

## ОСОБЛИВОСТІ АНАЛІЗУ БЕЗРОБІТТЯ ЗА РЕЗУЛЬТАТАМИ ВИБІРКОВИХ ОБСТЕЖЕНЬ

**Н. В. КОВТУН**, доктор економічних наук  
(Київський національний університет імені Тараса Шевченка);  
**І. С. МАРЧЕНКО**, кандидат економічних наук  
(Інститут демографії та соціальних досліджень  
імені М. В. Птухи НАН України)

***Анотація.** Розглянуто проблеми, що виникають під час використання в аналізі показників безробіття, визначених за результатами вибіркового обстеження. Показано, що за певних умов неврахування рівня точності показників може суттєво впливати на висновки, що формуються за результатами аналізу. Певну увагу приділено змісту характеристик точності оцінювання показників за результатами вибіркового обстеження.*

***Ключові слова:** показники безробіття, вибірково обстеження, точність показників.*

В Україні, як і у багатьох країнах сучасного світу, найважливіші показники, що характеризують функціонування національного та регіональних ринків праці, визначаються за результатами спеціально організованих вибіркового обстеження населення. Це насамперед обумовлено необхідністю отримання своєчасних даних щодо відповідних процесів за відносно незначних витрат коштів і часу на збір і обробку інформації.

Показники безробіття населення посідають особливе місце в системі інформаційного забезпечення соціально-економічної політики, оскільки характеризують не лише економічну активність населення та ефективність державної політики у сфері зайнятості населення, але є і важливими індикаторами потенційної соціальної напруги у суспільстві. Щоб визначити ці показники за результатами вибіркового обстеження, необхідно врахувати в разі їх використання характеристики статистичної точності (надійності) для уникнення помилкових результатів аналізу.

Проблеми точності вибіркового даних розглядаються у багатьох вітчизняних і закордонних публікаціях. Так, в Україні якість результатів вибіркового обстеження досліджують О. Гладун, О. Гончар, В. Саріогло [1–3]. Водночас особливостям використання результатів

вибіркового обстеження з урахуванням рівня їх надійності присвячено недостатньо уваги, що може суттєво погіршувати корисність відповідних даних для користувачів.

Метою статті є дослідження впливу точності визначення показників безробіття на особливості аналізу відповідних процесів. Певну увагу приділено змісту характеристик точності оцінювання показників за результатами вибіркового обстеження.

В Україні безробіття вимірюється за результатами державного вибіркового обстеження населення з питань економічної активності (далі ОЕАН) [4]. Це обстеження проводиться зі щомісячною періодичністю, на основі загальнодержавної територіально ймовірнісної вибірки неінституційних домогосподарств. Обсяг щомісячної вибіркової сукупності за період 2009–2013 рр. становив у середньому 16,6 тис. домогосподарств. Інформація збиралась по всіх членах відібраних домогосподарств віком від 15 до 70 років (включно). Результати ОЕАН поширюються на все населення України у зазначеному віці.

Слід відмітити, що річні масиви даних ОЕАН містять інформацію приблизно по 1 % осіб віком 15–70 років. Водночас частина сукупності обстежених осіб представлена у цих масивах двічі та навіть тричі відповідно до

схеми заміщення домогосподарств у вибірці [4].

Важливо зауважити, що зі статистичного погляду оцінка певного показника за результатами вибіркового обстеження характеризується високою точністю, якщо оцінки, побудовані на основі значної кількості можливих вибірок однакового обсягу та сформованих з використанням однакового механізму сконцентровані біля дійсного значення показника. Висока точність забезпечує високу імовірність того, що за результатами реалізації процедури формування вибірки лише один раз (це завжди відчувається під час організації реальних вибіркового обстежень) оцінки показників будуть близькими до дійсних їх значень, зазвичай, невідомих.

