



Doi: <https://doi.org/10.15407/dse2018.03.056>

УДК 314.84

JEL CLASSIFICATION: J11

## П.Є. ШЕВЧУК

канд. екон. наук, старш. наук. співроб.  
Інститут демографії та соціальних  
досліджень ім. М.В. Птухи НАН України  
01032, м. Київ, бул. Т. Шевченка, 60  
E-mail: pavlo-shevchuk@ukr.net

## ПОБУДОВА ТАБЛИЦЬ СМЕРТНОСТІ НА СУБРЕГІОНАЛЬНОМУ РІВНІ

*У зв'язку з розширенням повноважень місцевого самоврядування зростає потреба в адекватній оцінці соціально-економічної ситуації на низовому рівні, в тому числі її демографічної складової. Проблема, з якою стикаються дослідники під час розрахунку показників дожиття для малих популяцій полягає в ненадійності таких оцінок, що випливає зі стохастичної природи демографічних процесів. Тому для обчислення, зокрема, таблиць смертності малолюдних сукупностей, зростає потреба в явному окресленні границь довірчих інтервалів, у яких можуть знаходитися показники дожиття. Існує два підходи до розв'язання цієї проблеми: знаходження стандартної похибки за відповідними формулами та знаходження границь довірчих інтервалів на основі моделювання чисел померлих за методом Монте-Карло. Більш поширеним у практиці статистичних відомств є метод, розроблений Ч.Л. Чангом. Але він не враховує внеску відкритого вікового інтервалу. Ця проблема знайшла свій розв'язок у роботах П. Сіллокса та ін., а також Е. Ло та ін. У цьому дослідженні були використані дані про вікову структуру та природний рух кількох адміністративних районів України за період 2002–2015 рр. З урахуванням розподілу за статтю та типом поселення було досліджено 364 таблиці смертності для сукупностей від 2810 до 46 113 осіб. Стандартна похибка не перевищила 4,5 року для очікуваної тривалості життя при народженні. Мінімальна величина стандартної похибки склала 0,7 року для населення 37–42 тис. осіб. Для населення більше 40 тис. стандартна похибка не перевищила 0,85 року. Для таблиць смертності по відповідних областях стандартна похибка очікуваної тривалості життя при народженні не перевищила 0,36 року для населення близько 200 тис. осіб. Відкритий віковий інтервал дає додатковий внесок у дисперсію очікуваної тривалості життя при народженні 10–20 % є у понад половини досліджених таблиць. Показано, що порівняння величини очікуваної тривалості життя населення районів навіть однієї області не завжди дає певне уявлення, в якому з них смертність насправді нижче. Цей факт треба враховувати у порівняльній оцінці демографічного розвитку територій субрегіонального рівня.*

**Ключові слова:** таблиця смертності, стандартна похибка, довірчий інтервал.

© ШЕВЧУК П.Є., 2018

*П.Є. Шевчук*

канд. экон. наук, старш. наук. сотр.  
Институт демографии и социальных  
исследований им. М.В. Птухи НАН Украины  
01032, г. Киев, бул. Т. Шевченко, 60  
E-mail: pavlo-shevchuk@ukr.net

## ПОСТРОЕНИЕ ТАБЛИЦ СМЕРТНОСТИ НА СУБРЕГИОНАЛЬНОМ УРОВНЕ

*В связи с расширением полномочий местного самоуправления возрастает потребность в адекватной оценке социально-экономической ситуации на низовом уровне, в том числе ее демографической составляющей. Проблема, с которой сталкиваются исследователи, рассчитывая показатели дожития для малых популяций, состоит в ненадежности таких оценок, что следует из стохастической природы демографических процессов. Поэтому при расчете таблиц смертности для малочисленных совокупностей возрастает потребность в явном очерчивании границ доверительных интервалов, в которых могут находиться показатели дожития. Существует два подхода к решению данной проблемы: нахождения границ доверительных интервалов на основе моделирования чисел умерших по методу Монте-Карло. Более распространен в практике статистических ведомств метод, разработанный Ч.Л. Чангом. Однако он не учитывает вклада открытого возрастного интервала. Эта проблема нашла свое решение в работах П. Силкокса и др., а также Э. Ло и др. В этом исследовании были использованы данные о возрастной структуре и естественном движении нескольких административных районов Украины за период 2002–2015 гг. С учетом распределения по полу и типу поселения были исследованы 364 таблицы смертности для совокупностей от 2810 до 46113 человек. Стандартная погрешность не превысила 4,5 года для ожидаемой продолжительности жизни при рождении. Минимальная величина стандартной ошибки составила 0,7 года для населения 37–42 тыс. чел. Для населения более 40 тыс. стандартная ошибка не превысила 0,85 года. Открытый возрастной интервал дает дополнительный вклад в дисперсию ожидаемой продолжительности жизни при рождении 10–20 % более, чем в половине исследованных таблиц. Показано, что сравнение величины ожидаемой продолжительности жизни населения районов даже одной области не всегда дает корректное представление, в каком из них смертность на самом деле ниже. Этот факт следует учитывать в сравнительной оценке демографического развития территорий субрегионального уровня.*

**Ключевые слова:** *таблица смертности, стандартная ошибка, доверительный интервал.*

*P.E. Shevchuk*

PhD (Economics), Leading scientist  
Ptoukha Institute for Demography and Social  
Studies of the National Academy of Sciences of Ukraine  
01032, Ukraine, Kyiv, blvd Taras Shevchenko, 60  
E-mail: pavlo-shevchuk@ukr.net

## THE LIFE TABLES CONSTRUCTION AT SUBREGIONAL LEVEL

*In connection with the expansion of powers of local authority, there is a growing need for an adequate assessment of the socio-economic situation at the grassroots level, including its demographic component. The problem faced by researchers in calculating survival rates for small populations is the uncertainty of such estimates, which appears from the stochastic nature of demographic processes. When calculating the life tables for small populations, there is a growing need for a clear outline of the boundaries of confidence intervals in which indicators can exist. There are two approaches to solve this problem: finding the standard error with the corresponding formulas and finding the boundaries of confidence intervals based on the modelling of the death numbers via the Monte Carlo method. More common in the practice of statistical departments is the method developed by Ch.L. Chiang. But he does not take into account the contribution of the open age interval. This problem has found its solution in the works of P.B.S. Silcocks et al. and E. Lo et al. In this study data on the age structure and vital statistics of several administrative districts of Ukraine for the period of 2002–2015 were used. Taking into account the distribution by sex and urban-rural areas, 364 life tables for populations from 2810 to 46113 persons were investigated.*

*The standard error has not exceeded 4.5 years for the expected life expectancy at birth. The minimum value of the standard error was 0.7 years for the population of 37–42 thousand persons. For the population more than 40 thousand standard error does not exceed 0.85 years. For the life tables in the respective administrative regions (oblasts), the standard error of the expected life expectancy at birth did not exceed 0.36 years for the population of about 200 thousand persons. The open age interval gives an additional contribution to the dispersion of the life expectancy at birth from 10 to 20 % in more than half of the studied tables. It is also shown that the point estimates of the life expectancy by districts of even the same region does not always give a precise idea, in which of them mortality is actually lower because of too wide confidence intervals.*

**Keywords:** life table, standard error, confidence interval.

**Постановка й актуальність проблеми.** Проблема побудови таблиці смертності для малочисельного населення очевидна будь-кому, обізнаному із законом великих чисел: чим менша чисельність населення, тим менш надійними є його статистичні оцінки, зокрема ті, які містяться у таблиці смертності. Таким чином, таблиця смертності, розрахована для порівняно малочисельного населення, потребує явного окреслення границь довірчих інтервалів. Натомість поточна практика зазвичай оминає етап оцінювання надійності демографічних показників.

