Гомперца. Цей закон успішно використовується в біології, актуарній науці та демографії для опису закономірностей смертності багатьох видів (включно з людьми), визначення політики у

страхуванні, створення актуарних таблиць та моделей зростання. Закон має наступний вигляд:

$$\mu_{x+t} = A + Bc^{x+t}, t > 0,$$

де, A – постійна додатна компонента, що не залежить від віку індивіда.

Ймовірність виживання в моделі Мейхема виглядає наступним чином:

$${}_t p_x = e^{(-At - \frac{B}{\ln(c)} c^x (c^t - 1))}$$

Детально розподіл Гомпертца – Мейхема було розглянуто в роботі “On the Gompertz–Makeham law: A useful mortality model to deal with human mortality” (Кастелларес, Патрісіо та Лемонте, 2022) [3], де було виведено деякі структурні властивості моделі в статистиці, демографії та актуарних науках, а також на основі реальних статистичних даних про смертність обчислено очікувану тривалість життя для кількох країн і доведено, що модель, особливо при використанні розподілу Белла, дає належні результати для практичних розрахунків людської смертності і є надійною для використання.

Введемо деякі додаткові поняття що стосуються тривалості майбутнього життя індивіда.

$K = [T]$ – кількість повних років життя індивіда до настання моменту смерті.

Розподіл для величини буде мати наступний вигляд:

$$P(K = k) = P(k \leq T < k + 1) = {}_k p_x \cdot q_{x+k}$$

Математичне сподівання $E(K)$ або ж e_x позначає очікувану обмежену тривалість майбутнього життя:

$$e_x = \sum_{k=1}^{\infty} k P(K = k) = \sum_{k=1}^{\infty} k \cdot {}_k p_x \cdot q_{x+k}$$

або ж:

$$e_x = \sum_{k=1}^{\infty} P(K \geq k) = \sum_{k=1}^{\infty} k p_x$$

S – випадкова величина неперервно розподілена між 0 та 1 що позначає дробову частину року смерті індивіда, під час якої той був живий. Припускаючи що значення S наближається до $\frac{1}{2}$, отримаємо більш практичну формулу для розрахунку очікуваної тривалості життя:

$$e_x^{\circ} = e_x + \frac{1}{2}$$

1.2 Таблиці тривалості життя

Згадані в попередньому пункті 1.1 актуарні таблиці, або ж таблиці смертності (іноді таблиці тривалості життя) - таблиці, які показують ймовірність того, що людина у певному віці помре до свого наступного дня народження. Їх часто використовують компанії зі страхування життя для розрахунку очікуваної тривалості життя для людей різного віку та категорій населення, а також ймовірності індивіда дожити до певного року життя. Окрім математичного методу побудови таблиць смертності, вони також можуть будуватись на основі статистичних даних.

Ключовими параметрами при складанні таблиці смертності є наступні колонки [5]:

- x – позначає вік або вікову категорію досліджуваної групи;
- l_x – позначає кількість осіб з досліджуваної групи що є живими у віці x .

Маючи ці ключові параметри, можна розрахувати наступні важливі показники виживання та смертності для групи:

- d_x – позначає кількість осіб, що померли в проміжок від x до $x + 1$, що є рівним різниці між кількістю живих осіб певного вікового класу і наступним віковим класом і розраховується за наступною формулою:

$$d_x = l_x - l_{x+1}$$

• q_x – позначає ймовірність що індивід віку x помре протягом одного наступного року і розраховується за наступною формулою:

$$q_x = \frac{d_x}{l_x}$$

• s_x – позначає ймовірність доживання до віку x і розраховується за наступною формулою:

$$s_x = \frac{l_x}{l_0}$$

• L_x – позначає середню кількість осіб, що були живими в момент останнього дня народження для віку x (або ж в проміжку від x до $x + 1$) в будь-який момент часу і розраховується за наступною формулою:

$$L_x = l_x - \frac{d_x}{2}$$

• T_x – позначає кількість осіб, що були живими в момент останнього дня народження для віку x та пізніше в будь-який момент часу і розраховується за наступною формулою:

$$T_x = \sum_{n=x}^k L_n$$

• e_x – позначає очікувану тривалість майбутнього життя та за умови відомих значень l_x і розраховується за наступною формулою:

$$e_x = \frac{l_{x+1} + l_{x+2} + l_{x+3} + \dots}{l_x}$$

що, виходячи з вище наведених формул, є еквівалентним наступному запису:

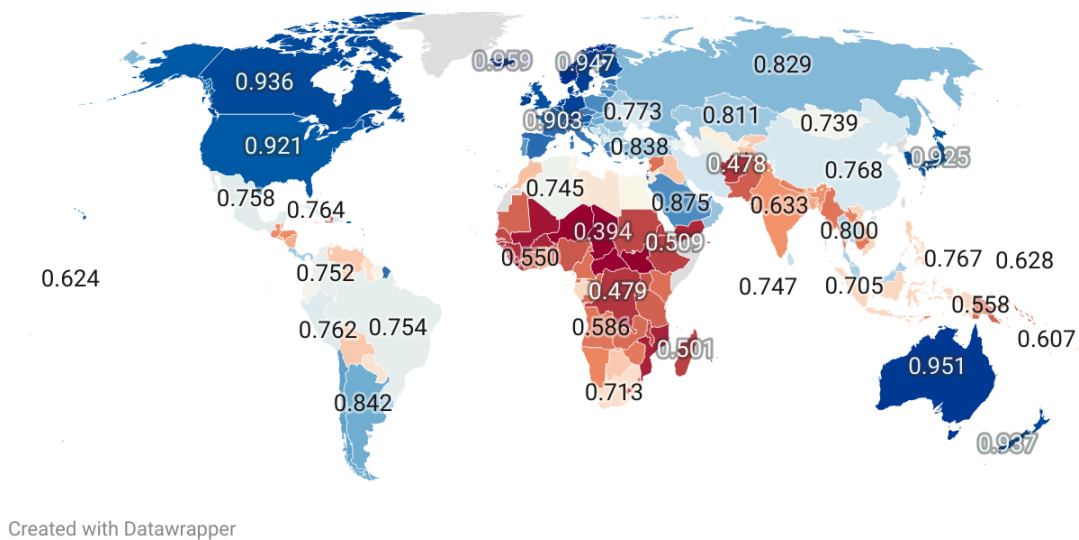
$$e_x = \frac{T_x}{l_x}$$

Таблиці смертності є необхідним інструментом для страхових компаній у визначенні ризиків та формуванні цін на страхові поліси, оскільки ці таблиці надають детальну статистику щодо тривалості життя в залежності від різних факторів, таких як вік, стать, стан здоров'я, спосіб життя та інші аспекти. Аналізуючи ці дані та утилізуючи їх для внутрішніх розрахунків, страхові компанії можуть прогнозувати, скільки часу вони можуть мати перед сплатою страхових виплат. Розуміння того, які фактори впливають на описані вище значення у таблицях смертності та загалом на тривалість життя осіб різних соціальних груп, допомагає точніше оцінювати ризики та встановлювати адекватні ціни на страхові продукти [6].

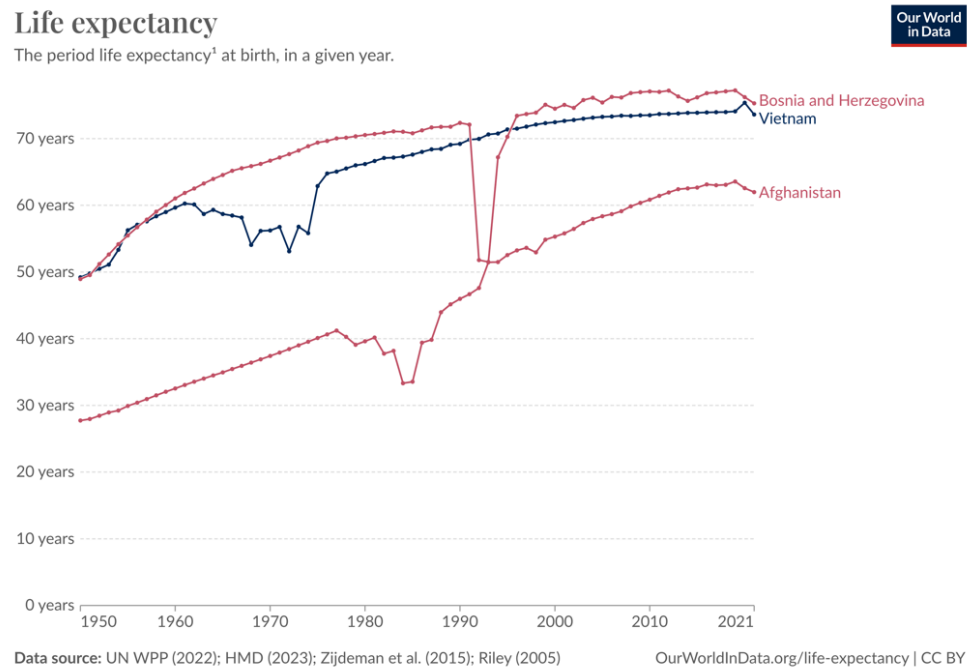
1.3 Залежність тривалості життя від сторонніх факторів

Тривалість життя індивіда залежить від різноманітних факторів, що виходять за межі здорового способу життя. Загальна кількість впливу різноманітних стресових факторів в життєвому та професійному середовищі, які накопичуються протягом життя індивіда, впливає як на якість, так і на тривалість його життя. Окрім того, надмірне споживання алкоголю, наркотиків та тютюну — лише деякі з відомих характеристик життєвого стилю, які мають різко негативні наслідки для організму та значно підвищують ризик розвитку різноманітних серйозних, іноді хронічних, хвороб [7]. Велика увага в даному контексті приділяється також харчуванню, яке може бути важливим фактором розвитку хвороб у випадках надмірного споживання шкідливих штучних добавок у харчовій продукції. Варто зазначити що в деяких дослідженнях з країн з низьким рівнем доходів дієта може вважатися здоровою, якщо вона містить достатню кількість тваринного білка, тоді як у дослідженнях з країн з високим рівнем доходів надмірне споживання білка з м'яса, риби, сиру та яєць часто вважається фактором ризику. Ця невідповідність є важливим індикатором того, наскільки нерівномірним є соціально-економічний стан у різних географічних регіонах, та про те, наскільки великим чином такі розбіжності можуть

Рис. 1.3.1. Очікувана тривалість життя для країн світу станом на 2024 рік [10]



Нижче наведено графік середньої тривалості життя трьох країн – Боснії і Герцеговини, В'єтнаму та Афганістану.



1. **Period life expectancy:** Period life expectancy is a metric that summarizes death rates across all age groups in one particular year. For a given year, it represents the average lifespan for a hypothetical group of people, if they experienced the same age-specific death rates throughout their whole lives as the age-specific death rates seen in that particular year. Learn more in our article: "Life expectancy" – What does this actually mean?

Рис. 1.3.3. Середня тривалість життя населення Боснії і Герцеговини, В'єтнаму та Афганістану з 1950 по 2021 рік [13]

Графік тривалості життя для Боснії і Герцеговини показує стабільність та помірне зростання з часом до певного моменту. Різке зниження тривалості життя у країні на початку 90-х років припадає на період Боснійської війни. У В'єтнамі спостерігається затяжний період зниження середньої тривалості життя, особливо під час війни, коли країна переживала інтенсивний конфлікт. Після закінчення воєнних дій простежується збільшення тривалості життя серед населення країни. В Афганістані графік показує стабільно низький рівень тривалості життя протягом тривалого періоду часу, але простежується також поступова позитивна тенденція. Така ситуація може бути наслідком багатьох факторів, включаючи війну, внутрішні конфлікти, бідність та низький рівень доступу до медичних послуг, і загалом

ситуація свідчить про серйозні проблеми у суспільстві та потребу в серйозних заходах для поліпшення ситуації зі здоров'ям та соціальним розвитком.

Задля детальнішого дослідження залежності тривалості життя для різних національних груп та додаткової аргументації зазначених вище припущень проаналізуємо актуальний датасет «Глобальний набір даних про країни» 2023 року (Global Country Information Dataset 2023) [14]. Для кожної країни в датасеті додамо колонку з позначенням рівня розвитку країни. Отриманий датасет містить наступні колонки:

- Country – назва країни;
- GDP – валовий внутрішній продукт країни станом на 2023 рік;
- Gross primary education enrollment - загальна кількість учнів, залучених у початковій освіті (незалежно від віку), виражена у відсотках від населення вікової групи, яка офіційно відповідає початковій школі, за рік для країни;
 - Life expectancy - очікувана тривалість життя для населення країни, базується на оцінці середнього віку громадян країни в момент їх смерті;
 - Gross tertiary education enrolment – загальна кількість учнів будь-якої вікової категорії, залучених у вищій освіті (включає державні та приватні університети, коледжі, технікуми та професійно-технічні навчальні заклади), виражена у відсотках від загальної кількості населення;
 - Population – чисельність населення країни станом на 2023 рік;
 - Unemployment rate – співвідношення непрацевлаштованих осіб до загальної кількості населення країни;
 - Urban population – чисельність населення, яке проживає на територіях, що мають більшу щільність населення, ніж сільська місцевість, і загалом є більш компактними, ніж сільська місцевість;
 - Physicians per thousand - кількість лікарів, включаючи лікарів загальної практики та лікарів-спеціалістів, на тисячу населення;

- Status – рівень розвитку країни. Країни класифіковані як розвинені або ті, що розвиваються, на основі валового внутрішнього продукту або валового національного доходу на душу населення, загального рівня життя, рівня індустріалізації і розвиненості інфраструктури та інших факторів.

