

ДЕРЖАВНИЙ КОМІТЕТ СТАТИСТИКИ УКРАЇНИ

ЗАТВЕРДЖЕНО
Наказ Державного комітету
статистики України
31.05.2010 р. № 202

МЕТОДИКА
ФОРМУВАННЯ ВИБІРКИ ДЛЯ ПРОВЕДЕННЯ ПОТОЧНИХ ОБСТЕЖЕНЬ
ПІДПРИЄМСТВ ІЗ ПИТАНЬ СТАТИСТИКИ ПРАЦІ

Київ – 2010

АНОТАЦІЯ

У методиці викладено теоретичні засади розробки дизайну та формування вибірки для проведення поточних обстежень підприємств із питань статистики праці. Розглянуто джерела інформації для визначення сукупності звітних одиниць обстежень, підходи до формування основи вибірки, оцінки обсягу вибірки, стратифікації генеральної сукупності, реалізації схеми ротації одиниць у вибірці, механізм відбору одиниць тощо. При розробці методики враховано досвід органів статистики інших країн щодо формування вибірок для проведення обстежень підприємств.

Процедури формування сукупності одиниць для обстеження розроблено з урахуванням вимог використання сучасних комп'ютерних технологій для обробки статистичної інформації, зокрема на основі пакета статистичних програм SPSS та його модуля "Complex Samples".

Значну увагу приділено питанням інформаційного забезпечення розрахунків та наведено конкретні приклади.

Методику призначено для використання у практичній діяльності працівниками органів державної статистики, насамперед на державному рівні. Вона може бути корисною також для користувачів інформації – фахівців наукових організацій і навчальних закладів, які проводять дослідження в галузі методології вибірових обстежень, аспірантів та студентів.

Методику розроблено фахівцями Інституту демографії та соціальних досліджень НАН України В.Г. Саріогло, Г.І. Терещенко за безпосередньої участі директора департаменту статистики праці Держкомстату Сенік І.В.

ЗМІСТ

	с.
Передмова	4
Визначення термінів	5
1. Загальні положення	7
2. Визначення сукупності звітних одиниць для обстеження	9
3. Дизайн вибірки	11
3.1. Стратифікація генеральної сукупності	11
3.2. Схема ротації одиниць вибірки	11
3.3. Визначення обсягу вибіркової сукупності	12
4. Відбір звітних одиниць	18
4.1. Відбір звітних одиниць за стратами	18
4.2. Відбір звітних одиниць за сукупністю „малих” страт і за стратами, в яких обсяг вибірки дорівнює обсягу генеральної сукупності	20
4.3. Коригування обсягу вибірки для забезпечення оцінювання показників за адміністративно-територіальними одиницями	20
5. Контроль якості формування вибірки	21
6. Актуалізація основи вибірки	22
Додатки	23
Перелік джерел	33

ПЕРЕДМОВА

Пріоритетним напрямом реалізації Стратегії розвитку державної статистики на період до 2012 року, схваленої розпорядженням Кабінету Міністрів України від 5 листопада 2008р. № 1413-р, визначено вдосконалення організації державних статистичних спостережень, подальший перехід до вибірових методів спостереження та підвищення якості статистичних даних.

Одним з основних етапів організації вибірового обстеження, який суттєво впливає на якість результатів, є розробка дизайну (плану) вибірки – визначення обсягу, структури та механізмів формування вибірової сукупності. Важливість цього етапу зростає для вибірових обстежень, які проводяться на постійній основі – щомісячно, щоквартально або щорічно, оскільки для таких обстежень необхідно забезпечити раціональне навантаження на респондентів, прийнятні для користувачів терміни оприлюднення даних тощо. Це призводить до необхідності задоволення низки додаткових вимог до дизайну вибірки.

Ця методика, розроблена у рамках виконання заходів щодо реалізації у 2009р. Стратегії розвитку державної статистики на період до 2012 року, визначає порядок формування сукупності звітних одиниць для проведення поточних обстежень із питань статистики праці. Для організації цих обстежень застосовується комбінований підхід, за яким частина сукупності звітних одиниць обстежується на суцільній основі, інша частина – з використанням вибірового методу.

Методика складається з шести розділів. У першому розділі наводяться основні методологічні положення щодо організації поточного обстеження підприємств із питань статистики праці: мета; одиниці спостереження та генеральна сукупність обстеження; розрізи, за якими передбачається оцінювати показники. Другий розділ містить інформацію щодо джерел даних, необхідних для визначення генеральної сукупності звітних одиниць. Розглядаються особливості формування сукупності звітних одиниць у частині суцільного обстеження та вибірової сукупності звітних одиниць. У третьому розділі висвітлюються основні питання розробки дизайну вибірки – стратифікації генеральної сукупності, побудови схеми ротації одиниць вибірки, визначення обсягу вибірової сукупності. Четвертий розділ присвячено особливостям реалізації механізмів відбору звітних одиниць за стратами. У п'ятому розділі наводяться підходи до контролю якості формування вибірки. У шостому розділі розглядається порядок актуалізації основи вибірки.

У додатках наведено технологічну схему формування сукупності звітних одиниць для обстеження та приклади реалізації окремих процедур і результати формування вибірових сукупностей.