Показники таблиці смертності використовують для аналізування демографічної ситуації, вони можуть слугувати індикаторами соціально-економічного становища в регіоні. Демографічні характеристики населення є важливими показниками, на основі яких можуть бути ухвалені управлінські рішення. Тому такий інструмент аналізу як таблиця смертності є необхідним для керівників різного рівня. Збільшення компетенцій на місцевому рівні, згідно з Законом України «Про внесення змін до деяких законодавчих актів України щодо розширення повноважень органів місцевого самоврядування та оптимізації надання адміністративних послуг» № 888-VIII від 10.12.2015 [1] зумовлює посилення уваги до статистичних даних відповідного територіального рівня. З викладеного очевидно, що наведення границь довірчих інтервалів для основних демографічних показників є актуальним.

**Аналіз останніх досліджень і публікацій та невирішені раніше частини загальної проблеми.** Першим струнку методологію оцінки статистичної похибки показників дожиття й очікуваної тривалості життя розробив Ч.Л. Чанг (Ch.L. Chiang), і виклав її у монографії [2]. Цим методом користуються статистичні служби Канади [3] та Великої Британії.

П. Сілкокс та ін. (Silcocks P.V.S. et al.) уперше привернули увагу до проблеми оцінки похибки у відкритому віковому інтервалі, яку Ч.Л. Чанг вважав нульовою, і запропонували власну формулу [4, с. 43].

Основними цілями Б. Тосон (B. Toson) і А. Бейкера (A. Baker) були порівняння методології обчислення таблиць смертності та знайдення придатної для використання у випадку малочисельного населення; визначення мінімальної чисельності, нижче якої очікувану тривалість життя розраховувати не доцільно, та розгляд впливу вікових груп із відсутністю смертних випадків. З приводу вибору методу обчислення таблиць вони знайшли, що відмінності на результатах майже не позначаються. Для встановлення мінімальної чисельності населення, для якої є сенс розраховувати таблицю смертності, вони виконали серію тестів із різною чисельністю населення – від 1 тис. до 80 тис. осіб. Тести являли собою симуляції за методом Монте-Карло для генерування значень тривалості життя з населенням, статеві-вікова структура якого була зафіксована відповідно до населення Англії та Уельсу у середині 1991 р. Покладені в основу коефіцієнти смертності базувалися на Англійських таблицях смертності, які є трирічними таблицями смертності для Англії та Уельсу за роки, суміжні з роком

перепису. На основі моделювання визначено, що п'ять тисяч осіб є мінімальною величиною [5, с. 21], для якої доцільно обчислювати тривалість життя. За меншої чисельності стандартна похибка швидко зростає, а ширина довірчих інтервалів сягає понад 15 років для населення в 1 тис. осіб [5, с. 14]. Зазначено також, що навіть якщо публікація довірчих інтервалів для тривалості життя за виборчими округами виявиться нездійсненою, ідея надання мірила варіації до цих показників матиме користь для інших регіональних результатів [5, с. 6], що є слушним і для України.

Д. Айрес (D. Eayres) та І. Вільямс (E. Williams) підтвердили мінімальний розмір населення у п'ять тисяч, для якого можна розрахувати таблицю смертності з прийнятною довірчою ймовірністю [6, с. 245] та відмітили, що зі зниженням розміру популяції таблиці смертності мають тенденцію до завищення очікуваної тривалості життя. Окремо увагу вони приділили проблемі вікових груп із відсутністю смертних випадків і дійшли висновку, що методи підстановки теоретичних величин не покращують оцінку.

С. Щербов (S. Scherbov) і Д. Едієв (D. Ediev) виконали моделювання, основане на використанні доступних таблиць смертності чоловіків і жінок деяких країн Європи та Японії [7]. Для кожної таблиці смертності вони знайшли вікові структури стабільного населення, які відповідають річним темпам росту населення, %: -2, -1, 0, 1 та 2. Чисельність цих населень була масштабована до восьми рівнів: 1, 5, 10, 25, 50, 100, 250 тис. і 1 млн осіб. Також дослідники порівняли показники повних і скорочених таблиць смертності та показали, що використання скороченої таблиці порівняно з повною суттєво менше впливає на точність оцінки [7, с. 529], ніж вибір відкритого інтервалу.

Е. Ло та ін. (Lo E. et al.) розвинули попередні дослідження та зосередили увагу саме на проблемі врахування відкритого вікового інтервалу таблиці смертності. Вони показали, що смертність у цьому інтервалі може істотно впливати [8, с. 411–414] на дисперсію тривалості життя, особливо зі зменшенням кількості людино-років під ризиком смерті. Також вони навели у вказаній праці виправлену формулу для оцінки дисперсії [8, с. 404].

Таким чином, у названих роботах проблема побудови таблиць смертності для малочисельних регіонів розв'язувалася в цілому, для будь-яких випадків, за винятком [8], у якій обчислення здійснено на реальних даних регіонів Канади. В Україні населення за віковими групами в малих регіонах розподілено вкрай нерівномірно. Тому доцільно розрахувати таблиці смертності з довірчими інтервалами для фактичних даних населення малих регіонів та порівняти з їхніми стаціонарними аналогами.

**Мета** статті: дослідити можливості побудови таблиць смертності для населення окремих адміністративних районів і міст України залежно від їх людності; вивчити залежність ширини довірчих інтервалів для показників очікуваної тривалості життя населення від чисельності населення під ризиком смерті; з'ясувати, наскільки виправданим є використання вікової структури теоретичного (стаціонарного) населення як заміника реального повікового спектру населення.

**Новизною** статті є порівняльний аналіз стандартного відхилення очікуваної тривалості життя при народженні для реальних популяцій субрегіонального рівня та їх теоретичних еквівалентів, а також практичні рекомендації щодо побудови таблиць смертності для субрегіонального рівня з обчисленням довірчих інтервалів.

**Виклад основного матеріалу.** На відміну від теоретичних популяцій або населення Англії та Уельсу в цілому, використаних у цитованих вище роботах [4–7], віковий спектр населення України та її регіонів є доволі строкатим, що зумовлено історичними особливостями демографічного розвитку. Зокрема, значні відмінності від

загальноукраїнської статеві-вікової структури можна знайти в малолюдних адміністративних районах із переважно сільським населенням. Більше того, навіть реальна вікова структура населення всієї України на межі тисячоліть сильно відрізнялася від свого стабільного еквівалента. Хоча С. Щербов і Д. Едієв погоджуються, що вікова структура може значно впливати на встановлення довірчих інтервалів і значимість варіації, для розрахунків вони обирають вікові структури стабільного населення просто масштабуючи його до різної чисельності. Населення всієї Англії теж більш рівномірно розподілене за віком, ніж в окремих її графствах.