Характеристикою точності оцінювання показників за результатами вибіркового обстеження є похибка вибірки, взаємопов'язаними характеристиками якої є стандартна похибка, гранична похибка, коефіцієнт варіації (відносна стандартна похибка), довірчий інтервал та ін. Власне стандартна похибка вибірки і роз-

раховується як корінь квадратний з дисперсії оцінок показників за різними вибірками. Слід зазначити, що для імовірнісних вибірок дисперсія вибіркового оцінок може бути розрахована на основі дисперсії ознаки, за якою визначається показник, по одиницях вибірки при одноразовій її реалізації. Гранична похибка визначає межі довірчого інтервалу для оцінки показника і є максимально можливою похибкою вибірки за певної довірчої імовірності. Довірчу імовірність встановлюють, зазвичай, на рівні 95 %, тобто передбачається, що для кожних 95 зі 100 вибірок цього плану та обсягу оцінка показника потрапляє до довірчого інтервалу. Зазначений рівень довірчої імовірності гранична похибка становить приблизно дві стандартні похибки [5]. Коефіцієнт варіації – це відношення величини стандартної похибки до оцінки показника. Коефіцієнт варіації є відносною характеристикою похибки вибірки. Оцінки рівня безробіття за регіонами України у 2012 р. і характеристики їх точності представлені у табл. 1.

Таблиця 1

**Оцінки надійності показника рівня безробіття населення у віці 15–70 років за регіонами проживання у 2012 р. [4]**

Адміністративно-територіальна одиниця, область, місто	Рівень безробіття населення у віці 15–70 років, %	Стандартна похибка вибірки ( $\mu$ ), %	Гранична похибка вибірки ( $\Delta$ ), %	Коефіцієнт варіації ( $V_{\mu}$ ), %
Україна	7,5	0,2	0,4	2,8
Автономна Республіка Крим	5,8	0,6	1,2	10,9
Вінницька	8,8	1,2	2,4	14,0
Волинська	8,1	1,3	2,6	16,3
Дніпропетровська	6,6	0,7	1,4	10,9
Донецька	8,0	0,7	1,3	8,5
Житомирська	9,7	1,2	2,3	12,0
Закарпатська	8,7	0,6	1,1	6,6
Запорізька	7,0	0,6	1,2	9,1
Івано-Франківська	7,9	1,2	2,3	14,8
Київська	6,3	0,8	1,6	13,3
Кіровоградська	8,4	0,7	1,3	8,2
Луганська	6,4	1,1	2,1	17,0
Львівська	7,5	0,7	1,4	9,7
Миколаївська	7,9	1,1	2,2	14,4
Одеська	5,8	1,1	2,2	19,5
Полтавська	8,6	1,5	3,0	17,7
Рівненська	9,8	1,9	3,8	19,9
Сумська	8,6	1,8	3,5	20,8

Продовж. табл. 1

Адміністративно-територіальна одиниця, область, місто	Рівень безробіття населення у віці 15–70 років, %	Стандартна похибка вибірки ( $\mu$ ), %	Гранична похибка вибірки, ( $\Delta$ ), %	Коефіцієнт варіації ( $V_\mu$ ), %
Тернопільська	9,8	1,6	3,1	16,2
Харківська	6,8	1,1	2,2	16,3
Херсонська	8,7	0,6	1,2	7,2
Хмельницька	8,6	1,1	2,2	13,1
Черкаська	9,0	1,4	2,7	15,2
Чернівецька	8,0	1,8	4,8	22,6
Чернігівська	9,8	0,8	1,5	7,9
м. Київ	5,5	0,6	1,2	10,8
м. Севастополь	5,9	1,1	2,1	17,9

Для ілюстрації впливу похибки вибірки на особливості аналізу безробіття розглянемо приклад. На рис. 1 представлено зміну рівнів безробіття та довірчих інтервалів, визначених за результатами ОЕАН для Тернопільської об-

ласті. З наведених даних видно, що межі довірчих інтервалів становлять у середньому 2,7 %. Це означає, що з імовірністю 95 % кожного року рівень безробіття міг би бути дещо більшим або меншим за наведений.