На основі наявних даних додані наступні колонки та обраховані значення для кожної країни:

- Urban population percentage – співвідношення чисельності населення, яке проживає на територіях, що мають більшу щільність населення, ніж сільська місцевість, і загалом є більш компактними, ніж сільська місцевість, до загальної чисельності населення країни;

- GDP per capita - валовий внутрішній продукт країни станом на 2023 рік на душу населення;

- Out of pocket health expenditure - витрати, які несе безпосередньо пацієнт, якщо страховка не покриває повної вартості медичного товару чи послуги. Такі витрати враховують самолікування та інші витрати, які оплачуються безпосередньо приватними домогосподарствами.

Опрацюємо дані задля подальшої зручності їх використання. Подальший аналіз виконано за допомогою мови програмування Python та бібліотек Pandas [15], Matplotlib [16], SciPy [17], Plotly [18], NumPy [19], Math [20] , Statsmodels [21] та Scikit-learn [22].

Для початку візуалізуємо наявні дані для більшої зручності аналізу.

На графіку нижче зображено розподіл середньостатистичної тривалості життя для країн різних рівнів розвитку (розвинених і тих, що розвиваються).

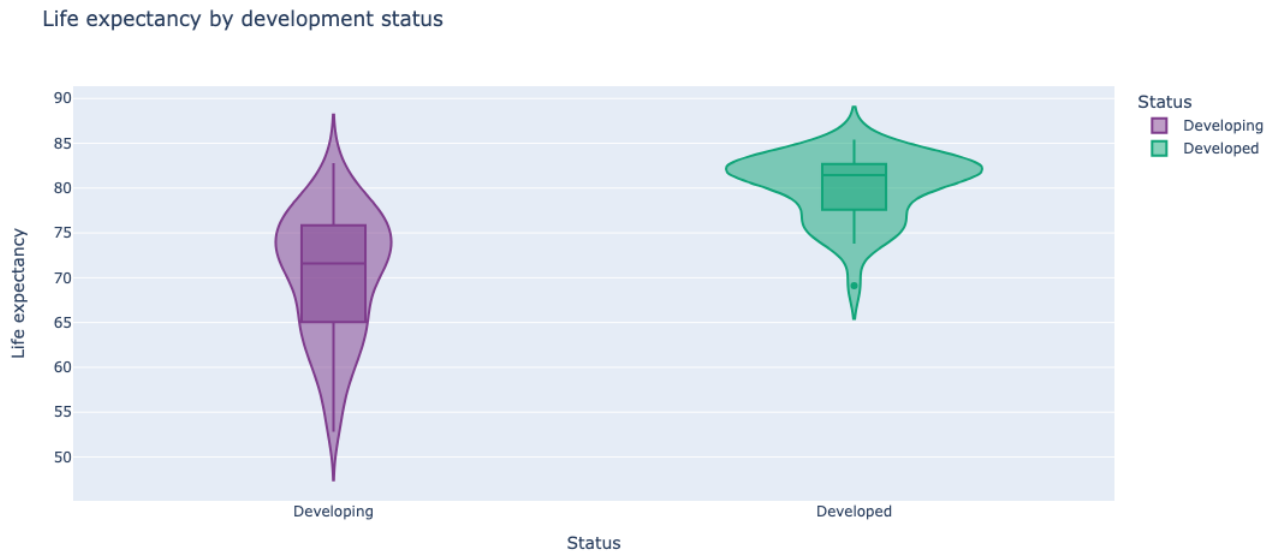


Рис. 1.3.4. Розподільчий графік значень очікуваної тривалості життя для розвинених країн та країн, що розвиваються

Помітна значна різниця між медіанами тривалості життя для розвинених країн (конкретно – 81,45) та країн що розвиваються (71,6). Також наявна велика різниця між мінімальними значеннями – 52,8 для країн що розвиваються та 69,1 для розвинених країн, в той час як настільки критичної різниці між максимальними значеннями немає (82,8 та 86,14 відповідно).

Для дослідження конкретних факторів впливу побудовано графіки кореляції тривалості життя зі значеннями інших колонок. Розглянемо деякі з побудованих графіків.

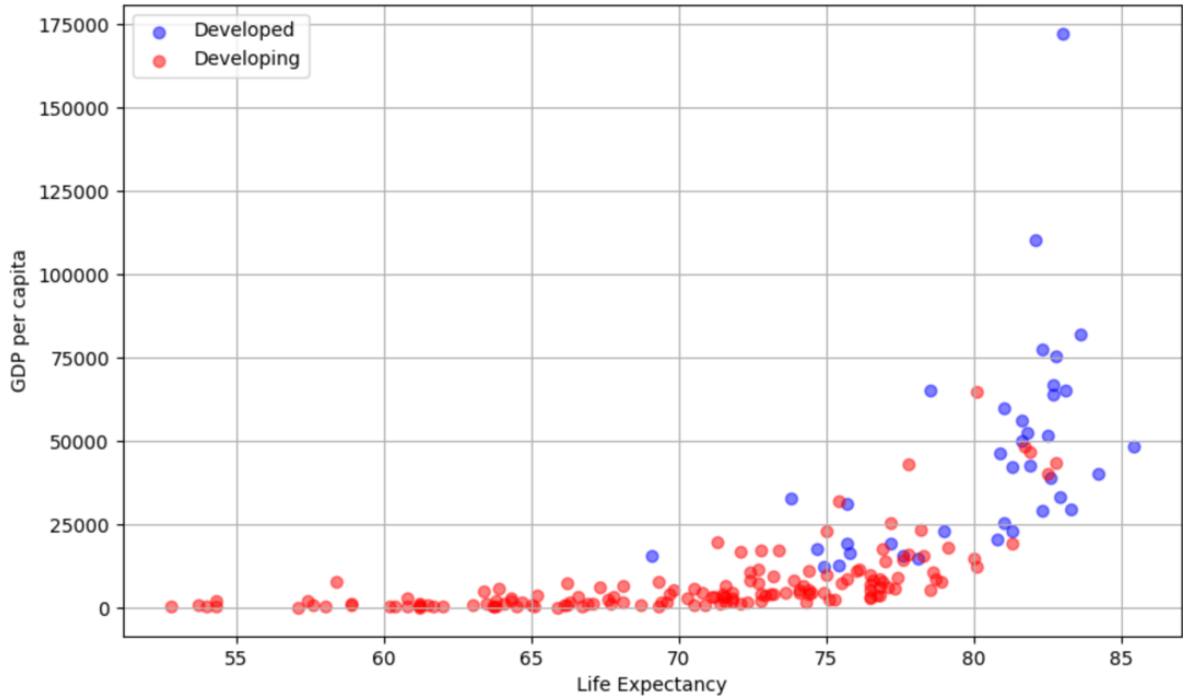


Рис. 1.3.5. Діаграма розсіювання значень очікуваної тривалості життя та ВВП на душу населення

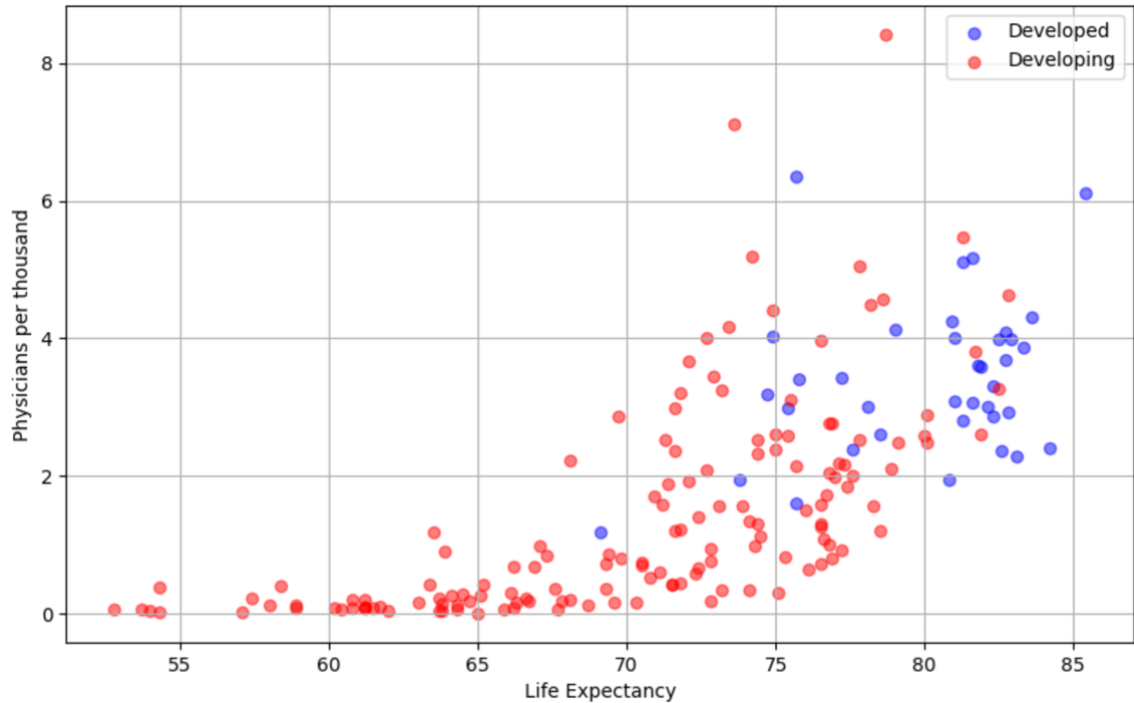


Рис.1.3.6 Діаграма розсіювання значень очікуваної тривалості життя та кількості лікарів на тисячу населення

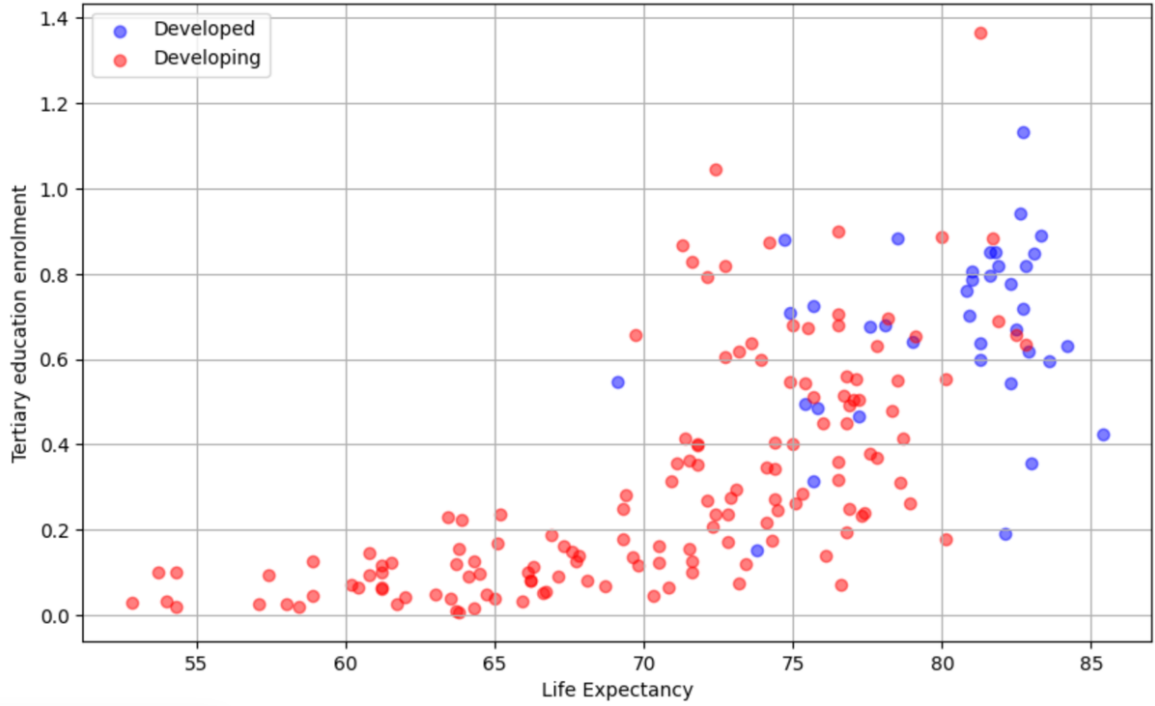


Рис.1.3.7. Діаграма розсіювання значень очікуваної тривалості життя та відсотків людей, що отримують вищу освіту

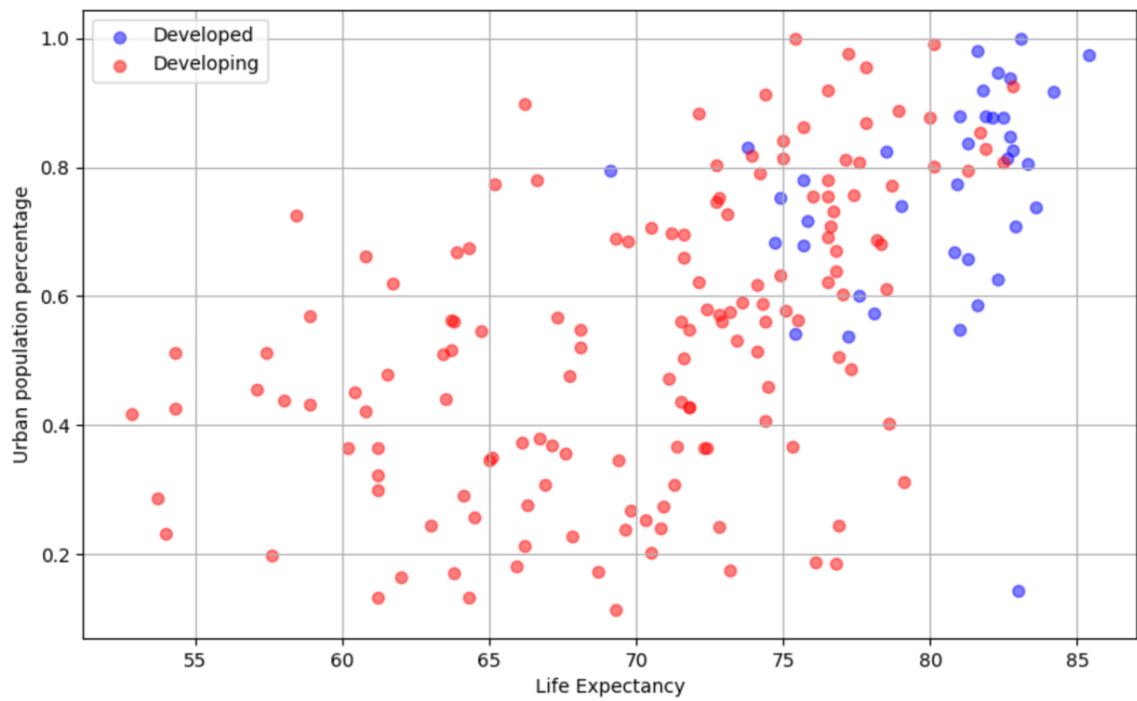


Рис.1.3.8. Діаграма розсіювання значень очікуваної тривалості життя та відсотків населення, що живе у містах