**Статистичні дані.** Обчислення здійснено на основі даних поточної (післяпереписної) оцінки чисельності населення за статтю та віком за період 2002–2015 рр. для Горохівського (Волинської), Миронівського (Київської), Біляївського, Захарівського (Одеської), Дергачівського (Харківської області) районів, м. Новоград-Волинський (Житомирської) та Новокаховської міськради (Херсонської області). Вибір саме цих адміністративних одиниць ґрунтується на досвіді попередніх розрахунків, які показали, що смертність у деяких із цих районів напрочуд точно відповідає обласному рівню (з урахуванням особливостей статеві-вікової структури), а інших навпаки – сильно відрізняється. Тому вирішено детальніше їх дослідити. Чисельність населення обраних адміністративних одиниць варіює від 2810 до 46113 осіб. Оскільки розглянутий період охоплює 14 календарних років, то чисельність досліджуваних популяцій становить 364 одиниці.

**Деталізація статеві-вікової структури:** за однорічними віковими групами до 69 років і укрупнена вікова група 70 років і старше, за винятком даних перепису 2001 р. Також були доступні числа померлих за статтю, віком (за однорічними віковими групами) і роком народження (за однорічними когортами). Оцінка чисельності населення за віковими групами в межах групи 70 і старше була здійснена за когортами від даних останнього перепису. Якщо виявлялось, що когорта вимирає повністю, і значення ставало від'ємним, вживали гіпотезу, що померли ті, хто прибув у цьому ж календарному році. Звісно, вони могли прибути й у попередні роки та не зареєструватися: гіпотеза є складнішою але не більш правомірною.

**Метод.** Скорочені таблиці смертності для вказаних вище адміністративних одиниць обчислено демографічним методом (на основі вікових коефіцієнтів смертності). Якщо в певному інтервалі (навіть п'ятирічному) не трапилося смертей, то для нього знаходили гіпотетичне число смертей, виходячи з коефіцієнта смертності в цьому віці для всього регіону. Такий підхід використовує статистична служба Канади [3, с. 5], а також Е. Ло та ін. [8, с. 450]. Звісно, це штучно завищує смертність. Проте нульовий коефіцієнт смертності визначає нульову ймовірність померти, хоча очевидно, що ймовірність смерті існує в будь-якому віці. Тобто, відсутність смертей у деяких вікових інтервалах зумовлює певне завищення очікуваної тривалості життя й, відповідно переоцінку «успіхів» [9, с. 51], досягнутих малими популяціями. Корекція саме числа померлих (а не ймовірності смерті, що було б логічно) в контексті даного дослідження зумовлена потребою в додатному числі смертей, оскільки їх відсутність, згідно з формулою (1) [2, с. 79], призводить до ділення на нуль:

$$S_{q_x}^2 = \frac{1}{D_x} q_x^2 (1 - q_x), \quad (1)$$

де  $S_{q_x}^2$  – вибіркова дисперсія ймовірності смерті,  $D$  – кількість померлих,  $q$  – ймовірність смерті,  $x$  – вік.

Тому, якщо у даних про кількість смертей у віковому інтервалі трапляється нуль, то відповідну чисельність населення слід помножити на коефіцієнт смертності у цьому ж віці в регіоні, частиною якого є досліджуване населення. Якщо в досліджуваному регіоні відсутнє населення у певному віці (навіть у п'ятирічному віковому інтервалі), що трапляється в найстарших вікових групах, то відкритий віковий інтервал слід знижувати до того віку, в якому коефіцієнт смертності буде вищий за попередній (молодший) віковий інтервал. Серед досліджуваного набору популяцій для 42 відкритий інтервал становить 100 років і старше, для 95 – відкритий інтервал 95 років і старше, для 125 – 90 років і старше, для 77 – 85 років і старше, для 25 виявилось необхідним знизити початок відкритого інтервалу до 80 років. Для зручності подальшого аналізу всі таблиці були приведені до відкритого інтервалу 85 років і старше.

Через громіздкість розрахунків і відповідних вихідних таблиць для наочного прикладу було обрано лише міських жінок Біляївського району Одеської області в 2003 р. Причинами цього були їхня досить мала чисельність (8428,5 осіб на середину 2003 р.) для генерування досить широкого довірчого інтервалу, але не надто мала, щоби демонструвати приклад, не рекомендований для розрахунків. Рік для прикладу було обрано 2003, оскільки він досить близький до даних перепису населення, але довірчий інтервал трохи ширший, ніж трапився у 2002 р.

Так, у наведеному прикладі для знаходження ненульових значень померлих у вікових групах 0, 1–4, 10–14, 15–19 та 20–24 роки чисельність міських жінок Біляївського району була помножена на коефіцієнти смертності міських жінок Одеської області в 2003 р. у відповідному віці. Результат наведено у ст. 4 (таблиця). Таким чином, із 364 досліджуваних популяцій до 14 не було додано померлих, до 200 додано менше 1 % померлих, до 128 – від 1 до 5 % і до 22 одиниць було додано більше 5 % додаткових померлих.

Слід зазначити, що існують й інші підходи до розв'язання проблеми відсутності смертей в інтервалі. Так, П. Сілкок із співавторами в таких випадках пропонують вводити додатне мале число 0,693 або 3 [4, с. 40]. В. Тосон і А. Бейкер [5, с. 11] і Є. Андреев (E. Andreev) та В. Школьніков (V. Shkolnikov) [10, с. 2] використали метод Монте-Карло для симуляції випадкових величин чисел померлих.

Частку вікового інтервалу, яку прожили ті, хто в ньому помер (ст. 5 у таблиці 1) розраховано на основі повних таблиць смертності відповідного регіону:

$${}_n a_x = \frac{{}_n L_x - l_{x+n}}{l_x - l_{x+n}}, \quad (2)$$

де  ${}_n L_x$  – стаціонарне населення у віці від  $x$  до  $x+n$ ;  $n$  – ширина вікового інтервалу,  $l_x$  – кількість тих, хто доживає до початку вікового інтервалу  $x$ ;  $l_{x+n}$  – кількість тих, хто доживає до початку вікового інтервалу  $x+n$ .

Величину  ${}_1 a_0$  отримуємо прямо із повної таблиці за алгоритмом, описаним у Ч.Л. Чанга [2, с. 143]. Значення відкритого вікового інтервалу приймаємо тотожним очікуваній тривалості життя при його досягненні, поділеному на його ширину.

Вікові коефіцієнти смертності (ст. 6) обчислюються за звичайною формулою співвідношення числа померлих (ст. 4) до середньорічної чисельності населення в тому ж віці (ст. 2).

Таблиця. Розрахунок дисперсії очікуваної тривалості життя на прикладі міських жінок Білявського р-ну Одеської обл. в 2003 р.