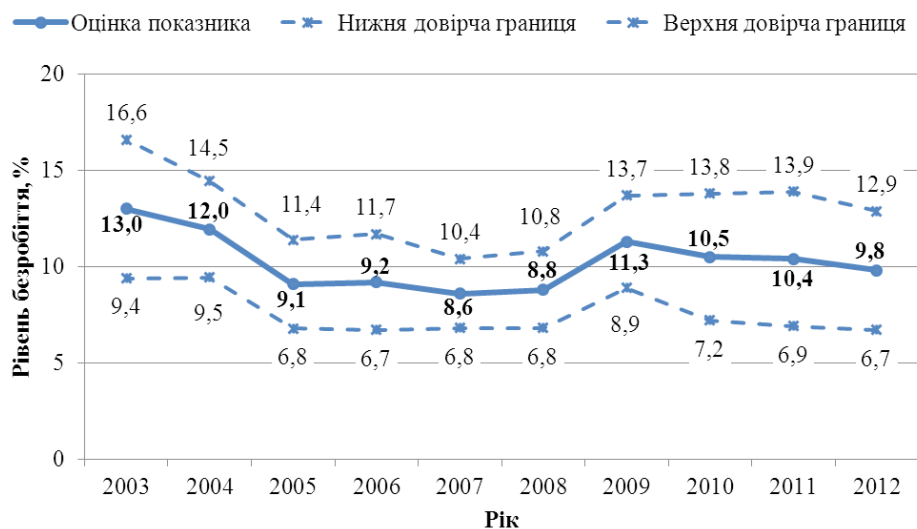


Рис. 1. Динаміка зміни оцінок рівня безробіття та їх довірчих меж для Тернопільської області (складено авторами за даними Державної служби статистики України)

Для ілюстрації та пояснення цього ефекту проведемо розрахунковий експеримент. Згенеруємо умовну генеральну сукупність осіб віком 15–70 років. Це можна зробити в EXCEL або у будь-якому статистичному пакеті. Нехай обсяг генеральної сукупності становить 35000 осіб. Випадково розподілимо у цій сукупності ознаку безробітного так, щоб частка безробітних складала 10 %, приблизно як у Тернопільській області у 2011, 2012 рр. Тобто 3500 осіб з цієї сукупності, обрані випадковим чином із рівними ймовірностями, матимуть

ознаку безробітного. Далі, використовуючи процедуру простого випадкового відбору, відберемо 1 % генеральної сукупності, що становить 350 осіб, і оцінимо частку безробітних у вибірці. Для такої вибірки дисперсія вибіркових оцінок визначається за формулою:

$$\delta = \frac{P \cdot (100 - P)}{n} = \frac{10 \cdot (100 - 10)}{350} \approx 2,57, \quad (1)$$

де  $P$  – рівень безробіття (у цьому випадку частка безробітних осіб);  
 $n$  – обсяг вибірки.

Якщо при розрахунках дисперсії вибірко-вих оцінок за формулою (1) дійсне значення показника невідомо, то використовують його оцінку за вибіркою. При цьому, відповідно, і дисперсія також оцінюється з певною похибкою.

Для прикладу, що розглядається, стандартна похибка становить  $\mu = \sqrt{\delta} \approx 1,60\%$ , гранична похибка – 3,2 %, а коефіцієнт варіації – 16,0 %. При цьому слід враховувати, що стандартна та гранична похибки вимірюються у тих же одиницях, що і показник, а коефіцієнт

варіації показує, скільки відсотків оцінки показника становить стандартна похибка.

Якщо сформувати таку вибірку незалежно 10 разів, то результат, що буде отримано, може мати вигляд, як на рис. 2. З наведених даних видно, що у першу вибірку не потрапило жодної безробітної особи. Цей випадок належить до тих 5 % імовірності отримати оцінку показника, що відхиляються від дійсного його значення на величину більшу, ніж дві стандартні похибки. Всі інші оцінки потрапили у довірчий інтервал, як це і прогнозувалося.