Оскільки взаємозв'язок змінних на графіках є монотонним, для знаходження коефіцієнтів кореляції для зображених на графіках вище факторів можна застосувати коефіцієнт рангової кореляції Спірмена - непараметричну міру рангової кореляції (статистичної залежності між рангами двох змінних), що оцінює наскільки добре зв'язок між двома змінними можна описати за допомогою монотонної функції. Кореляція Спірмена між двома змінними дорівнює кореляції Пірсона між ранговими значеннями цих двох змінних, але в той час як кореляція Пірсона оцінює лінійні зв'язки, кореляція Спірмена оцінює монотонні зв'язки (лінійні або нелінійні). Коефіцієнт кореляції Спірмена представлений наступним чином:

$$\rho = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}},$$

де n – розмір вибірки, \bar{x} та \bar{y} – середні значення для рядів досліджуваних змінних, позначених X та Y відповідно, x_i та y_i – ранги елементів множин X та Y , тобто середні числа позицій елементів, відсортованих в порядку зростання [23].

Для розрахунку коефіцієнтів кореляції Спірмена використано функцію `scipy.stats.spearmanr(a, b)`.

```
[ ] import scipy

[ ] life_expectancy = df["Life expectancy"]
   gdp_per_capita = df["GDP per capita"]
   scipy.stats.spearmanr(life_expectancy, gdp_per_capita)
```

Рис.1.3.9. Приклад використання бібліотеки SciPy для обчислення коефіцієнту кореляції Спірмена між очікуваною тривалістю життя та ВВП на душу населення

В результаті виконання функції для чотирьох показників, зображених на графіках 1.3.5, 1.3.6, 1.3.7 та 1.3.8, отримані наступні коефіцієнти кореляції:

- Тривалість життя та ВВП на душу населення – 0,86;

- Тривалість життя та кількість лікарів на тисячу населення – 0,81;
- Тривалість життя та відсоток громадян що отримують вищу освіту – 0,78;
- Тривалість життя та відсоток осіб що проживають у містах – 0,61.

```

▶ print(scipy.stats.spearmanr(life_exp, gdp_per_capita).statistic)
⇒ 0.85699092834884

[16] print(scipy.stats.spearmanr(life_exp, doctors).statistic)
    0.8074706842910471

[14] print(scipy.stats.spearmanr(life_exp, education).statistic)
    0.7796352618190041

[12] print(scipy.stats.spearmanr(life_exp, urban).statistic)
    0.6112289477246587

```

Рис.1.3.10. Результати обчислення коефіцієнтів кореляції Спірмена для графіків 1.3.5, 1.3.6, 1.3.7 та 1.3.8

Для перших трьох пунктів значення кореляції можна класифікувати як сильний зв'язок, для останнього – як зв'язок середньої сили.

Для подальшого аналізу даних застосовано регресійний аналіз з використанням методу найменших квадратів (OLS, англійською: Ordinary Least Squares) - статистичний метод, який використовується для оцінки залежності між залежною змінною та однією або кількома незалежними змінними. Основна мета OLS регресії — знайти найкращі коефіцієнти, які мінімізують суму квадратів відхилень спостережуваних значень залежної змінної від передбачених значень [24]. Для незалежних змінних, де спостерігається наявність нелінійної кореляції, було застосовано логарифмічне перетворення задля покращення показників коректності моделі. Нижче на рисунку 1.3.11 наведені результати OLS регресії.

| OLS Regression Results | | | | | | |
|--|------------------|---------------------|----------|-------|-----------|----------|
| Dep. Variable: | Life expectancy | R-squared: | 0.773 | | | |
| Model: | OLS | Adj. R-squared: | 0.758 | | | |
| Method: | Least Squares | F-statistic: | 51.88 | | | |
| Date: | Tue, 21 May 2024 | Prob (F-statistic): | 1.07e-35 | | | |
| Time: | 16:42:13 | Log-Likelihood: | -353.89 | | | |
| No. Observations: | 131 | AIC: | 725.8 | | | |
| Df Residuals: | 122 | BIC: | 751.7 | | | |
| Df Model: | 8 | | | | | |
| Covariance Type: | nonrobust | | | | | |
| | coef | std err | t | P> t | [0.025 | 0.975] |
| const | 39.9234 | 5.363 | 7.444 | 0.000 | 29.306 | 50.541 |
| Gross primary education enrollment | 5.9610 | 2.865 | 2.080 | 0.040 | 0.289 | 11.634 |
| log_Gross tertiary education enrolment | 10.4531 | 3.180 | 3.288 | 0.001 | 4.159 | 16.747 |
| Unemployment rate | -4.2312 | 7.476 | -0.566 | 0.572 | -19.032 | 10.569 |
| Urban_population | 1.394e-10 | 3.74e-09 | 0.037 | 0.970 | -7.26e-09 | 7.54e-09 |
| log_Physicians per thousand | 3.1303 | 1.219 | 2.567 | 0.011 | 0.717 | 5.544 |
| Urban_population_percentage | -2.4213 | 2.424 | -0.999 | 0.320 | -7.219 | 2.377 |
| log_GDP_per_capita | 2.5796 | 0.527 | 4.895 | 0.000 | 1.536 | 3.623 |
| Out of pocket health expenditure | -1.0730 | 2.115 | -0.507 | 0.613 | -5.260 | 3.114 |
| Omnibus: | 5.512 | Durbin-Watson: | 2.125 | | | |
| Prob(Omnibus): | 0.064 | Jarque-Bera (JB): | 5.028 | | | |
| Skew: | -0.458 | Prob(JB): | 0.0810 | | | |
| Kurtosis: | 3.284 | Cond. No. | 2.20e+09 | | | |

Рис.1.3.11. Результати регресійного аналізу для моделі залежності очікуваної тривалості життя від незалежних змінних у датасеті

В даній моделі значення R-квадрат (0,773) вказує на те, що приблизно 77,3% варіації в тривалості життя пояснюється моделлю. Висока F-статистика і дуже низьке р-значення вказують на те, що отримана модель являється статистично значущою. Отримано наступні коефіцієнти для незалежних змінних моделі:

- Gross primary education enrollment: коефіцієнт – 5,9610, р-значення – 0,040. Ця змінна позитивно впливає на очікувану тривалість життя і є статистично значущою.
- Gross tertiary education enrolment: коефіцієнт – 10,4531, р-значення – 0,001. Ця змінна має позитивний і дуже значущий вплив на очікувану тривалість життя.
- Unemployment rate: коефіцієнт - -4,2312, р-значення – 0,572. Ця змінна має негативний вплив на очікувану тривалість життя, але він не є статистично значущим.

- Urban population: коефіцієнт – $1,394e-10$, р-значення – 0,970. Ця змінна має дуже невеликий позитивний вплив на очікувану тривалість життя і не є статистично значущою.
- Physicians per thousand: коефіцієнт – 3,1303, р-значення – 0,011. Ця змінна має позитивний і статистично значущий вплив на очікувану тривалість життя.
- Urban population percentage: коефіцієнт - -2,4213, р-значення – 0,320. Ця змінна має негативний вплив на очікувану тривалість життя, але не є статистично значущою.
- GDP per capita: коефіцієнт – 2,5796, р-значення – 0,000. Ця змінна має позитивний і дуже значущий вплив на очікувану тривалість життя.
- Out of pocket health expenditure: коефіцієнт - -1,0730, р-значення – 0,613. Ця змінна має негативний вплив на очікувану тривалість життя, але не є статистично значущою.

Для отримання більш точних результатів та кращого врахування потенційних нелінійних залежностей між залежними та незалежними змінними можна застосувати модель Random forest (англ. випадковий ліс) - популярний метод машинного навчання, який використовується для вирішення задач регресії та класифікації. Випадковий ліс складається з великої кількості дерев рішень, кожне з яких тренується на випадковій підмножині даних та ознак. При побудові кожного дерева використовується випадковий вибір ознак, що допомагає зменшити кореляцію між деревами та покращити загальну продуктивність моделі. Використання моделі випадкового лісу дозволяє ідентифікувати ознаки які є найбільш важливими для прогнозування очікуваної тривалості життя [25].

```

from sklearn.ensemble import RandomForestRegressor
from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_squared_error
import numpy as np

X = df[['Gross primary education enrollment',
        'Gross tertiary education enrolment', 'Unemployment rate',
        'Urban_population', 'Physicians per thousand',
        'Urban_population_percentage', 'GDP per capita',
        'Out of pocket health expenditure']]
y = df['Life expectancy']
X_train, X_test, y_train, y_test =
train_test_split(X, y, test_size=0.2, random_state=42)

rf = RandomForestRegressor(n_estimators=100, random_state=42)

rf.fit(X_train, y_train)

y_pred = rf.predict(X_test)

mse = mean_squared_error(y_test, y_pred)
rmse = np.sqrt(mse)
print(f'RMSE: {rmse}')

importances = rf.feature_importances_
feature_names = X.columns
sorted_indices = np.argsort(importances)[::-1]

for i in sorted_indices:
    print(f'{feature_names[i]}: {importances[i]}')

```

Рис.1.3.12. Програмний код з використання методу Random Forest та визначенням важливості характеристик

```

➡ RMSE: 3.191151295785365
Gross tertiary education enrolment: 0.4236178041998351
Physicians per thousand: 0.2472819383609869
GDP per capita: 0.2364596753341272
Urban_population: 0.022835555483011936
Urban_population_percentage: 0.020029793027997372
Out of pocket health expenditure: 0.019565956923332115
Gross primary education enrollment: 0.018906279704780853
Unemployment rate: 0.01130299696592843

```

Рис.1.3.13. Результати виконання коду на рисунку 1.3.12

Підсумуємо результати виконання програми 1.3.12:

- Ефективність моделі пояснюється значенням середньоквадратичної похибки (RMSE): 3,191. Це значення вказує на середню похибку прогнозів

моделі щодо очікуваної тривалості життя. Отримане значення RMSE свідчить про досить добру якість прогнозування очікуваної тривалості життя.

- Модель показує важливість кожної ознаки для прогнозування очікуваної тривалості життя. Відсоток охоплення населення вищою освітою – найвагоміший (42,36%) предиктор у моделі, що свідчить про те, що вищий рівень охоплення вищою освітою тісно пов'язаний з вищою очікуваною тривалістю життя. Кількість лікарів на тисячу осіб є другим за важливістю фактором (24,73%), що підкреслює вирішальну роль доступності охорони здоров'я у визначенні тривалості життя. Вищий ВВП на душу населення також є основним предиктором (23,65%). Загальна кількість міського населення робить свій внесок у прогноз, але цей внесок є менш значущим порівняно з першими трьома факторами (2,28%). Відсоток населення, що проживає в містах, має помірний вплив на очікувану тривалість життя (2,00%). Сума, витрачена “з власної кишені” на охорону здоров'я, є ще одним фактором, хоча його вплив є відносно незначним порівняно з провідними предикторами (1,96%). Зарахування до початкової освіти також відіграє певну роль, але є менш важливим порівняно з вищою освітою (1,89%). Рівень безробіття має найменший вплив серед ознак (1,13%).

Отримані результати підтверджують наведені попередньо потенційні причини нижчої тривалості життя у різних країнах. Більше лікарів може означати більшу доступність профілактичних заходів та скринінгових програм, які допомагають виявляти захворювання на ранніх стадіях, коли вони легше піддаються лікуванню. Люди, які мають можливість легко отримати консультацію або лікування, швидше звертаються за допомогою при медичних проблемах, що дозволяє їм уникнути ускладнень та погіршення стану здоров'я. Варто звернути увагу і на зв'язок між вищою тривалістю життя та вищим рівнем освіти, який може бути пояснений кількома чинниками. Освічені люди зазвичай мають кращі знання про те, як

піклуватися про своє здоров'я. Окрім того, люди з вищою освітою зазвичай мають доступ до вищих оплачуваних робіт та стабільних кар'єрних можливостей. Це дозволяє їм мати фінансову спроможність для доступу до високоякісної медичної допомоги та інших послуг, що сприяють довголіттю. Зв'язок між відсотком людей, які проживають у містах, та тривалістю життя є менш виражений зв'язок порівняно з іншими факторами. З однієї сторони, міське середовище може забезпечити доступ до кращої медичної інфраструктури та послуг, так як міста зазвичай мають більше лікарень, клінік та спеціалізованих закладів охорони здоров'я. Також міста зазвичай мають кращий доступ до освіти, розваг та культурних заходів, що може позитивно впливати на психічне здоров'я та загальний стан населення. З іншої сторони, міське середовище може бути пов'язане з вищим рівнем стресу та психічних навантажень через швидкий темп життя, конкуренцію та надмірну кількість подразників. Міста часто мають вищий рівень забруднення повітря, шуму та інших форм забруднення довкілля. Експозиція до цих негативних факторів може призвести до розвитку різних захворювань, таких як захворювання дихальних шляхів, серцево-судинні захворювання та інші хронічні захворювання, що можуть скоротити тривалість життя.