Вік $x$	Дані						Таблиця смертності						Дисперсія очікуваної тривалості життя, за Ч.Л. Чаном				Поправка з урахуванням відкритого інтервалу			
	$n$	$P$	$D$	$D$ ( <i>corr</i> )	$a$	$m$	$q$	$l$	$e$	Вибіркова дисперсія $q_x$ (формула 1)	(формула 3)	Кумулята ст. II знизу	Дисперсія очікуваної тривалості життя (ст. 12; (ст. 8)²)	$P_{ov}$	Поправочний член (формула б)	Дисперсія очікуваної тривалості життя (ст. 13 + ст. 15)				
A	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16				
0	1	67	0	0,54990	0,10	0,00827	0,00821	100000	70,55	0,00012	6131674935	34823930986	3,4824	0,0641	0,0072	3,4896				
1	4	297	0	0,19801	0,35	0,00067	0,00266	99179	70,14	0,00004	1667390411	28692256051	2,9169	0,0646	0,0073	2,9242				
5	5	461	1	0,50	0,00217	0,01080	0,01080	98915	66,32	0,00012	4695086289	27024865641	2,7621	0,0648	0,0073	2,7694				
10	5	589	0	0,16427	0,54	0,00028	0,00139	97847	62,01	0,00001	399354659	22329779351	2,3323	0,0655	0,0075	2,3398				
15	5	527	0	0,22934	0,50	0,00044	0,00217	97710	57,10	0,00002	587818661	21930424693	2,2970	0,0656	0,0075	2,3045				
20	5	568	0	0,51236	0,57	0,00090	0,00450	97498	52,22	0,00004	921023949	21342606032	2,2452	0,0658	0,0076	2,2528				
25	5	619	2	0,55	0,00323	0,01605	0,01605	97059	47,44	0,00013	2461729945	20421582083	2,1678	0,0661	0,0076	2,1754				
30	5	602	3	0,53	0,00499	0,02465	0,02465	95501	43,17	0,00020	3108631354	17959852138	1,9692	0,0671	0,0079	1,9771				
35	5	567	2	0,50	0,00353	0,01748	0,01748	93147	39,19	0,00015	1817721971	14851220785	1,7117	0,0688	0,0083	1,7200				
40	5	609	6	0,53	0,00985	0,04815	0,04815	91519	34,85	0,00037	3523666420	13033498813	1,5561	0,0701	0,0086	1,5647				
45	5	649	4	0,53	0,00617	0,03040	0,03040	87112	31,47	0,00022	1503972763	9509832394	1,2532	0,0736	0,0095	1,2626				
50	5	670	6	0,54	0,00896	0,04388	0,04388	84465	27,38	0,00031	1456949694	8005859631	1,1222	0,0759	0,0101	1,1322				
55	5	385	4	0,55	0,01039	0,05076	0,05076	80759	23,51	0,00061	1909265608	6548909937	1,0041	0,0794	0,0110	1,0151				
60	5	576	8	0,51	0,01390	0,06723	0,06723	76660	19,62	0,00053	1033806461	4639644329	0,7895	0,0836	0,0122	0,8017				
65	5	448	9	0,53	0,02011	0,09600	0,09600	71506	15,85	0,00093	1011229087	3603837868	0,7048	0,0897	0,0140	0,7189				
70	5	334	12	0,55	0,03597	0,16626	0,16626	64642	12,25	0,00192	1047940414	2592608781	0,6205	0,0992	0,0172	0,6376				
75	5	272	18	0,51	0,06624	0,28522	0,28522	53894	9,15	0,00323	797130135	1544668366	0,5318	0,1190	0,0247	0,5565				
80	5	118	9	0,49	0,07647	0,32011	0,32011	38522	6,78	0,00774	464846213	747538232	0,5037	0,1664	0,0484	0,5521				
85	5	51	13	0,46	0,25489	0,75522	0,75522	26191	3,82	0,01074	282692019	282692019	0,4121	0,2448	0,1047	0,5168				
90	15	24	7	0,23	0,28593	1,00000	1,00000	6411	3,50	0	0	0	0	1	1,7474	1,7474				

Закінчення таблиці

Вік	3 дисперсії чисельності населення			Дисперсія з негативною біноміальною наддисперсією							Дисперсія з негативною біноміальною наддисперсією та врахуванням відкритого вікового інтервалу			
	Другий доданок (добуток квадратів) формули 7	Дисперсія оцікваної тривалості життя з поправкою на похибку чисельності населення (ст. 16 + ст. 17)	Формула (9) Р	Р <sup>3</sup> ст. 19 – ст. 2	ст. 2 : (ст. 19 × ст. 20)	(ст. 8) <sup>2</sup> × [(1 – ст. 5) × ст. 1 + наступний рядок ст. 9] 2 × ст. 21	Кумулята ст. 22 знизу	Негативна біноміальна дисперсія, (ст. 8) <sup>2</sup>	ст. 142 : (ст. 19 × ст. 20) × (ст. 19 / ст. 20)	Дисперсія оцікваної тривалості життя з наддисперсією та врахуванням відкритого вікового інтервалу, ст. 24 + ст. 25	Стандартне відхилення, ст. 26	Внесок останнього вікового інтервалу з наддисперсією, (ст. 26 : 13 – 1) %		
А	17	18	19	20	21	22	23	24	25	26	27	28		
0	0,00003	3,48961	67	66	0,00012	623358,3993	42742723287	4,27427	0,02258	4,29686	2,07289	23,4%		
1	0,00003	2,92425	74	74	0,00004	1676303795	36509139294	3,71159	0,02296	3,73455	1,93250	28,0%		
5	0,00003	2,76946	93	92	0,00012	4798163846	34832835499	3,56011	0,02308	3,58319	1,89293	29,7%		
10	0,00003	2,33986	118	118	0,00001	400471006	30034671652	3,13710	0,02359	3,16069	1,77783	35,5%		
15	0,00003	2,30458	106	105	0,00002	590382257	29634200646	3,10392	0,02365	3,12758	1,76850	36,2%		
20	0,00003	2,25279	114	113	0,00004	929372003	29043818389	3,05535	0,02376	3,07911	1,75474	37,1%		
25	0,00003	2,17544	125	123	0,00013	2542706950	28114446386	2,98439	0,02397	3,00837	1,73446	38,8%		
30	0,00003	1,97709	122	119	0,00021	3267740703	25571739435	2,80377	0,02476	2,82853	1,68182	43,6%		
35	0,00004	1,72000	114	112	0,00016	1882982947	22303998732	2,57065	0,02603	2,59668	1,61142	51,7%		
40	0,00004	1,56472	125	119	0,00041	3889153958	20421015785	2,43813	0,02696	2,46509	1,57006	58,4%		
45	0,00004	1,26268	132	128	0,00024	1599747209	16531861827	2,17852	0,02976	2,20828	1,48603	76,2%		
50	0,00004	1,13228	137	131	0,00034	1593735408	14932114618	2,09302	0,03165	2,12467	1,45763	89,3%		
55	0,00005	1,01519	79	75	0,00068	2118898396	13338379210	2,04515	0,03463	2,07978	1,44214	107,1%		
60	0,00005	0,80177	119	111	0,00061	1190490363	11219480814	1,90915	0,03843	1,94758	1,39556	146,7%		
65	0,00006	0,71893	94	85	0,00113	1237396316	10028990451	1,96142	0,04417	2,00559	1,41619	184,6%		
70	0,00008	0,63772	72	60	0,00276	1507578589	8791594135	2,10398	0,05404	2,15802	1,46902	247,8%		
75	0,00011	0,55664	63	45	0,00632	1560230219	7284015546	2,50777	0,07775	2,58552	1,60795	386,2%		
80	0,00021	0,55235	28	19	0,01675	1005608011	5723785327	3,85709	0,15218	4,00927	2,00232	695,9%		
85	0,00046	0,51726	17	4	0,17924	4718177316	4718177316	6,87812	0,32921	7,20733	2,68465	1648,9%		
90	0,00764	1,75502	10	3	0	0	0	0	5,49457	5,49457	2,34405	X		

\* Джерело: обчислено автором за методологією Е. Ло та ін. [8].