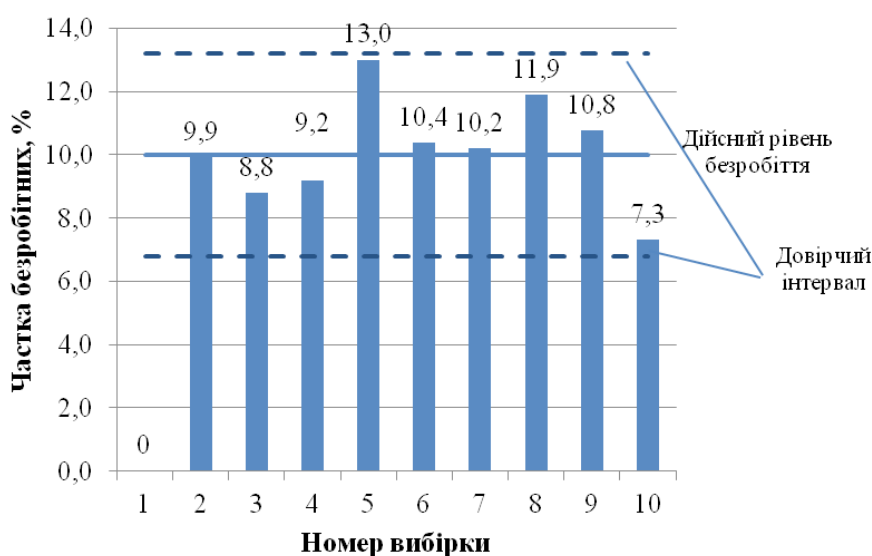


Рис. 2. Розподіл оцінок частки безробітних осіб за незалежними вибірками з модельної генеральної сукупності (складено авторами за власними розрахункам)

Розподіл оцінок рівня безробіття має характер випадкових коливань навколо дійсного значення. При цьому амплітуда коливань у 95 % випадків обмежується довірчим інтервалом. Загалом, аналогічні коливання за порівнюваними величинами похибок вибірки характерні і для результатів ОЕАН. Що це означає з погляду можливості використання даних для аналізу відповідних процесів?

По-перше, завжди слід враховувати, що рівень безробіття, визначений за результатами ОЕАН, наприклад, для Тернопільської області у 2012 р., становить не точно 9,8 %, а з імовірністю 95 % може дорівнювати будь-якому значенню з інтервалу  $9,8 \pm 3,1\%$ . При цьому оцінки показника в межах довірчого інтервалу розподілені за нормальним законом і довірчий

інтервал будується не навколо невідомого дійсного значення показника, а навколо його оцінки [6].

По-друге, при аналізі динаміки процесів зміна оцінки показника на величину меншу ніж гранична похибка фактично може не враховуватись, оскільки висновок про зростання або зниження рівня безробіття буде статистично необґрунтованим. Наприклад, для Тернопільської області за період 2009–2012 рр. усі зміни оцінок рівня безробіття не перевищують граничних похибок, тому не мають братися до уваги, особливо у кількісному аналізі. Все це стосується і порівнянь рівня безробіття між регіонами.

Доцільно зауважити, якщо в аналізі необхідно оцінювати і враховувати зміни оцінок

рівня безробіття, то точність визначення таких змін залежить від точності оцінок самих показників і може бути гіршою за неї. Зокрема, за умов незалежності вибірок за два суміжні роки  $t$  і  $t - 1$ , стандартна похибка оцінювання різниці оцінок рівнів безробіття визначається за формулою (2):

$$\mu(\hat{P}_t - \hat{P}_{t-1}) = \sqrt{(\mu(\hat{P}_t))^2 + (\mu(\hat{P}_{t-1}))^2}. \quad (2)$$

Якщо оцінки показника за два суміжні роки характеризуються однаковою стандартною похибкою, що становить, наприклад, 1,6 %, то стандартна похибка оцінки зміни показника дорівнюватиме  $\sqrt{(1,6)^2 + (1,6)^2} = 1,6 \cdot \sqrt{2} \approx 2,3$  %.

Відповідно, зміна оцінки показника, що може становити біля 3 в. п., визначатиметься з точністю  $\pm 4,6$  в. п.

Наведене свідчить, що використання даних вибірових обстежень при аналізі безробіття потребує певної підготовки та відповідних знань. Враховуючи те, що вибірові дані можуть характеризуватися значною похибкою, на практиці доцільно застосовувати відомий метод триангуляції. Сутність якого, в аспекті, що розглядається, полягає у використанні при аналізі даних з різних джерел. Наприклад, у ході аналізу рівня безробіття для певному регіоні за умов значної похибки оцінювання необхідно враховувати не лише поточну оцінку показника за результатами ОЕАН, а й дані за попередній період, дані по суміжних регіонах, дані про рівень зареєстрованого безробіття тощо. Власне, досвідчені користувачі саме так і роблять. Крім того, існують статистичні процедури, що дають можливість ефективно оцінювати показники на основі інформації з різних джерел [7].