В контексті сьогодення в Україні особливу увагу варто звернути на відтік медичних працівників за кордон та нестабільність економічної ситуації. Згідно з дослідженням Українського центру охорони здоров'я, кількість медичного персоналу в Україні у 2022 році зменшилася на 14% порівняно з попереднім роком. Особливо суттєве зменшення відбулося серед молодшого медперсоналу, а саме на 16,5% [26]. Це свідчить про серйозні виклики, які війна створює для системи охорони здоров'я країни. Зменшення кількості медичного персоналу, особливо на прифронтових і тимчасово окупованих територіях, може бути наслідком небезпечних умов роботи, втрати доступу до медичної допомоги або страху за власну безпеку. Це може погіршити доступність медичних послуг для мешканців

цих регіонів і загострити ситуацію з медичним обслуговуванням. За результатами аналізу наявних даних, така ситуація разом з відтоком студентів та випускників шкіл за кордон може мати значний негативний вплив на очікувану тривалість життя населення в Україні.

Дослідження впливу на тривалість життя груп осіб на рівні країн дають важливу інформацію про загальні тенденції та фактори, що впливають на тривалість життя, але не дають можливості глибше зрозуміти, як ці фактори взаємодіють на мікрорівні. Для досягнення більш комплексного розуміння необхідно розглянути дані про смертність конкретних осіб та характеристики способу життя цих осіб. Подібний аналіз на індивідуальному рівні відіграє важливу роль як у наукових дослідженнях, так і в практичному застосуванні, оскільки він дозволяє визначити конкретні фактори, що сприяють смертності, такі як вік, стать та соціально-економічні умови, та виявити і класифікувати вразливі групи населення.

Для подальшого аналізу використано дані з датасету «Annual Health Survey: Mortality Schedule» [27]. Датасет містить загалом 770 тисяч спостережень і 121 змінну, що робить його одним з найбільших у своєму роді. Збір даних було проведено у дев'яти штатах Індії, які мають відносно високі показники народжуваності та смертності - ці дев'ять штатів, які складають близько 48% загального населення Індії, є критично важливими для аналізу, оскільки вони відповідають за 59% народжень, 70% дитячих смертей, 75% смертей дітей до 5 років та 62% смертей при пологах у країні. Цей датасет містить детальну інформацію про смертність серед звичайних жителів домогосподарств протягом фіксованого періоду з 2010 по 2013 роки. Він включає інформацію про стать померлого, дату смерті, вік на момент смерті, реєстрацію смерті та джерело медичної допомоги, отриманої перед смертю. Крім того, дані включають інформацію про різні фактори материнської смертності, такі як випадки смерті, пов'язані з вагітністю, інформацію про фактори, що спричинили або сприяли смерті, симптоми, що передували смерті,

та час між початком ускладнень і смертю. Проаналізуємо значення віку смерті для різних груп індивідів, поділених за деякими важливими для дослідження категоріями, за допомогою графіків, наведених нижче.

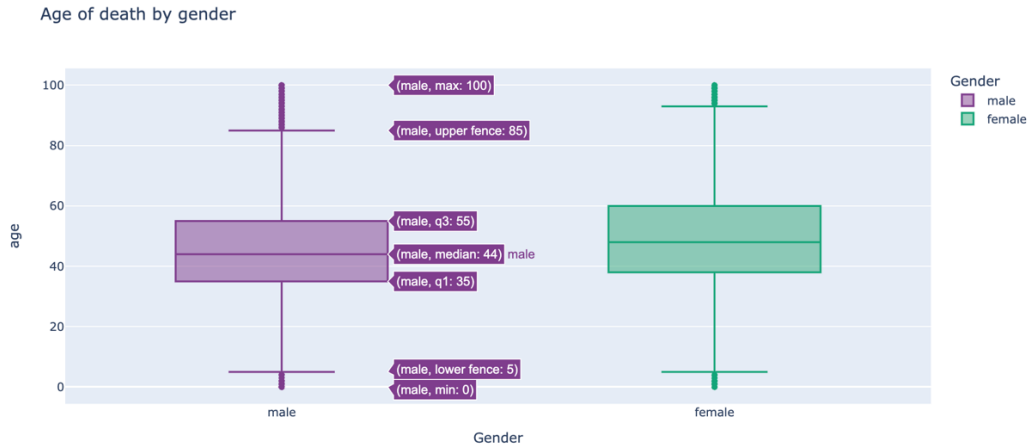


Рис.1.3.14. Графік розподілу віку смерті за статевою ознакою(зі значеннями для чоловіків)

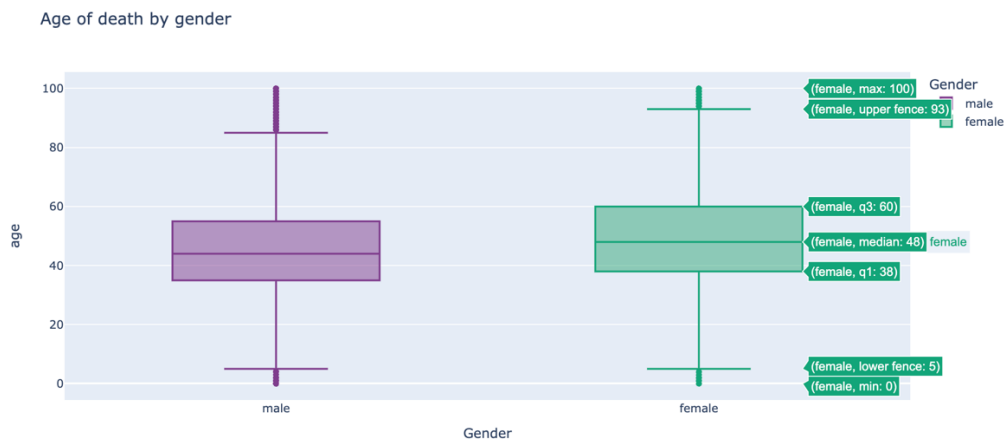


Рис.1.3.15. Графік розподілу віку смерті за статевою ознакою(зі значеннями для жінок)

На графіці 1.3.14 – 1.3.15 зображене порівняння значень тривалості життя для чоловічої та жіночої частини населення. З графіку можна зробити наступні висновки:

- Середній вік смерті (медіана): жінки даного регіону в середньому живуть довше за чоловіків (48 років проти 44 років);

- Міжквартильний розмах розподілу віку смерті у жінок більший (22 роки) порівняно з чоловіками (20 років), що свідчить про більшу варіабельність віку смерті серед жінок.

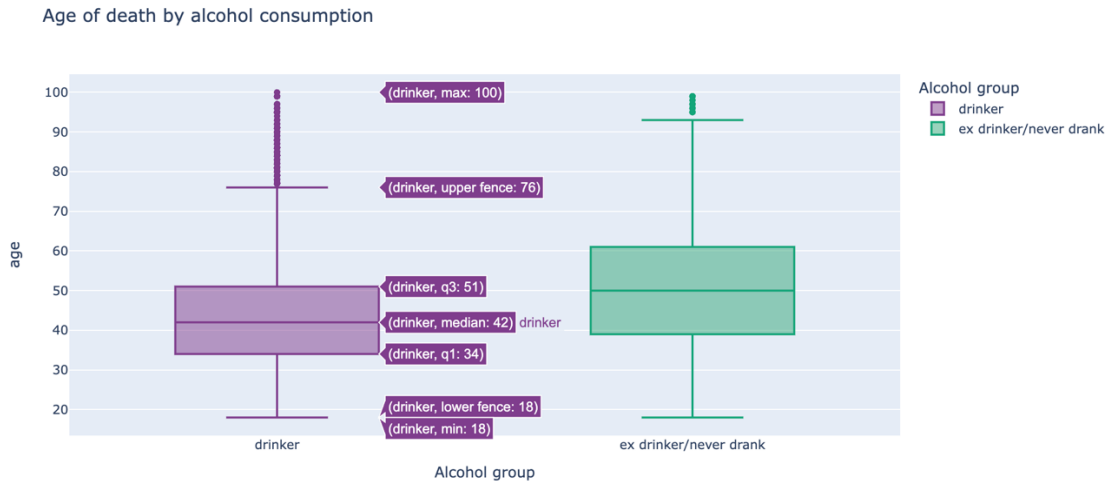


Рис.1.3.16. Графік розподілу віку смерті за споживанням алкоголю (зі значеннями для осіб що споживали алкоголь на момент смерті)

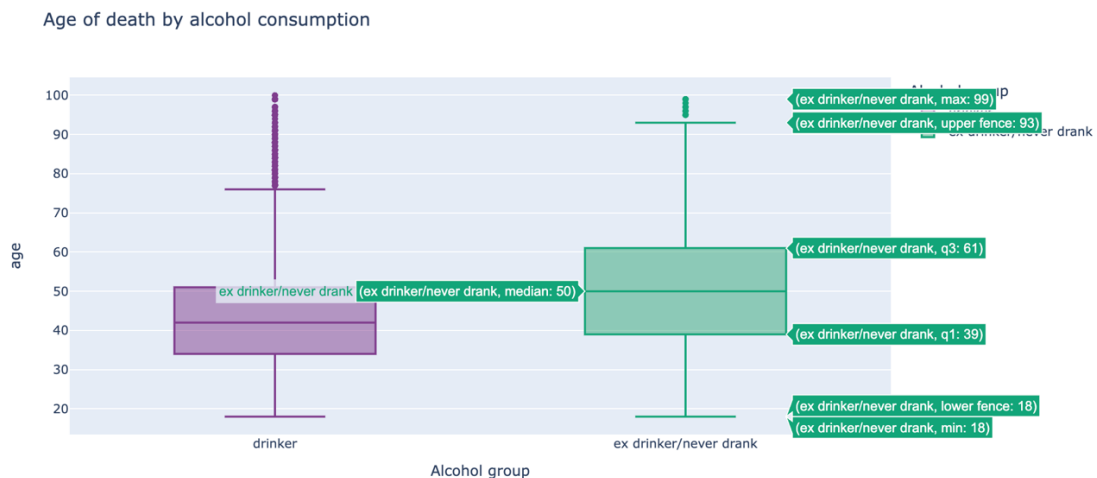


Рис.1.3.17. Графік розподілу віку смерті за споживанням алкоголю (зі значеннями для осіб що не споживали алкоголь на момент смерті)

У випадку графіку 1.3.16 – 1.3.17 осіб було класифіковано як тих, хто на момент смерті активно вживали алкогольні напої та тих, хто на момент смерті не вживав алкоголь. Також задля кращої репрезентативності графіку були враховані лише ті

записи, де померла особа була повнолітньою. З графіку можна зробити наступні висновки:

- Середній вік смерті (медіана): особи, які не вживали алкоголь або припинили вживати до моменту смерті, прожили у середньому довше (50 років) у порівнянні з тими, хто вживав активно вживав алкоголь (42 роки).

Age of death by smoking habit

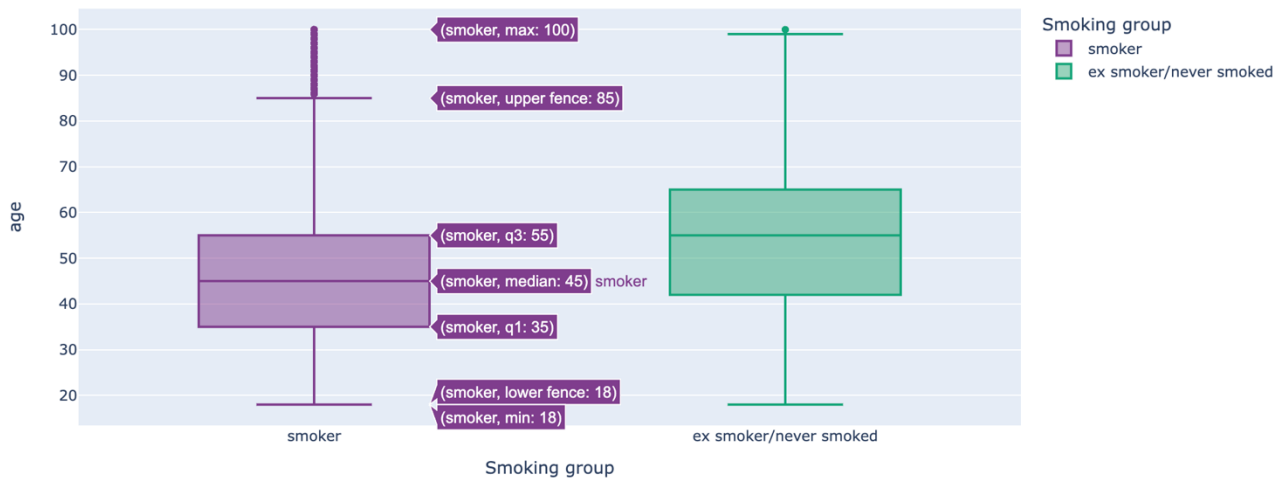


Рис.1.3.18. Графік розподілу віку смерті за споживанням тютюнової продукції (зі значеннями для осіб що палили на момент смерті)