Таблиця смертності (ст. 7–9 у таблиці) розрахована за традиційними формулами [2, с. 140–142]. Деякі показники таблиці смертності тут не наведено через її громіздкість. Вибіркова дисперсія ймовірності померти (ст. 10) визначена за наведеною вище формулою Ч.Л. Чанга (1). Величини у ст. 11 обчислюються за формулою [2, с. 163]:

$$l_x^2 \left( (1 - a_x) n_x + e_{x+n} \right)^2 S_{q_x}^2, \quad (3)$$

де  $e_{x+n}$  – очікувана тривалість життя при досягненні наступного вікового інтервалу.

Величини у ст. 12 є накопиченими сумами ст. 11 знизу догори (таблиця). Дисперсія очікуваної тривалості життя (ст. 13) розраховується за формулою [2, с. 163]:

$$S_{e_x}^2 = \sum_x^{\omega-1} p_x^2 \left( (1 - a_x) n_x + e_{x+n} \right)^2 S_{R_x}^2, \quad (4)$$

де  $p_x$  – імовірність дожити від початку до кінця вікового інтервалу  $x$ .

Оскільки ймовірність померти та ймовірність дожити доповнюють одна одну до одиниці, то їхні вибіркові дисперсії тотожні [2, с. 153]:

$$S_{q_x}^2 = S_{p_x}^2. \quad (5)$$

Ч.Л. Чанг постулює: оскільки ймовірність смерті у відкритому віковому інтервалі дорівнює одиниці, а ймовірність дожиття відповідно дорівнює нулю, то дисперсія цієї величини теж дорівнює нулю [2, с. 106]. Проте стохастична варіативність показників таблиці смертності впливає не з імовірності дожиття як такого, а з малих чисел померлих [8, с. 406]. Тому Е. Ло та ін., з огляду на непевність показників смертності відкритого вікового інтервалу, запропонували додавати поправочний член [8, с. 404]:

$$\sigma_{adj}^2 = S_{e_x}^2 + \frac{P_{0\omega}^2}{M_{\omega}^3 P_{\omega}}, \quad (6)$$

де  $S_{e_x}^2$  – вибіркова дисперсія очікуваної тривалості життя за методом Ч.Л. Чанга;  $P_{0\omega}$  – ймовірність дожити від народження до початку відкритого вікового інтервалу;  $P_{\omega}$  – чисельність населення у відкритому віковому інтервалі;  $M_{\omega}$  – коефіцієнт смертності у відкритому віковому інтервалі.

Таким чином, у ст. 14 (таблиця) записано ймовірність дожити від початку поточного до початку відкритого вікового інтервалу. У ст. 15 – її квадрат, поділений на добуток чисельності населення та кубу коефіцієнту смертності відкритого вікового інтервалу (6). Сума ст. 13 і 15 по рядках дає вибіркову дисперсію очікуваної тривалості життя з урахуванням внеску відкритого вікового інтервалу (ст. 16 у табл.).

Хоча чисельність населення має незначний вплив [8, с. 405], її внесок у дисперсію показників дожиття теж можна знайти [8, с. 405]:

$$\sigma_{adj+PE}^2 = \sigma_{adj}^2 + \left( \frac{P_{0\omega}}{M_{\omega}} \right)^2 \cdot \left( \frac{0,05}{2} \right)^2, \quad (7)$$

де 0,05 та 2 відповідають 95 % довірчому інтервалу оцінки чисельності населення [8, с. 434]. Приклад розрахунку наведено у ст. 17 і 18 таблиці.

Дисперсія очікуваної тривалості життя з наддисперсією й урахуванням внеску відкритого вікового інтервалу впливає з формули (4) [8, с. 406]:

$$\sigma_{adjNB}^2 = \sum_x^{\omega-1} (p_{0x})^2 ((1-a_x)n_x + e_{x+n})^2 S_{p_x}^2 \left( \frac{P_x^-}{P_x^+} \right)^2 + \frac{P_{0\omega}}{M_{\omega}^3 P_{\omega}} \left( \frac{P_{\omega}^-}{P_{\omega}^+} \right), \quad (8)$$

де  $P^+ = P^- - D_x$ , а

$$P^- = \frac{P_x}{n} + (1-a_x)D_x \quad (9)$$

Іншими словами, перший доданок формули (8) тотожний формулі дисперсії Ч.Л. Чанга (4), але зі своєю дисперсією ймовірності дожиття. Відповідно й обчислюється аналогічно, що показано в ст. 19–24.

Приклад розрахунку другого доданку формули (8) наведено в ст. 25, а саме: квадрат кожного рядка ст. 14 треба розділити на добуток кубу коефіцієнту смертності відкритого вікового інтервалу та чисельності населення в тому ж інтервалі та помножити на співвідношення останніх рядків ст. 19 та 20 (таблиця). Таким чином, дисперсія очікуваної тривалості життя з наддисперсією та врахуванням відкритого вікового інтервалу (ст. 26) буде сумою ст. 24 і 25 по рядках. Щоб отримати стандартне відхилення очікуваної тривалості життя (ст. 27) слід знайти квадратний корінь зі значень у ст. 26 (таблиця). У ст. 28 наведена відносна похибка (співвідношення дисперсій) з урахуванням відкритого вікового (ст. 27) інтервалу та без нього.

**Результати.** Як впливає з таблиці, смертність у відкритому віковому інтервалі справді може давати помітний внесок у дисперсію очікуваної тривалості життя. З формул (6) і (8), запропонованих Е. Ло та ін. [8], видно, що для цього повинна скластися ситуація з порівняно малим коефіцієнтом смертності у відкритому віковому інтервалі та порівняно малою чисельністю населення в ньому ж. Так, у 202 із 364 побудованих таблиць смертності відкритий інтервал 85 років і старше додатково вніс у дисперсію очікуваної тривалості життя при народженні від 10 до 20 % (рис. 1).

Обчислення показали, що ширина довірчих інтервалів, у яких лежить очікувана тривалість життя при народженні з 95 % імовірністю, швидко спадає зі збільшенням чисельності населення під ризиком смерті (рис. 2). Для населення менше п'яти тисяч стандартна похибка становить від 2,12 до 4,48 року, для населення від п'яти до 10 тисяч – від 1,56 до 2,71 року, від 40 до 47 тисяч не перевищує 0,85 року. Найбільша стандартна похибка очікуваної тривалості життя при народженні виявилася в сільських чоловіків міськради Нової Каховки – від 3,05 до 4,48 року. Звісно, це найменша сукупність із досліджених – менше трьох тисяч осіб.