Отже, за результатами виконаного дослідження встановлено, що в ході використання показників безробіття, розрахованих за результатами вибірових обстежень, необхідно враховувати точність їх оцінювання. За певних умов недостатня точність показників може суттєво впливати на висновки та рекомендації, що формулюються за результатами аналізу.

Для забезпечення потреб користувачів органів державної статистики мають використовувати всі можливості для підвищення точності оцінювання показників безробіття або їх змін за результатами вибірових обстежень. При цьому, вимоги до точності оцінювання повинні враховувати й очікувані рівні показників і їх змін, і основні напрями використання показників у аналізі. Враховуючи імовірнісний характер оцінок показників і, зокрема, рівня безробіття за результатами вибірових обстежень, під час аналізу доцільно максимально використовувати наявну додаткову інформацію – дані попередніх періодів, дані з інших джерел тощо.

## ЛІТЕРАТУРА

1. Гладун О. М. Проблеми отримання надійних даних для регіонального рівня при проведенні вибірових обстежень домогосподарств / О. М. Гладун // Проблеми статистики : щорічник наук. пр. – К. : НДІ статистики Держкомстату України, 2001. – Вип. 3. – С. 154–158.
2. Гончар О. В. Забезпечення якості статистичної інформації: методологія та організація : монографія / О. В. Гончар. – К. : ВПД «Формат», 2011. – 372 с.
3. Саріогло В. Г. Методологічні підходи до забезпечення надійності результатів вибірових обстежень населення / В. Г. Саріогло // Науковий вісник ДАСОА : зб. наук. пр. – К. : [б. в.], 2004. – Вип. 5. – С. 24–29.
4. Економічна активність населення України 2012 : стат. зб. / відп. за вип. І. В. Сенік. – К. : Держ. служба стат. України, 2013. – 208 с.
5. Єріна А. М. Організація вибірових обстежень : [навч. посіб.] / А. М. Єріна. – К. : КНЕУ, 2004. – 127 с.
6. Кокрен У. Методы выборочного исследования / У. Кокрен. – М. : Статистика, 1976. – 440 с.
7. Саріогло В. Г. Оцінювання соціально-економічних показників: прикладні аспекти застосування непрямих методів : монографія / В. Г. Саріогло. – К. : Інст-т демогр. та соц. досліджень імені М. В. Птухи НАН України, 2023. – 139 с.

**Н. В. Ковтун**, доктор економічних наук (Київський національний університет імені Тараса Шевченка); **І. С. Марченко**, кандидат економічних наук (Інститут демографії та соціальних досліджень імені Н. В. Птухи НАН України). **Особенности анализа безработицы по результатам выборочных обследований.**

**Аннотація.** Розглянуті проблеми, що виникають при використанні в аналізі показників безработиці, визначених по результатам виборочних досліджень. Показано, що при певних умовах пренебреження точністю показників може суттєво впливати на висновки, сформульовані по результатам аналізу. Визначено увагу на вміст характеристик точності оцінювання показників по результатам виборочних досліджень.

**Ключевые слова:** показники безработиці, виборочні дослідження, точність показників.

**N. Kovtun**, Dc. Econ. Sci., Professor (Kiev National University Taras Shevchenko); **I. Marchenko**, Cand. Econ. Sci., associate professor (Ptoukha Institute for Demography and Social Studies of the National Academy of Sciences of Ukraine). **Features of unemployment analysis by the sample surveys results.**

**Summary.** The paper examines the problems with the unemployment indicators analyzing by the sample surveys results. The accuracy of estimate could be affecting the conclusions by the analysis results. Some attention is paid to the contents of accuracy characteristics of indicators estimation by the sample surveys results.

**Keywords:** unemployment rate, sample surveys, precision performance.