Age of death by smoking habit

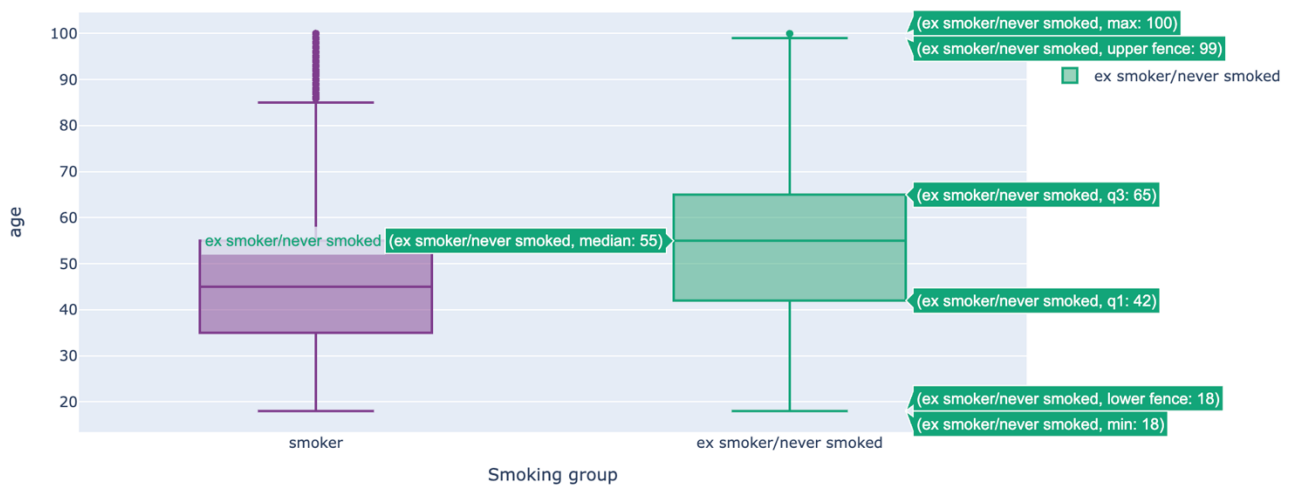


Рис.1.3.19. Графік розподілу віку смерті за споживанням тютюнової продукції (зі значеннями для осіб що не палили на момент смерті)

Аналогічний підхід застосовано для графіку 1.3.18 – 1.3.19 – осіб класифіковано як тих, хто активно вживав нікотинову продукцію на момент своєї смерті та тих, хто до моменту смерті припинив палити або взагалі ніколи не палив. Задля кращої репрезентативності графіку були враховані лише ті записи, де померла особа була повнолітньою.

З графіку можна зробити наступні висновки:

- Середній вік смерті (медіана): особи, які не вживали нікотинову продукцію або припинили вживати до моменту смерті, прожили у середньому довше (55 років) у порівнянні з тими, хто вживав активно вживав нікотинові вироби (45 років).

Для того, щоб краще дослідити важливі фактори впливу на тривалість життя осіб різних груп, скористаємось методом Random forest та проаналізуємо значущі фактори впливу на залежну змінну, в нашому випадку – вік, у якому особа померла.

```
[ ] features = ['state', 'district', 'rural', 'treatment_source',
               'nearest_medical_facility', 'sex', 'religion', 'social_group_code',
               'marital_status', 'highest_qualification', 'occupation_status',
               'disability_status', 'illness_type', 'sought_medical_care',
               'regular_treatment', 'regular_treatment_source', 'smoke', 'alcohol',
               'drinking_water_source', 'is_water_filter',
               'household_have_electricity', 'land_possessed',
               'iscoveredbyhealthscheme']

df_cleaned = df.copy()

for feature in features:
    most_frequent_value = df_cleaned[feature].mode()[0]
    df_cleaned[feature].fillna(most_frequent_value, inplace=True)

most_frequent_age = df_cleaned['age'].mode()[0]
df_cleaned['age'].fillna(most_frequent_age, inplace=True)

X = df_cleaned[features]
y = df_cleaned['age']

[ ] X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
                                                    random_state=42)

rf_model = RandomForestRegressor(n_estimators=100, random_state=42)
rf_model.fit(X_train, y_train)
y_pred_rf = rf_model.predict(X_test)

[ ] importances = rf_model.feature_importances_
feature_names = X.columns
sorted_indices = np.argsort(importances)[::-1]

for i in sorted_indices:
    print(f'{feature_names[i]}: {importances[i]}')
```

Рис.1.3.20. Приклад підготовки даних, використання алгоритму Random Forest та визначенням важливості характеристик

```

↔ occupation_status: 0.24749472526805863
  district: 0.1548389845883548
  treatment_source: 0.08774080003435153
  marital_status: 0.08166761081592272
  highest_qualification: 0.05279556013745943
  drinking_water_source: 0.05067425837336753
  state: 0.04591747631381313
  land_possessed: 0.043623253903364426
  illness_type: 0.024262075701996528
  smoke: 0.02307966776061368
  alcohol: 0.021694400295305304
  social_group_code: 0.021690147506375865
  regular_treatment_source: 0.018694315019379522
  religion: 0.017990074205777168
  rural: 0.01520355318926956
  iscoveredbyhealthscheme: 0.01474969030758132
  household_have_electricity: 0.014320619133024284
  is_water_filter: 0.014310442591875937
  sex: 0.013847646626379667
  sought_medical_care: 0.012143413578986744
  regular_treatment: 0.010916249037673057
  disability_status: 0.007115076220340544
  nearest_medical_facility: 0.005229959390728562

```

Рис.1.3.21. Результати виконання коду з рисунку 1.3.20

Після застосування методу до даних про смертність були виявлені наступні найбільш значущі фактори, що впливають на вік смерті:

- Вид/статус професійної зайнятості (`occupation_status`): статус зайнятості виявився найважливішим фактором, що впливає на вік смерті. Це може бути пов'язано з різницею в умовах праці, рівнем стресу та рівнем оплати праці для осіб різних видів зайнятості.
- Район проживання (`district`): даний географічний фактор також значно впливає на тривалість життя, що може бути обумовлено рівнем розвитку інфраструктури, екологічними умовами та доступністю медичних послуг.
- Джерело лікування (`treatment_source`): джерело отримання медичної допомоги є важливим фактором, оскільки якість та своєчасність медичних послуг безпосередньо впливають на перебіг хвороби, а відповідно і на тривалість життя пацієнта.

- Сімейний стан (`marital_status`): сімейний стан також може впливати на тривалість життя, оскільки люди, що перебувають у шлюбі, зазвичай мають кращу соціальну підтримку та стабільне місце проживання
- Рівень отриманої освіти (`highest_qualification`): освіта може мати позитивний вплив на обізнаність щодо здорового способу життя та доступ до стабільного джерела доходу
- Джерело питної води (`drinking_water_source`): якість питної води безпосередньо впливає на здоров'я індивіда
- Штат (`state`): регіональні особливості також мають значний вплив на тривалість життя через різні рівні економічного розвитку регіонів
- Землеволодіння (`land_possested`): особи з більшою площею земельної ділянки у багатьох випадках є більш заможними, отже, мають менше перешкод до доступу до якісних та своєчасних медичних послуг
- Тип захворювання (`illness_type`): різні захворювання мають різних вплив на якість життя особи та його тривалість
- Паління (`smoke`) та вживання алкоголю (`alcohol`): негативні звички знижують тривалість життя через підвищений ризик розвитку різних захворювань та злоякісних пухлин.

1.4 Метод Каплана-Мейєра для дослідження смертності населення

Однією з переваг даного датасету є можливість використання метода Каплана-Мейєра (англійською: Kaplan–Meier estimator).

Метод Каплана-Мейєра є одним з поширених методів для оцінки функції виживання. Він дозволяє оцінювати ймовірність виживання індивіда певного віку в певний фіксований момент часу. Цей метод широко використовується в медичних дослідженнях, дослідженнях соціальних наук та інших областях, де важливо оцінити тривалість часу до настання певної події, в даному випадку – смерті індивіда.

Оцінка Каплана-Мейєра має загальноприйняте позначення $S(t)$, де t позначає певний момент часу, та розраховується за наступною формулою:

$$S(t) = \prod_{t_i \leq t} \left(1 - \frac{d_i}{n_i}\right),$$

де,

- t_i - моменти часу, коли відбувається подія
- d_i - кількість подій (в контексті даного дослідження – смертей) у момент часу t_i

- n_i - кількість осіб що перебувають під ризиком у момент часу t_i , в контексті даного дослідження – кількість осіб, що є живими у момент часу t_i . На самому початку дослідження, тобто при $t = 0$, всі індивіди вважаються під ризиком

Оцінка Каплана-Мейєра зазвичай представляється графічно у вигляді кривої. На осі абсцис (x) буде позначений вік індивіда, а на осі ординат (y) — ймовірність виживання індивіда $S(t)$ [28].

На основі даних позначень розрахуємо функцію виживання Каплана-Мейєра для деяких груп індивідів.

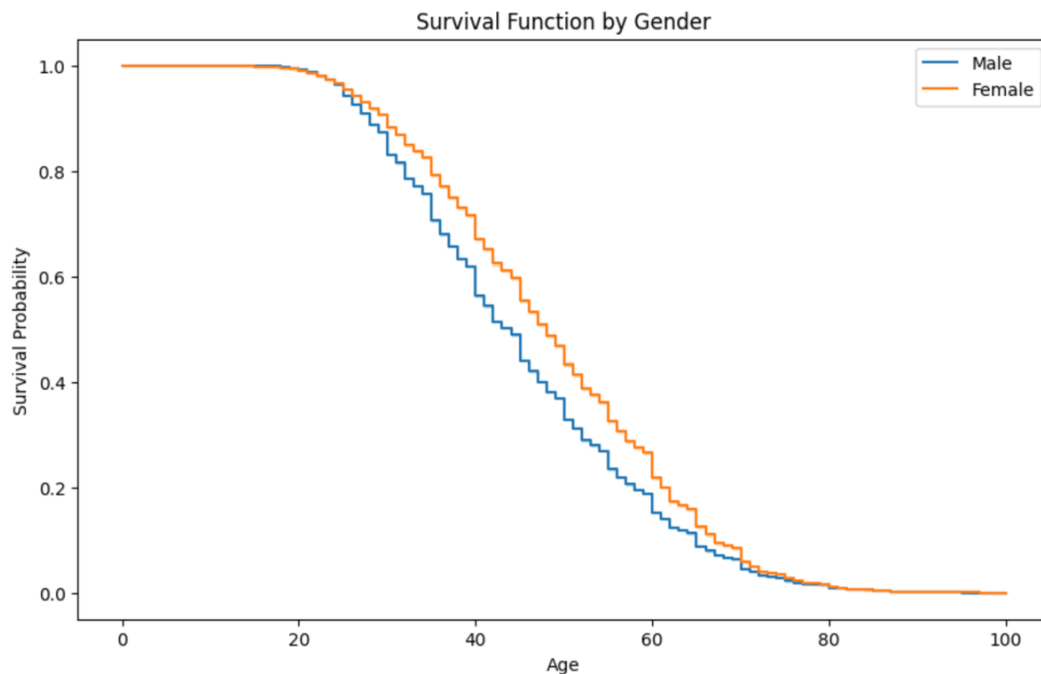


Рис.1.4.1. Функція виживання за методом Каплана-Мейєра за статевою ознакою

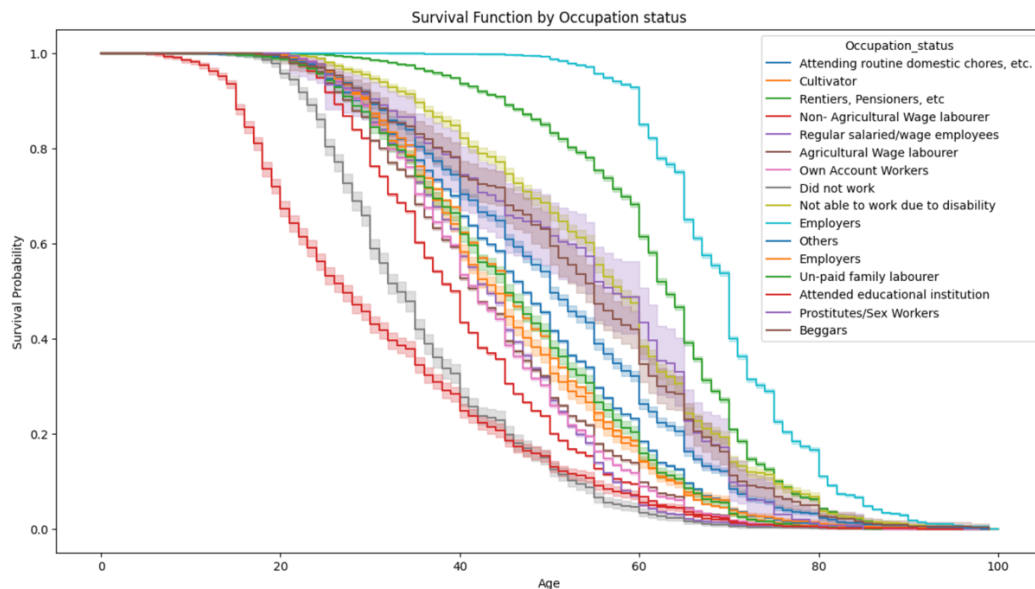


Рис.1.4.2. Функція виживання за методом Каплана-Мейєра за статусом зайнятості осіб

На графіку 1.4.2 представлено ймовірність виживання залежно від віку для людей з різним статусом зайнятості. Кожна лінія представляє окрему категорію зайнятості, зазначену в легенді. Люди, які є роботодавцями, мають найбільші шанси дожити до старшого віку, що обумовлено кращим фінансовим положенням та доступом до більш якісних медичних послуг. Найбільш вразливі групи населення – це непрацевлаштоване населення, що ймовірно спричинено нестачею фінансів для підтримки задовільної якості життя.