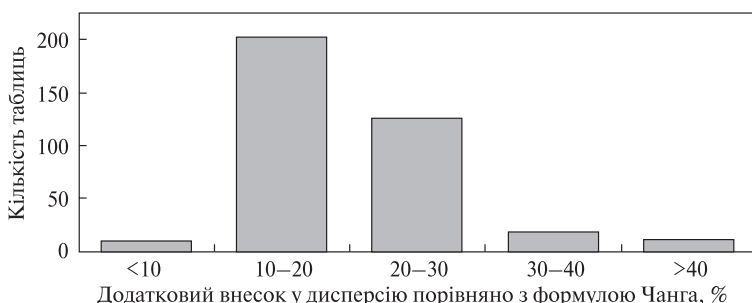


Рис. 1. Додатковий внесок відкритого інтервалу з дисперсією чисельності населення та негативною біноміальною наддисперсією у дисперсію очікуваної тривалості життя при народженні

Джерело: розраховано автором за даними Держстату України.

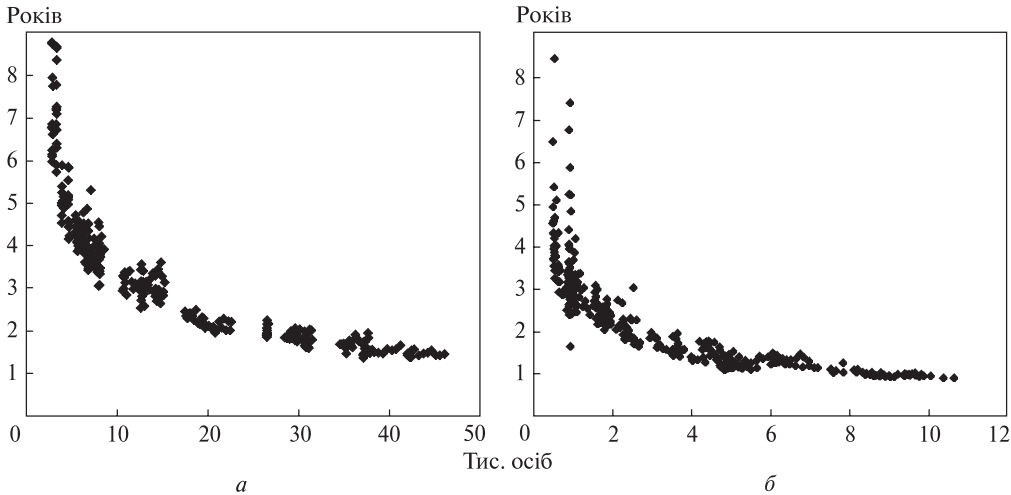


Рис. 2. Залежність 1,96 стандартного відхилення очікуваної тривалості життя при народженні від усієї чисельності населення (а) та при досягненні 60 років від чисельності населення у віці 60 років і старше (б)

Джерело: розраховано автором за даними Держстату України.

Цікаво, що за абсолютною величиною стандартна похибка очікуваної тривалості життя при досягненні 60 років близька до відповідної при народженні. Так, графік 1,96 стандартного відхилення виглядає дуже подібно (рис. 2), хоча, звісно, відносна похибка є значно більшою для віку 60 років: для очікуваної тривалості життя при народженні стандартна похибка не перевищує 7 %, а при досягненні 60 років може становити 31 %.

Цікавою проблемою є правомірність використання теоретичного населення, еквівалентного реальному. Розрахунки показують, що стандартна похибка очікуваної тривалості життя при народженні у теоретичного населення, за загальною чисельністю тотожного реальному, зазвичай нижча. Із 364 аналізованих випадків така ситуація є у 267 (рис. 3). Причому суттєва відмінність може трапитися навіть для населення понад 35 тисяч, як це сталося у сукупності жінок Нової Каховки – відмінність склала понад 10 % (рис. 3). На загал, особливої залежності від людності не спостерігається, і навіть за чисельності населення близько п'яти тисяч осіб стандартна похибка очікуваної тривалості життя для реального населення може бути близькою до його стаціонарного еквіваленту (рис. 3). Звісно, тут дається взнаки, з одного боку, ступінь близькості вікового розподілу реального і стаціонарного населення, а з другого – величина похибки, значення якої є більшим у меншій популяції.

Так, у згаданому випадку вікова структура жінок Нової Каховки (37 тис. осіб) помітно відрізняється від свого стаціонарного еквіваленту (рис. 4). Стандартна похибка склала 0,87 року проти 0,79 року (різниця у 10,2 %) в стаціонарному населенні. Якщо взяти до порівняння населення майже удесятеро менше – міських чоловіків Захарівського району (4 тис. осіб), то стандартна похибка в реальному і стаціонарному населенні виявилася тотожною до другого знаку після коми (2,58 року), а відносна різниця склала лише двадцяту частину відсотка. На рис. 4 видно, що вікова структура міських чоловіків Захарівського району є помітно ближчою до свого стаціонарного еквіваленту, ніж структура жінок Нової Каховки. Таким чином, використання теоретичної вікової структури є виправданим, якщо досліджено її подібність до реальної.

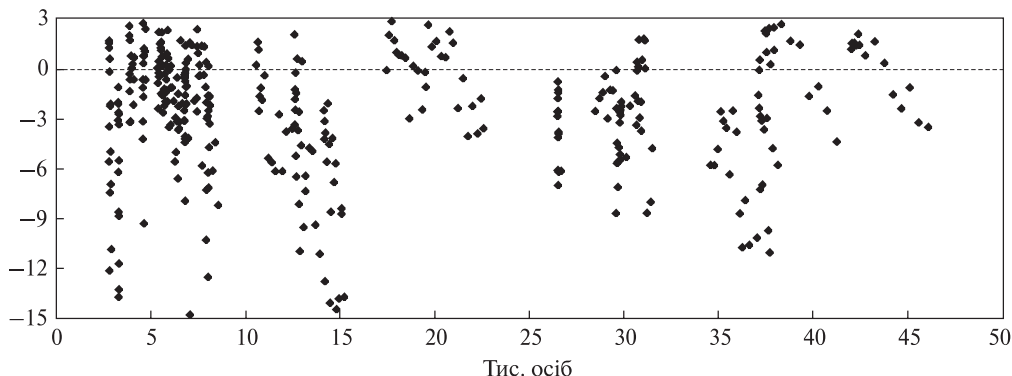


Рис. 3. Різниця стандартного відхилення очікуваної тривалості життя при народженні в стаціонарному та реальному населенні залежно від чисельності населення

Джерело: розраховано автором за даними Держстату України.