Розділ 2. Страхування життя – види, специфіка та розрахунки

2.1 Загальна інформація про страхування життя та основні терміни

В сучасному світі страхування життя є одним із інструментів забезпечення фінансової стабільності та захисту для індивідів та їхніх сімей. Від свого виникнення до сучасності методи страхування життя пройшли довгий шлях розвитку, відбиваючи зміни в суспільстві, технологіях та фінансових підходах. Існують різні види страхування життя, кожен з яких має свої унікальні особливості та призначення. У цьому розділі розглянуто різні види страхування життя та їхні особливості.

Види страхування життя поділяють на дві основні групи – тимчасові та довічні. Тимчасовий поліс має фіксовану тривалість дії, в той час як довічний, або ж безстроковий може діяти до кінця життя індивіда. Тимчасове страхування життя, як правило, найкраще підходить для тих, кому потрібен доступний захист та відчуття стабільності на певний період життя. Довічне страхування життя користується попитом серед осіб, що прагнуть мати стабільний та передбачуваний поліс страхування життя, який забезпечить можливість залишити після себе значну суму грошей і надасть доступ до фінансів протягом усього життя [29].

Розглянемо детальніше різні види страхування життя та їх специфіку. Для початку введемо основні терміни.

Страховик – страхова компанія що надає послуги.

Страхувальник – особа, що укладає договір страхування зі страховиком та за договором бере на себе відповідальність за сплату внесків.

Застрахована особа – особа, якій страховик надає захист [30].

Страхова подія – виживання або смерть застрахованої особи (залежить від виду страхування) [31].

Брутто-премія – сума, що сплачується страхувальником згідно договору страхування та покриває нетто премію та додаткові витрати, наприклад вартість підписання страхового договору [32].

Разова нетто-премія – одноразова плата за страхування, що сплачується страхувальником і в основі формування вартості якої лежить ймовірність настання страхової події. На вартість впливають такі фактори як тривалість дії страхування за договором і таблиці смертності [33].

Вигодонабувач – особа, яка за договором страхування життя у разі настання страхової події отримає страхову виплату.

2.2 Основні види страхування життя з виплатою страхової суми в кінці року смерті застрахованої особи

Вважається, що час і розмір виплати страхової суми (разовий платіж) можуть бути функціями випадкової величини T – тривалості майбутнього життя індивіда, введеної в розділі 1, отже, час і розмір виплати можуть бути випадковими величинами.

Через Z надалі буде позначатись поточна вартість страхової виплати. Математичне сподівання E величини Z позначає разову нетто-премію поточного договору страхування життя.

Надалі при розгляді видів страхування життя з фіксованою сумою страхової виплати будемо вважати, що передбачається виплата суми в розмірі 1 при настанні страхової події.

Розглянемо довічне страхування життя – вид страхування, що забезпечує покриття протягом всієї тривалості життя застрахованої особи. Такі безстрокові поліси поєднують у собі виплату в разі смерті з накопичувальним компонентом. В цьому випадку розмір виплати страхової суми є не випадковою, а фіксованою величиною, в той час як момент виплати є випадковою величиною і позначається як

$K + 1$, де K - кількість повних років життя індивіда до настання моменту смерті.
Поточна вартість виплати виражається через формулу:

$$Z = v^{K+1},$$

де v – коефіцієнт дисконтування, що виражається як

$$v = \frac{1}{1+i},$$

i – фактична річна відсоткова ставка, визначена договором страхування життя.

Розподіл величини Z отримується відповідно розподілу величини K , оскільки значення величини витікає з ймовірності прожиття індивідом певної кількості років:

$P(Z = v^{k+1}) = P(K = k) = {}_k p_x \cdot q_{x+k}$ (${}_k p_x$ – ймовірність того що особа після віку x прожила k років, q_{x+k} – ймовірність що особа померла протягом року після настання віку $x+k$).

Разова нетто-премія для довічного страхування життя позначається A_x і розраховується наступним чином:

$$A_x := E(v^{K+1}) = \sum_{k=0}^{\infty} v^{k+1} \cdot {}_k p_x \cdot q_{x+k}$$

Рекурентно формула виражається як

$$A_x = vq_x + vA_{x+1}p_x$$

або

$$A_x = vA_{x+1} + v(1 - A_{x+1})q_x$$

Дисперсія величини Z може бути необхідною для подальшого врахування ризику який несе страхувальник і розраховується за формулою

$$D(Z) = E(Z^2) - A_x^2, E(Z^2) = E(e^{-2\delta(K+1)})$$

Значення $E(Z^2)$ у формулі отримано при заміні v на $e^{-\delta}$, де δ – норма відсоткового прибутку, що виражається як

$$\delta = \ln(1 + i)$$

та має наступний зв'язок з v :

$$v^h = e^{-\delta h}$$

Тепер розглянемо тимчасове страхування життя – вид страхування, призначений для забезпечення фінансового захисту на випадок смерті застрахованої особи протягом певного періоду, найчастіше кількості років (надалі позначатиметься як n), відомого як термін дії полісу. В такому випадку страховою подією буде вважатись смерть застрахованого індивіда протягом n років дії страхування, і прийнята сума страхової виплати 1 буде виплачена в разі настання цієї страхової події протягом n років, і фактичний час виплати суми – в кінці року смерті індивіда. В такому випадку поточна вартість виплати Z виражається наступним чином:

$$Z = \begin{cases} v^{K+1}, & K < n \\ 0, & K \geq n \end{cases}$$

Для тимчасового страхування життя маємо наступне позначення і формулу розрахунку разової нетто-премії:

$$A_{x:\overline{n}|}^1 = \sum_{k=0}^{n-1} v^{k+1} \cdot {}_k p_x \cdot q_{x+k}$$

В такому випадку $A_{x:\overline{\infty}|}^1 = A_x$.

Ще одним видом страхування життя є чисте доживання - це вид тимчасового страхування життя, за якого страхова сума виплачується лише у разі, якщо страхувальник залишається живим до кінця зазначеного терміну n років. Якщо страхувальник помирає протягом цього періоду, страхова сума не виплачується. Основною причиною купівлі полісу чистого доживання є зменшення ризиків для осіб, що мають заключений договір тимчасового страхування життя та сумніваються у ймовірності отримання виплати у разі ненастання смерті протягом n років. Поточна вартість виплати у випадку чистого доживання виражається через формулу:

$$Z = \begin{cases} 0, & K < n \\ v^{K+1}, & K \geq n \end{cases}$$

Разова нетто-премія позначається як $A_{x:\overline{n}|}^1$ і розраховується за формулою:

$$A_{x:\overline{n}|}^1 = v^n \cdot {}_n p_x$$

Дисперсія величини Z виглядає як

$$D(Z) = v^{2n} \cdot {}_n p_x \cdot {}_n q_x$$

Окрім чистого доживання існує такий вид страхування життя як доживання - якщо смерть індивіда настає протягом зазначеного терміну (перші n років), страхова сума виплачується в кінці року, в якому настала смерть, в протилежному випадку, тобто якщо застрахована особа доживає до зазначеного терміну (через n років), страхова сума виплачується в кінці n -го року. Поточна вартість виплати у випадку доживання виражається як

$$Z = \begin{cases} v^{K+1}, & K < n \\ v^n, & K \geq n \end{cases}$$

Разова нетто-премія для доживання позначається як $A_{x:\overline{n}|}$. Оскільки за своєю сутністю доживання є комбінацією тимчасового страхування (далі виражене як Z_1) та чистого доживання (далі виражене як Z_2), то Z також можна виразити як

$$Z = Z_1 + Z_2$$

Виходячи з цього, нетто-премія для доживання виражається як

$$A_{x:\overline{n}|} = E(Z) = A_{x:\overline{n}|}^1 + A_{x:\overline{n}|}^{\frac{1}{}}$$

Дисперсія для величини Z виражається наступним чином:

$$D(Z) = D(Z_1) + 2\text{Cov}(Z_1 Z_2) + D(Z_2)$$

$$\text{Cov}(Z_1 Z_2) = E(Z_1 Z_2) - E(Z_1)E(Z_2) = 0 - E(Z_1)E(Z_2) = -A_{x:n}^1 A_{x:n}^1$$

$$D(Z) = D(Z_1) + D(Z_2) - 2A_{x:\overline{n}|}^1 A_{x:\overline{n}|}^{\frac{1}{}}$$

Маючи наведені вище означення, можна також сформулювати означення для довічного відкладеного на m років страхування - типу страхового продукту, який покриває весь залишок життя застрахованої особи, але з відкладенням першого можливого моменту виплат до певного часу в майбутньому, визначеного строком m років. Тобто в разі смерті застрахованого індивіда протягом m років страхова сума

не буде виплачена. Поточна вартість страхової виплати для даного виду страхування життя визначається наступним чином:

$$Z = \begin{cases} 0, & K < m \\ v^{K+1}, & K \geq m \end{cases}$$

Разова нетто-премія в цьому випадку позначається символами ${}_m|A_x$ і за своєю сутністю виражається як різниця між нетто-премією для довічного страхування і тимчасового страхування з терміном m років для особи віку x

$${}_m|A_x := {}_m p_x \cdot v^m \cdot A_{x+m}$$

$${}_m|A_x = A_x - A_{x:\overline{m}|}^1$$

2.3 Основні види страхування життя з виплатою в момент смерті застрахованої особи

В описаних у розділі 2.2 елементарних видах страхування передбачалось, що страхова виплата відбувається під кінець року смерті індивіда. Така практика, хоча і дає перевагу обчислення значень лише за допомогою таблиць смертності, не дуже відповідає реальній практикованій ситуації у сфері страхування життя. Розглянемо наведені вище види страхування життя, але модифікуємо вирази задля врахування виплати безпосередньо в момент смерті застрахованої особи. Тепер для довічного страхування поточна вартість страхової суми буде виражена як

$$Z = v^T,$$

а разова нетто-премія буде представлена як \bar{A}_x .

$$\bar{A}_x = \int_0^{\infty} v^t {}_t p_x \mu_{x+t} dt$$

Якщо S – дробова частина року смерті, протягом якої застрахована особа залишається живою і дана випадкова величина незалежна від K і неперервно розподілена на проміжку від 0 до 1, то тривалість майбутнього життя T застрахованої особи можна виразити як

$$T = K + S = (K + 1) - (1 - S)$$

З цього маємо

$$E((1+i)^{1-S}) = \int_0^1 (1+i)^u du = \bar{s}_{\overline{1}|} = \frac{i}{\delta}$$

$$\bar{A}_x = E(v^{K+1})E((1+i)^{1-S}) = \frac{i}{\delta}A_x$$

Отже бачимо, що обчислення разової нетто-премії для довічного страхування з виплатою у момент смерті особи звелось до обчислення разової нетто-премії для довічного страхування з виплатою в кінці року смерті.

Аналогічно можна отримати формули для розрахунку разової нетто-премії для тимчасового страхування життя.

$$\bar{A}_{x:\overline{n}|}^1 = \int_0^n v^t {}_t p_x \mu_{x+t} dt = \frac{i}{\delta} A_{x:\overline{n}|}^1$$

Для доживання формула виглядає наступним чином:

$$\bar{A}_{x:\overline{n}|} = \bar{A}_{x:\overline{n}|}^1 + A_{x:\overline{n}|}^{\frac{1}{n}} = A_{x:\overline{n}|} + \left(\frac{i}{\delta} - 1\right) A_{x:\overline{n}|}^1$$

Окрім розглянутих вище елементарних видів страхування життя, розглянемо також деякі загальні типи страхування життя. Якщо раніше розмір страхової суми був фіксованим і дорівнював 1, то тепер модифікуємо попередні означення довічного страхування життя для врахування зміни розміру страхової суми з року в рік з виплатою суми під кінець року смерті. Нехай c_j позначає страхову виплату, що має бути виплачена після j років дії страхування. Тоді величина Z виражена як

$$Z = c_{K+1}v^{K+1}$$

За своєю сутністю даний вид страхування відповідає сумі відкладених довічних страхувань життя, кожному з яких відповідає деяка стала страхова сума. Тоді для розрахунку разової нетто-премії застосовується формула

$$E(Z) = c_1 A_x + (c_2 - c_1) {}_1|A_x + (c_3 - c_2) {}_2|A_x + \dots$$

Аналогічно можемо виразити формулу для тимчасового страхування життя на n років, тільки використовуючи комбінацію тимчасових страхувань з різними страховими сумами які починаються в момент укладання договору:

$$E(Z) = c_n A_{x:\overline{n}|}^1 + (c_{n-1} - c_n) A_{x:\overline{n-1}|}^1 + (c_{n-2} - c_{n-1}) A_{x:\overline{n-2}|}^1 + \dots$$

Аналогічно розглянемо також ситуацію, яка частіше зустрічається на практиці – страхування з виплатою страхової суми в момент смерті застрахованої особи. В даному випадку припускаємо, що страхова сума в певний момент часу є функцією c від часу t . Тоді для довічного страхування життя Z визначається як

$$Z = c(t) \cdot v^T,$$

і разова нетто-премія розраховується як

$$E(Z) = \int_0^{\infty} c(t) v^t {}_t p_x \cdot \mu_{x+t} dt$$

Дане обчислення можна звести до обчислення в дискретній моделі:

$$E(Z) = \sum_{k=0}^{\infty} c_{k+1} v^{k+1} {}_k p_x q_{x+k}, c_{k+1} = E(c(k+S)(1+i)^{1-S} | K=k)$$

2.4 Розрахування нетто-премій на основі таблиць тривалості життя та аналіз результатів

На основі матеріалів з розділу було виконано розрахунки нетто-премій для різних груп населення з використанням мови програмування Python.

Нижче представлено створену програму для розрахунку нетто-премій для осіб різного віку та отриманий графік, за яким можна відстежити зміну у вартості нетто-премій для тимчасового страхування життя з терміном у три роки для осіб чоловічої статі двох країн – Ізраїлю та Боснії і Герцеговини. Розрахунки були виконані на основі таблиць тривалості життя та смертності загальної чоловічої популяції Ізраїлю [34] та Боснії і Герцеговини [35] станом на кінець 10-х років 21 століття.

```
[ ] life_table_israel = pd.read_excel("israel.xlsx")

[ ] life_table_bh = pd.read_excel("bh.xlsx")

[ ] def calculate_sum(i, x, n, data, limit):
    total_sum = 0
    if x < limit:
        for k in range(n):
            age_range_data = data[(data['age'] >= x) & (data['age'] <= x + k)]
            p_product = age_range_data['p'].prod()
            q_value = data.loc[data['age'] == x + k, 'q'].values[0]
            total_sum += p_product * ((1 / (1 + i)) ** (k + 1)) * q_value
        return total_sum

    def add_sum_column(df, limit):
        df['sum'] = df.apply(lambda row: 1 * (0.05/math.log(1 + 0.05)) * calculate_sum(0.05, row['age'], 3, df, limit), axis=1)

    add_sum_column(life_table_israel, 95)
    add_sum_column(life_table_bh, 95)
```

Рис.2.4.1. Програмний код для розрахування вартості нетто-премій для тимчасового страхування

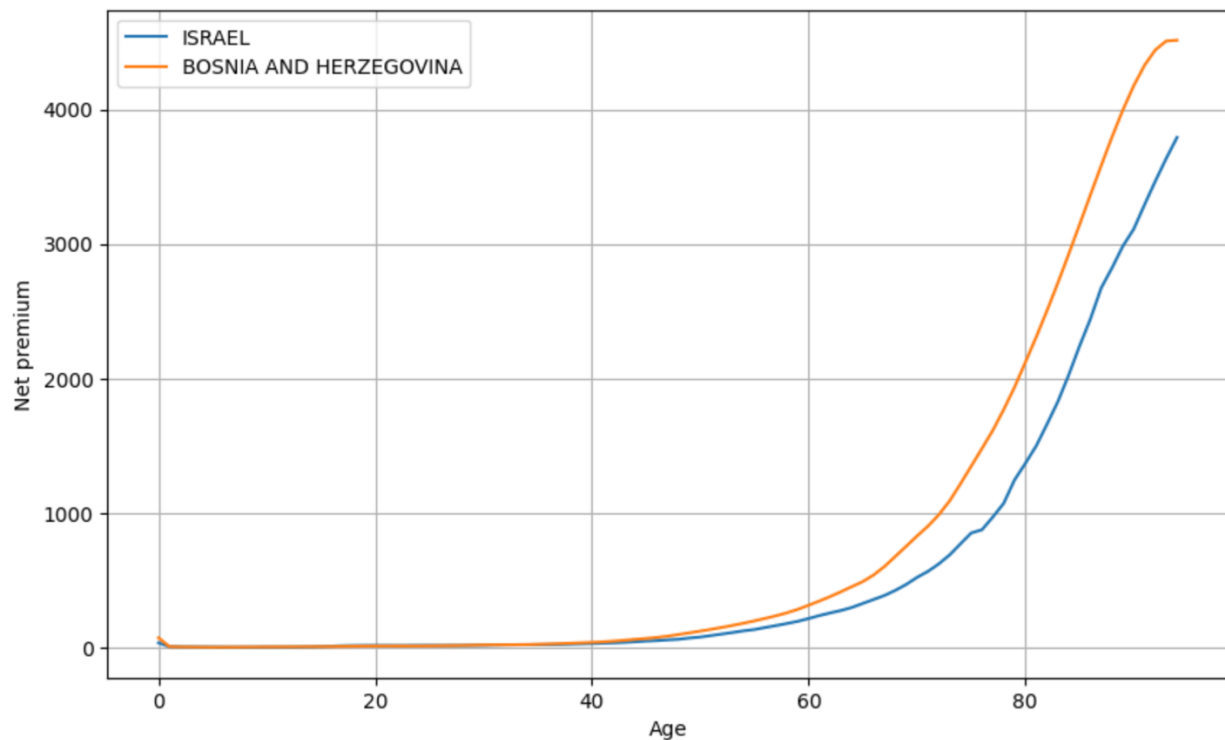


Рис.2.4.2. Графік вартості нетто-премій для тимчасового страхування життя для чоловічого населення Ізраїлю та Боснії і Герцеговини за віком

З графіку 2.4.2 видно, що нетто-премії для Боснії і Герцеговини стають значно вищими у старших вікових групах. Така різниця може бути пояснена декількома факторами, включаючи соціально-економічні умови, рівень медичного обслуговування та загальний стан здоров'я населення. Як вже було зазначено, одним

з важливих факторів, що впливають на тривалість життя та, відповідно, на вартість нетто-премій, є воєнні дії та геополітичні конфлікти. Історія Боснії і Герцеговини насичена конфліктами, включаючи боснійську війну 1992-1995 років, яка мала значний вплив на здоров'я та тривалість життя населення. Війна призводить до підвищення рівня смертності, що в свою чергу навіть через багато років може відобразитись на актуарних розрахунках та підвищувати вартість страхових полісів. З іншого боку, Ізраїль, незважаючи на власні зовнішньополітичні конфлікти, має добре розвинену систему охорони здоров'я та соціальне забезпечення, що сприяє вищій тривалості життя та, відповідно, нижчим ризикам смертності і відповідно нижчим нетто-преміям у порівнянні з Боснією і Герцеговиною. Важливо зазначити, що активні бойові дії в Ізраїлі тривали відносно давно, що дозволило країні сконцентруватися на відновленні та покращенні якості життя населення. Це показує, наскільки важливими є стабільність та соціальні інвестиції для покращення здоров'я населення та стабілізації ризиків підвищеної смертності населення.

Аналогічний підхід можна застосувати для простеження змін вартості нетто-премій впродовж років в певних регіонах для конкретних груп населення.

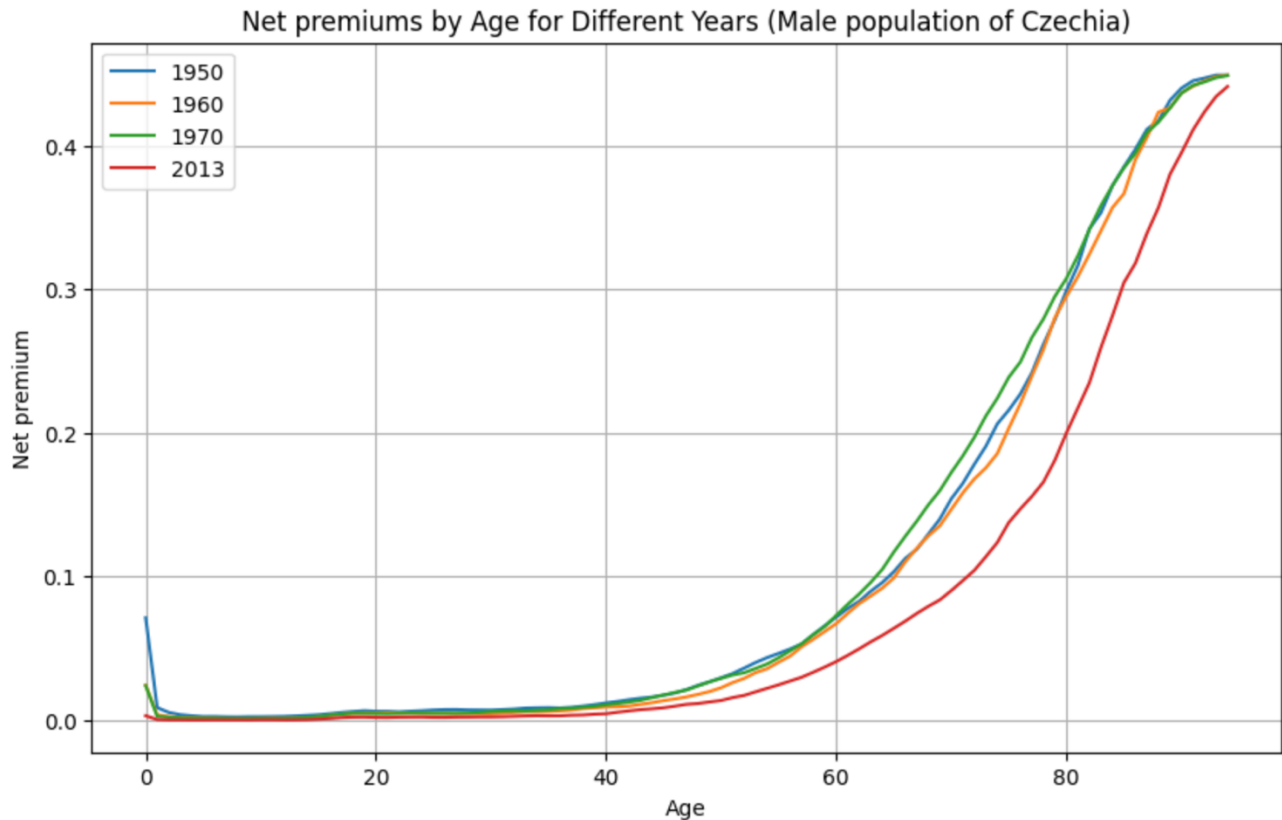


Рис.2.4.3. Графік вартості нетто-премій для тимчасового страхування життя на основі таблиць смертності чоловічої популяції Чеської Республіки (до 1993 року – Чехословаччини) для років 1950, 1960, 1970 та 2013 [36]

На графіку 2.4.3 зображено потенційні вартості нетто-премії за віком осіб чоловічої статі для різних років (1950, 1960, 1970, 2013) у Чехії. Таблиці смертності, на основі яких проводились розрахунки, включають роки після Другої світової війни та період Празької весни. Хоча сфера страхування життя існувала в Чехословаччині в середині 20-го століття, її структура і доступність страхування сильно залежали від політичного та економічного контексту. Епоха холодної війни та подальші політичні зміни у Східній Європі, включаючи Празьку весну 1968 року та остаточне падіння комунізму, мали значний вплив на економіку та здоров'я населення, а отже і на рівень смертності. Як вже було зазначено, вищий рівень смертності під час та одразу після війни може призвести до нижчої очікуваної тривалості життя, а отже, до помітної різниці у розрахунках вартості страхових

внесків. Десятиліття покращень у сфері охорони здоров'я, медичних технологій та рівня життя загалом збільшило очікувану тривалість життя, що в свою чергу відображається у нижчих вартостях для більшості вікових груп в останні роки (на графіку - 2013 р.).

Окрім використання таблиць тривалості життя для груп населення різних країн, можна також простежити різницю в ціні на страхування життя для окремих груп населення, наприклад для чоловіків-курців [37] та чоловіків, які не курять [38].

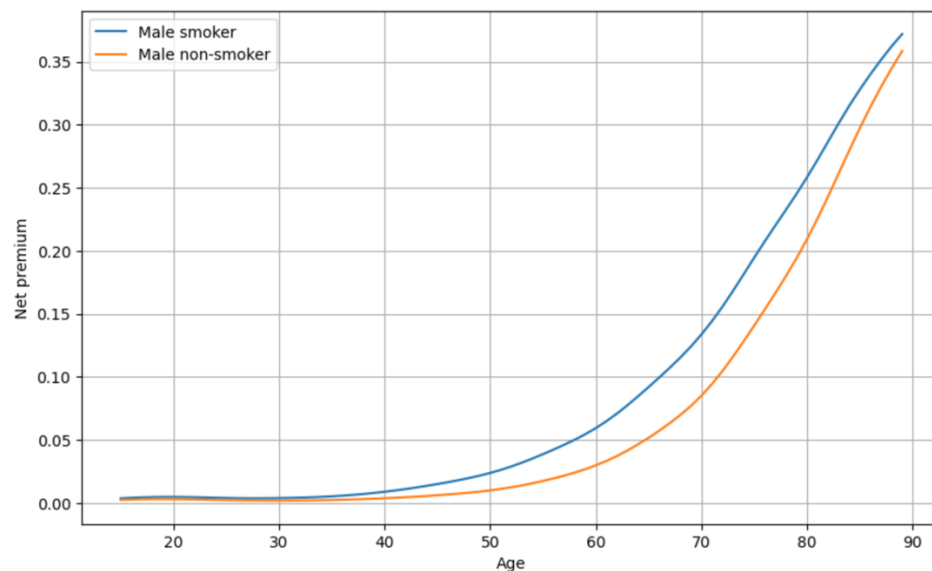


Рис.2.4.4. Графік вартості нетто-премій для чоловіків-курців та чоловіків, які не курять

В даному випадку значна різниця цін безпосередньо пов'язана з впливом негативних звичок на очікувану тривалість життя індивіда, який був досліджений попередньо в розділі 1.3. Як курці, так і некурці з віком стикаються зі збільшенням ризиків для здоров'я, але темпи погіршення стану здоров'я у курців є більш швидкими через ускладнення наслідків захворювань, пов'язаних з курінням. Отже, курці відносяться до групи підвищеного ризику, що призводить до вищих страхових внесків. Некурці, які зазвичай мають нижчі ризики для здоров'я, можуть очікувати нижчі страхові внески.

Розділ 3. Тривалість життя та страхування життя у воєнні та після воєнні часи – аналіз джерел та потенційний вплив на ситуацію в Україні

Війни мають катастрофічний вплив на населення – вони призводять до негайного збільшення смертності через бойові дії, і також спричиняють довготривалі наслідки, включаючи психологічні травми, голод, поширення інфекційних захворювань та погіршення доступу населення до медичної допомоги. Спираючись на дослідження, проведене в рамках цієї роботи та описане в розділах 1 та 2, можемо стверджувати про те, що результатами збройних конфліктів є значне зниження тривалості життя та зміни в сфері страхування життя. Аналіз історичних даних різних країн, що пройшли крізь різноманітні збройні конфлікти, дозволяє отримати краще розуміння специфіки воєнних реалій та передбачити потенційний вплив війни на тривалість життя населення та вартість страхування життя для України.