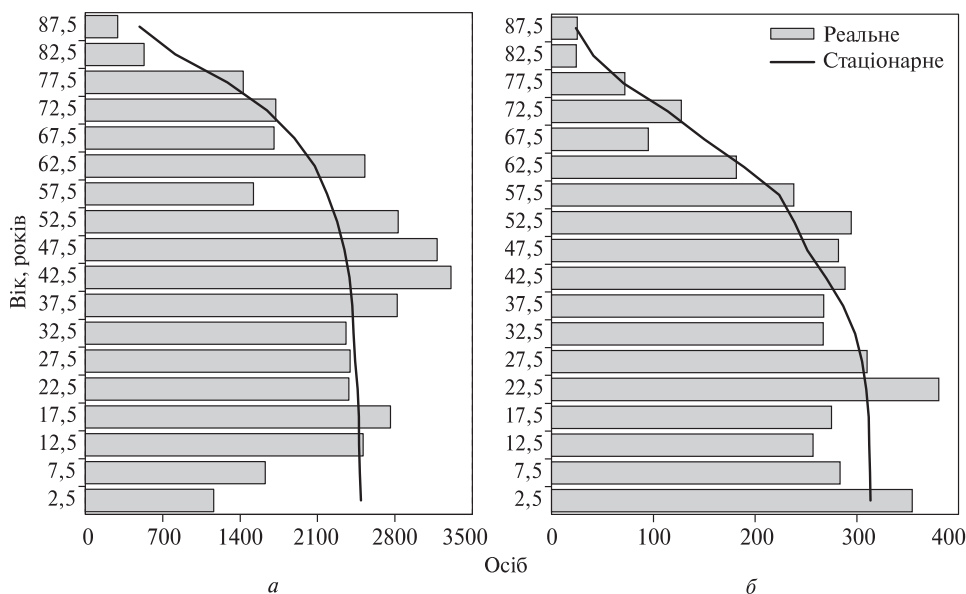


Рис. 4. Вікові діаграми реального та стаціонарного населення жінок Нової Каховки в 2002 р. (а) та міських чоловіків Захарівського району (б) в 2011 р.

Джерело: розраховано автором за даними Держстату України.

Ч.Л. Чанг розглядає число померлих як біноміальну випадкову змінну від чисельності населення [2, с. 79]. Звідси випливають можливості моделювання чисел померлих за методом Монте-Карло (генерування випадкових чисел), що було здійснено в роботах [7] та [10]. Порівняння методу, запропонованого Е. Ло та ін. з алгоритмом Є. Андрєєва та В. Школьнікова дає подібні результати. Наприклад, довірчі інтервали очікуваної тривалості життя при народженні для міських жінок Біляївського району досить близькі (рис. 5). Метод Е. Ло та ін. завжди має ширші границі довірчого інтервалу внаслідок урахування додаткових факторів невизначеності (про що йшлося вище), які не враховував Ч.Л. Чанг.

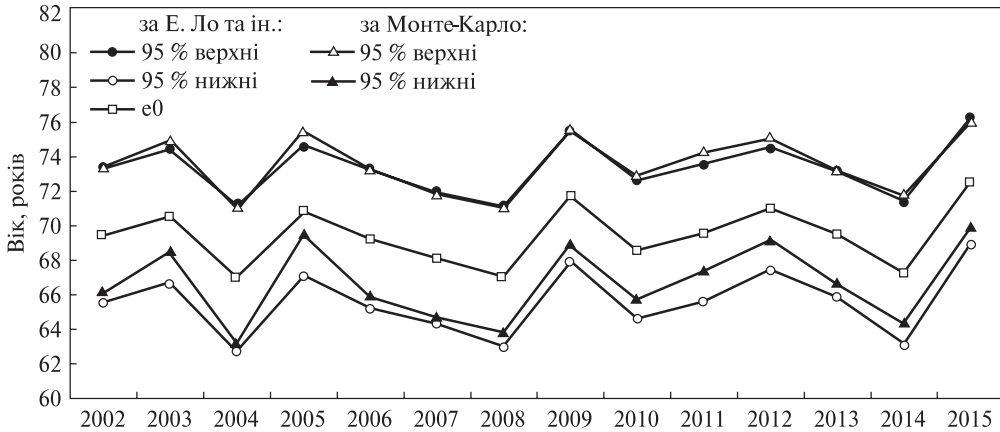


Рис. 5. Точкова оцінка очікуваної тривалості життя при народженні (e0) та границі довірчих інтервалів із 95 % імовірності за різними методами (міські жінки, Біляївський район)

Джерело: розраховано автором за даними Держстату України.

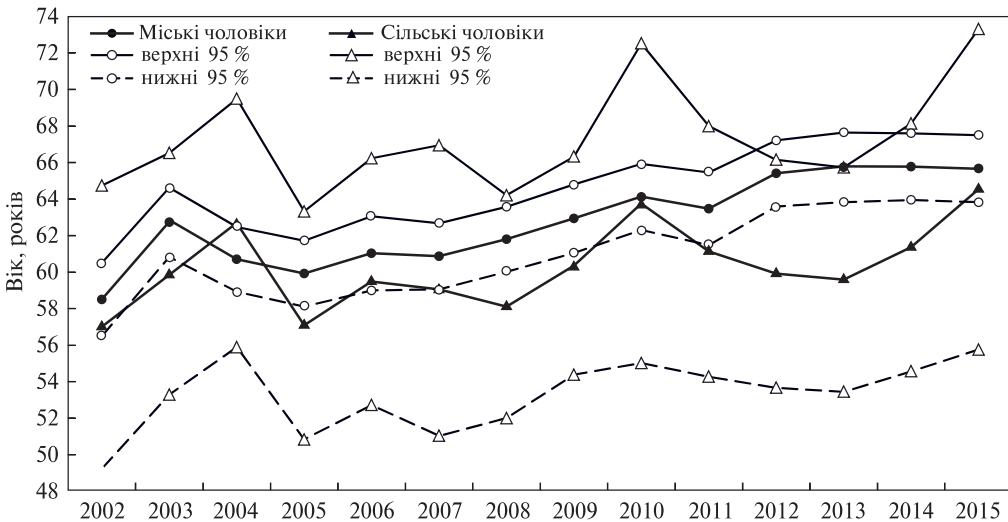


Рис. 6. Очікувана тривалість життя при народженні чоловіків міськради Нової Каховки за типом поселення з 95 % довірчим інтервалом

Джерело: розраховано автором за даними Держстату України.

Наприклад, зрозуміло, що тривалість життя міських і сільських чоловіків однієї міськради (Нової Каховки) навряд чи може сильно відрізнятися. Розрахунок відповідних таблиць смертності показує, що сільські чоловіки живуть у середньому помітно менше, за винятком 2004 та 2010 рр. (рис. 6).

Проте, враховуючи невизначеність, зумовлену малою людністю сільських чоловіків міськради та, відповідно, малим числом померлих, можна бачити, що ширина довірчих інтервалів очікуваної тривалості життя при народженні міських чоловіків майже повністю перекривається довірчим інтервалом сільських чоловіків.