Дослідження, проведене Широ Хоріучі [39], було сфокусоване на впливі Першої світової війни на старші покоління в Німеччині. Воно виявило, що ветерани війни зазнали підвищеної смертності в старшому віці, порівняно з тими, хто не брав участі у бойових діях. Це пояснюється тривалими наслідками фізичних і психологічних травм, отриманих під час війни. Інше дослідження, проведене Яном Саарелою та Фьяларом Фіннасом [40], зосередилося на ветеранах Другої світової війни у Фінляндії. На відміну від ситуації після впливу Першої світової війни у Німеччині, у Фінляндії не спостерігалось значного підвищення довгострокової смертності серед ветеранів, хоча відразу після війни рівень смертності був високим. Це пов'язано з вчасним розвитком фармацевтики та покращенням медичної допомоги, що суттєво знизило рівень смертності у загальному населенні країни навіть у післявоєнну епоху. Дослідження впливу війни на таблиці смертності в Чехії, проведене Петром Мазоучем [41], показало, що війна спричинила короткострокове

підвищення рівня смертності, проте довготривалі наслідки, хоч і були помітними, були не були настільки вираженими, як очікувалося.

У дослідженні, продемонстрованому у розділах 1 та 2, чітко прослідковується залежність очікуваної тривалості життя індивіда на вартість різних видів страхування життя. Очевидно, що через це під час війни страхові компанії стикаються з підвищеним рівнем ризиків через збільшення смертності. Зміни в роботі страхових компаній, пов'язані зі збройними конфліктами, можуть полягати у перегляді умов полісів, підвищення премій та введення нових продуктів, таких як страхування від ризиків, пов'язаних з війною.

Згідно з дослідженням, проведеним The Conversation [42], війна в Україні станом на зараз вже призводить до підвищення рівня дитячої смертності та зниження тривалості життя загалом. Страхові компанії в Україні будуть все більше й більше стикатися з викликами, пов'язаними з оцінкою ризиків та визначенням доцільних страхових премій задля покриття ризиків. Така ситуація може призвести до тимчасового припинення продажу нових полісів або до значного підвищення премій, що зробить страхування життя менш доступним для населення, в особливості його вразливих соціальних груп. Окрім того, виключення воєнних ризиків (англійською: war exclusion clause) є стандартним положенням у багатьох страхових полісах по всьому світу, яке виключає покриття страхових випадків, спричинених воєнними діями [43]. Це положення може допомогти страховим компаніям захиститися від значних фінансових втрат під час війни, коли рівень смертності різко зростає. Імплементация практики виключення воєнних ризиків може призвести до того, що велика кількість сімей не отримає страхові виплати у разі смерті члена сім'ї через воєнні дії. Це може створити додаткові фінансові труднощі для населення в умовах війни. Водночас, це положення дозволяє страховим компаніям залишатися фінансово стійкими та продовжувати надавати послуги іншим клієнтам.

Висновки

В результаті дослідження було здійснено комплексний аналіз факторів, що впливають на тривалість життя та вартість страхування життя для різних груп індивідів. Основні результати практичної частини роботи підтверджують значущість таких факторів, як економічний розвиток, доступ до медичних послуг, зокрема кількість лікарів у регіоні, рівень доступності освіти. Використання методів математичного навчання дозволило визначити додаткові фактори впливу на тривалість життя на рівні індивідуального способу життя особи, серед яких значне місце займають спосіб життя, професія, наявність шкідливих звичок тощо. Ці фактори, в свою чергу, впливають на зниження або збільшення вартості страхування життя. Підтверджено, що в країнах з вищою очікуваною тривалістю життя страхові премії є нижчими, оскільки ризики, пов'язані з виплатами, є меншими. Окремий аналіз показав, що збройні конфлікти мають значний негативний вплив на тривалість життя населення та підвищують вартість страхування життя через зростання ризиків. Поточні події в Україні мають та матимуть подалі негативний вплив на рівень смертності серед населення країни, що вимагає адаптації страхових продуктів до нових умов.

Програми та моделі, створені під час практичної частини дослідження, можуть бути використані для прогнозування тривалості життя та розрахунків вартості страхування життя. Отримані дані можуть також слугувати основою для подальших досліджень у сфері демографії, медицини та актуарної науки.

Список використаних джерел

1. Gerber H. Life Insurance Mathematics / Hans Gerber., 1997. – 221 с. – (3).
2. R. Norberg. Basic Life Insurance Mathematics / Norberg Ragnar., 2001. – 149 с.
3. Castellares F. On the Gompertz–Makeham law: A useful mortality model to deal with human mortality / F. Castellares, S. Patrício, A. Lemonte. // Brazilian Journal of Probability and Statistics. – 2022. – №3. – С. 613–639.
4. Smith F. The Force of Mortality Function / F.C. Smith. // The American Mathematical Monthly. – 1948. – №5. – С. 277–284.
5. Bell F. Actuarial Study No. 116 [Електронний ресурс] / F. Bell, M. Miller – Режим доступу до ресурсу:
https://www.ssa.gov/oact/NOTES/as116/as116_IV.html.
6. Life Tables [Електронний ресурс] – Режим доступу до ресурсу:
https://bio.libretexts.org/Courses/Gettysburg_College/01%3A_Ecology_for_All/10%3A_Population_modeling/10.07%3A_Life_Tables.
7. McCarthy C. 5 risk factors that affect life expectancy [Електронний ресурс] / Claire McCarthy // The Harvard Gazette. – 2022. – Режим доступу до ресурсу:
<https://news.harvard.edu/gazette/story/2022/06/5-risk-factors-that-affect-life-expectancy/>.

8. MONDAL N. Impact of Socio-Health Factors on Life Expectancy in the Low and Lower Middle Income Countries [Электронный ресурс] / N. MONDAL, M. SHITAN – Режим доступа до ресурсу:
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4441932/>.
9. Christensen K. Determinants of longevity: genetic, environmental and medical factors [Электронный ресурс] / К. Christensen, J. Vaupel. – 1996. – Режим доступа до ресурсу: <https://pubmed.ncbi.nlm.nih.gov/9010380/>.
10. Life Expectancy By Country [Электронный ресурс] – Режим доступа до ресурсу: <https://www.datapandas.org/ranking/life-expectancy-by-country>.
11. HDI By Country [Электронный ресурс] – Режим доступа до ресурсу:
<https://www.datapandas.org/ranking/hdi-by-country>.
12. Estimating indirect mortality impacts of armed conflict in civilian populations: panel regression analyses of 193 countries, 1990–2017 [Электронный ресурс]. – 2020. – Режим доступа до ресурсу:
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7487992/>.
13. Life Expectancy [Электронный ресурс] – Режим доступа до ресурсу:
<https://ourworldindata.org/life-expectancy>.
14. Global Country Information Dataset 2023 [Электронный ресурс] – Режим доступа до ресурсу:
<https://www.kaggle.com/datasets/nelgiriyeewithana/countries-of-the-world-2023>.
15. pandas [Электронный ресурс] – Режим доступа до ресурсу:
<https://pandas.pydata.org/>.

16. Matplotlib: Visualization with Python [Электронный ресурс] – Режим доступа до ресурсу: <https://matplotlib.org/>.
17. SciPy [Электронный ресурс] – Режим доступа до ресурсу: <https://scipy.org/>.
18. Plotly Open Source Graphing Library for Python [Электронный ресурс] – Режим доступа до ресурсу: <https://plotly.com/python/>.
19. NumPy documentation [Электронный ресурс] – Режим доступа до ресурсу: <https://numpy.org/doc/stable/index.html>.
20. math — Mathematical functions [Электронный ресурс] – Режим доступа до ресурсу: <https://docs.python.org/3/library/math.html>.
21. statsmodels 0.14.1 - statistical models, hypothesis tests, and data exploration [Электронный ресурс] – Режим доступа до ресурсу: <https://www.statsmodels.org/stable/index.html>.
22. scikit-learn - Machine Learning in Python [Электронный ресурс] – Режим доступа до ресурсу: <https://scikit-learn.org/stable/>.
23. Spearman's Rank-Order Correlation [Электронный ресурс] – Режим доступа до ресурсу: <https://statistics.laerd.com/statistical-guides/spearmans-rank-order-correlation-statistical-guide.php>.
24. Burton A. OLS (Linear) Regression / Alexander L. Burton // The encyclopedia of research methods in criminology and criminal justice / Alexander L. Burton., 2021.

25. Liaw A. Classification and Regression by RandomForest [Електронний ресурс] / A. Liaw, M. Wiener. – 2001. – Режим доступу до ресурсу: https://www.researchgate.net/publication/228451484_Classification_and_Regression_by_RandomForest.
26. Татаренко О. В Україні скоротилась кількість медиків: кого бракує у галузі і скільки знайшли роботу в Польщі [Електронний ресурс] / О. Татаренко, А. Бачук, І. Власюк // Суспільне Рівне. – 2023. – Режим доступу до ресурсу: <https://suspilne.media/rivne/469031-v-ukraini-skorotilas-kilkist-medikiv-kogo-brakue-u-galuzi-i-skilki-znajsli-robotu-v-polsi/>.
27. Predict Mortality/Death Rate [Електронний ресурс] – Режим доступу до ресурсу: <https://www.kaggle.com/datasets/rajanand/mortality/data>.
28. Goel M. Understanding survival analysis: Kaplan-Meier estimate / M. Goel, P. Khanna, J. Kishore. // International Journal of Ayurveda Research. – 2010.
29. Types Of Life Insurance Policies [Електронний ресурс] // THE OFFICIAL WEBSITE OF NEW YORK STATE – Режим доступу до ресурсу: https://www.dfs.ny.gov/consumers/life_insurance/types_of_policies.
30. Шевченко Л. Страховик та страхувальник - права і обов'язки [Електронний ресурс] / Лідія Шевченко // British Insurance. – 2021. – Режим доступу до ресурсу: <https://britishinsurance.com.ua/article/strakhovik-ta-strakhuvalnik---prava-i-obov-yazki>
31. ФІНАНСОВИЙ СЛОВНИК [Електронний ресурс] // PORSCHE FINANCE GROUP – Режим доступу до ресурсу: <https://www.porschefinance.ua/servisy/finansovyy-slovnyk>.

- 32.Кравченко В. Брутто-премія [Електронний ресурс] / Віталій Кравченко // LivingFo. – 2022. – Режим доступу до ресурсу: <https://livingfo.com/brutto-premiia>
- 33.Козьменко О. В. Актуарні розрахунки / О. В. Козьменко, О. В. Кузьменко. – Суми: Університетська книга, 2011. – 224 с.
- 34.COMPLETE LIFE TABLES OF ISRAEL [Електронний ресурс]. – 2023. – Режим доступу до ресурсу:
https://www.cbs.gov.il/he/publications/DocLib/2023/1911_life_tables_2017_2021/e_print.pdf
- 35.DETAILED LIFE TABLE FOR THE POPULATION OF FEDERATION OF BOSNIA AND HERZEGOVINA [Електронний ресурс]. – 2017. – Режим доступу до ресурсу: https://fzs.ba/wp-content/uploads/2017/12/DETAILED-LIFE-TABLE-2012-2014_01122017n.pdf.
- 36.Czechia, Life tables (period 1x1), Males [Електронний ресурс]. – 2023. – Режим доступу до ресурсу:
https://www.mortality.org/File/GetDocument/hmd.v6/CZE/STATS/mltper_1x1.txt
- 37.1980 CSO Basic Table - Male Smoker, ANB [Електронний ресурс] – Режим доступу до ресурсу: <https://mort.soa.org/ViewTable.aspx?&TableIdentity=22>.
- 38.1980 CSO Basic Table - Male Nonsmoker, ANB [Електронний ресурс] – Режим доступу до ресурсу: <https://mort.soa.org/ViewTable.aspx?&TableIdentity=21>.

39. Shiro H. The long-term impact of war on mortality: Old-age mortality of the First World War survivors in the Federal Republic of Germany [Электронный ресурс] / Horiuchi Shiro. – 1983. – Режим доступа до ресурсу: https://www.researchgate.net/publication/11142996_The_long-term_impact_of_war_on_mortality_Old-age_mortality_of_the_First_World_War_survivors_in_the_Federal_Republic_of_Germany.
40. Jan S. Long-term Mortality of War Cohorts: The Case of Finland [Электронный ресурс] / J. Saarela, F. *Finnäs* // European Journal of Population / Revue européenne de Démographie. – 2011. – Режим доступа до ресурсу: <https://link.springer.com/article/10.1007/s10680-011-9246-x>.
41. Petr M. WAR EFFECT TO COHORT LIFE TABLES IN THE CZECH REPUBLIC [Электронный ресурс] / Mazouch Petr. – 2016. – Режим доступа до ресурсу: https://www.researchgate.net/publication/316990693_WAR_EFFECT_TO_COHORT_LIFE_TABLES_IN_THE_CZECH_REPUBLIC.
42. Ukraine war: the devastating effects of conflict on infant mortality rates – new research [Электронный ресурс] // The Conversation. – 2023. – Режим доступа до ресурсу: <https://theconversation.com/ukraine-war-the-devastating-effects-of-conflict-on-infant-mortality-rates-new-research-203187>.
43. Kagan J. What Is a War Exclusion Clause in an Insurance Contract? [Электронный ресурс] / Julia Kagan // Investopedia. – 2023. – Режим доступа до ресурсу: <https://www.investopedia.com/terms/w/war-exclusion-clause.asp>.