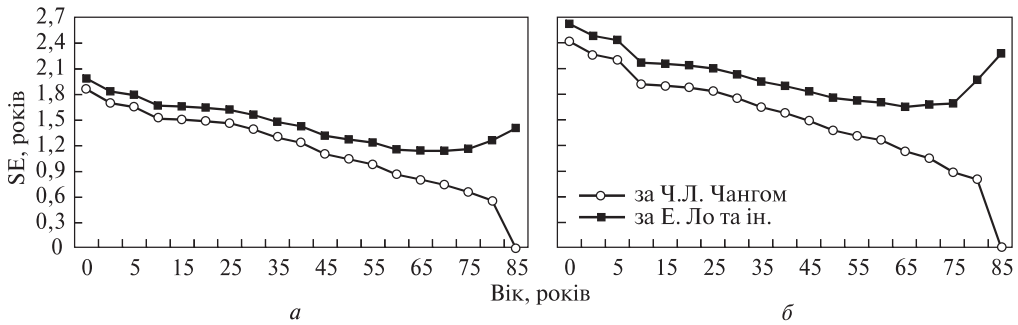


Рис. 7. Стандартна похибка очікуваної тривалості життя за віком для міських жінок Біляївського (а) та Захарівського районів (б) у 2003 р. за різними методами

Джерело: розраховано автором за даними Держстату України.

За методом Л.М. Чанга, стандартна похибка очікуваної тривалості життя з віком знижується. За методом, запропонованим Е. Ло та ін., – спершу знижується, потім у старших вікових групах зростає (рис. 7). Це зумовлено переважно врахуванням внеску відкритого вікового інтервалу.

**Висновки.** Методи, використані в цьому дослідженні, дають змогу обчислити стандартні похибки та довірчі інтервали для очікуваної тривалості життя при досягненні точного віку. Розрахунок таблиць смертності для регіонального (та субрегіонального) рівня із відповідною додатковою інформацією про довірчі інтервали, в яких знаходяться окремі показники, сприятиме глибшому розумінню динаміки процесів смертності, що, у свою чергу, уможливить коректніше порівняння показників регіонів із різною чисельністю та статеві-віковою структурою населення та ухвалення більш виважених управлінських рішень. Відповідно, розв'язується проблема, для якої чисельності населення доцільно робити подібні обчислення, оскільки користувач отримує конкретне значення ширини довірчого інтервалу, отже, може вирішити, чи задовольняє така точність його потреби.

У подальших дослідженнях доцільно у явному вигляді враховувати стохастичність демографічних процесів, зокрема у процесі складання демографічних прогнозів, що має особливе значення для малолюдних регіонів.

## ЛІТЕРАТУРА

1. Про внесення змін до деяких законодавчих актів України щодо розширення повноважень органів місцевого самоврядування та оптимізації надання адміністративних послуг: Закон України від 10.12.2015 № 888-VIII // Відомості Верховної Ради. – 2016. – № 3. – ст. 30.
2. Chiang C.L. Life Table and Its Applications. – Malabar, Florida: Robert E. Krieger Publishing Company, 1984. – 316 p.
3. Martel L., Provost M., Lebel A., Coulombe S., Sherk A. Methods for Constructing Life Tables for Canada, Provinces and Territories. – Statistics Canada, 2017. – 15 p.
4. Silcocks P.B.S., Jenner D.A., Reza R. Life expectancy as a summary of mortality in a population: statistical considerations and suitability for use by health authorities // Journal of Epidemiology & Community Health. – 2001. – 55. – P. 38–43. – doi:10.1136/jech.55.1.38
5. Toson B., Baker A., Office for National Statistics. Life expectancy at birth: Methodological options for small populations // National Statistics Methodological Series. – 2003. – № 33. – P. 1–27.
6. Eayres D., Williams E.S. Evaluation of methodologies for small area life expectancy estimation // Journal of Epidemiology & Community Health. – 2004. – 58. – P. 243–249. – doi:10.1136/jech.2003.009654

7. Scherbov S., Ediev D. Significance of life table estimates for small populations: Simulation-based study of standard errors // *Demographic Research*. – 2011. – 24. – P. 527–550. – doi: 10.4054/DemRes.2011.24.22
8. Lo E., Vatnik D., Benedetti A., Bourbeau R. Variance models of the last age interval and their impact on life expectancy at subnational scales // *Demographic Research*. – 2016. – 35, Art. 15. – P. 399–454. – doi:10.4054/DemRes.2016.35.15
9. Андреев Е., Жданов Д., Ясилонис Д. Доверительное оценивание демографических коэффициентов на примере коэффициентов смертности // *Демографическое обозрение*. – 2015. 1, № 1. – С. 24–55.
10. Andreev E., Shkolnikov V. Spreadsheet for calculation of the confidence limits for any life table or healthy-life table quantity // MPIDR technical report 2010-005. – 2010. – 4 p.

## REFERENCES

1. Zakon Ukrayiny vid 10.12.2015 № 888-VIII «Pro vnesennya zmin do deyakykh zakonodavchykh aktiv Ukrayiny shchodo rozshyrennya povnovazhen' orhaniv mistsevoho samovryaduvannya ta optymizatsiyi nadannya administratyvnykh posluh» [Law of Ukraine of 10.12.2015 № 888-VIII «On Amendments to Certain Legislative Acts of Ukraine on Extending Authorities of Local Self-Government Bodies and Optimizing the Provision of Administrative Services»]. (2016). *Vidomosti Verkhovnoyi Rady - Bulletin of the Verkhovna Rada*, 3, 30 [in Ukrainian].
2. Chiang, C.L. (1984). *Life Table and Its Applications*. Malabar, Florida: Robert E. Krieger Publishing Company.
3. Martel, L., Provost, M., Lebel, A., Coulombe, S., & Sherk, A. (2017). *Methods for Constructing Life Tables for Canada*, Provinces and Territories. Statistics Canada.
4. Silcocks, P.B.S., Jenner, D.A., & Reza, R. (2001). Life expectancy as a summary of mortality in a population: statistical considerations and suitability for use by health authorities. *Journal of Epidemiology & Community Health*, 55. doi:10.1136/jech.55.1.38
5. Toson, B., & Baker, A. (2003). Life expectancy at birth: Methodological options for small populations. Office for National Statistics. *National Statistics Methodological Series*, 33.
6. Eayres, D., & Williams, E.S. (2004). Evaluation of methodologies for small area life expectancy estimation. *Journal of Epidemiology & Community Health*, 58. doi:10.1136/jech.2003.009654
7. Scherbov, S., & Ediev, D. (2011). Significance of life table estimates for small populations: Simulation-based study of standard errors. *Demographic Research*, Vol. 24. doi: 10.4054/DemRes.2011.24.22
8. Lo, E., Vatnik, D., Benedetti, A., & Bourbeau, R. (2016). Variance models of the last age interval and their impact on life expectancy at subnational scales. *Demographic Research*, Vol. 35. doi:10.4054/DemRes.2016.35.15
9. Andreev, E., Jdanov, D., & Jasilionis, D. (2015). Doveritel'noye otsenivaniye demograficheskikh koeffitsiyentov na primere koeffitsiyentov smertnosti [Confidence estimation of demographic rates on example of mortality rates]. *Demograficheskoye obozreniye - Demographic Review*, 1. [in Russian].
10. Andreev, E., & Shkolnikov, V. (2010). *Spreadsheet for calculation of the confidence limits for any life table or healthy-life table quantity*. MPIDR technical report 2010-005.

Стаття надійшла до редакції журналу 25.06.2018.