Розроблені процедури реалізації методики передбачають автоматизацію на основі пакета статистичних програм SPSS.

ВИЗНАЧЕННЯ ТЕРМІНІВ

Відбір систематичний – відбір для вибіркового спостереження, суть якого полягає в тому, що складається список одиниць генеральної сукупності і залежно від кількості одиниць (серій), що відбирають, установлюють крок відбору, тобто через який інтервал слід брати для спостереження одиниці (серії). Початок відбору визначають або з використанням генератора випадкових чисел із номерів одиниць спостереження, які відповідають першому інтервалу, або способом жеребкування, або через розрахунок, наприклад як частку від ділення кроку відбору на два.

Відбір імовірнісний – відбір, оснований на об'єктивних правилах випадкового відбору, під час якого кожна одиниця сукупності має відомий ненульовий шанс бути відбраною до вибірки.

Вибіркова сукупність (вибірка) – сукупність одиниць, відібраних за певними правилами із генеральної сукупності для статистичного спостереження. Вибірка повинна бути представницькою (репрезентативною), тобто повно й адекватно відбивати властивості та пропорції генеральної сукупності.

Вибірковий метод – система правил відбору одиниць і способів характеристики сукупності досліджуваних одиниць, що вивчають. Вибірковий метод дає змогу розповсюдити висновки, отримані на основі вивчення частини сукупності (вибірки), на всю сукупність (генеральну).

Генеральна сукупність – множина одиниць, характеристики яких необхідно вивчати чи досліджувати і щодо яких буде зроблене узагальнення, ґрунтуючись на вибірці. Для вибіркового спостереження генеральною сукупністю є вся сукупність наявних об'єктів, із яких тим чи іншим чином формують вибірку сукупність.

Дисперсія вибіркових оцінок – середній квадрат відхилень оцінок показника за всіма можливими вибірками певного дизайну від середнього значення оцінки за всіма вибірками.

Звітна (облікова) одиниця (респондент статистичного спостереження) – одиниця, від якої отримують інформацію в ході статистичного спостереження.

Кластер – група одиниць спостереження, об'єднаних за певною ознакою.

Коефіцієнт варіації (відносна стандартна похибка) – відношення середньоквадратичного відхилення вибіркових оцінок показника за всіма можливими вибірками певного дизайну до середнього значення оцінки показника за всіма вибірками.

Надійність – ступінь близькості розрахованих оцінок показників до їх дійсних (точних) значень.

Обсяг вибірки – кількість одиниць у вибірковій сукупності. Необхідний для забезпечення репрезентативності та належної якості результатів вибіркового обстеження обсяг вибірки розраховують залежно від способу

відбору та типу вибірки. Величина обсягу вибірки залежить від варіації досліджуваних ознак та припустимої похибки вибірки.

Одиниця статистичного спостереження – первинний елемент об'єкта статистичного спостереження, який є носієм ознак, що підлягають реєстрації у ході статистичного спостереження.

Основа вибірки – упорядкований перелік, список одиниць відбору. При використанні багатоступеневого відбору на кожному ступені формується своя основа вибірки.

План (дизайн) вибірки – організаційно-логічна модель структури вибіркової сукупності та принципи її формування. Якість плану вибірки підтверджує репрезентативність отриманої за визначеними у плані правилами вибіркової сукупності.

Регіони – Автономна Республіка Крим, 24 області, міста Київ та Севастополь.

Ротація одиниць вибірки – заміна частини діючої вибіркової сукупності (або для окремих вибіркових обстежень – її повна заміна) за розробленими схемами при реалізації чергових раундів вибіркового обстеження, яке проводиться на постійній основі.

Стратифікація – поділ генеральної сукупності (або її частини) на групи, однорідні за певною ознакою.

Щодо визначень наведених термінів див. також [1].

Перелік скорочень

ЄДРПОУ	Єдиний державний реєстр підприємств та організацій України
ПА	Ітераційна пропорційна апроксимація
КВЕД	Класифікація видів економічної діяльності
КОПФГ	Класифікація організаційно-правових форм господарювання
КОАТУУ	Класифікатор об'єктів адміністративно-територіального устрою України
КОДУ	Класифікація органів державного управління
ОПСП	Поточне обстеження підприємств із питань статистики праці
СРП	Статистичний реєстр підприємств

1. ЗАГАЛЬНІ ПОЛОЖЕННЯ

Головною метою поточних обстежень підприємств із питань статистики праці (далі – ОПСП) є отримання надійної інформації щодо оплати праці та таких аспектів соціально-трудових відносин, як динаміка кількості зайнятих працівників та рух кадрів, обсяги і структура фонду оплати праці та робочого часу, розподіл працівників за розмірами заробітної плати та стан її виплати, стан колективно-договірного регулювання тощо. Зазначена інформація відповідає потребам основних користувачів і використовується для теоретичних узагальнень і аналізу відповідних явищ і процесів. За результатами спостережень розраховуються абсолютні та відносні (в розрахунку на одного працівника) показники, а також відповідні показники динаміки.

За часом реєстрації даних ОПСП складається зі щомісячних та щоквартальних спостережень. Щомісячні спостереження відрізняються від щоквартальних більш вузьким колом питань, що вивчаються, та коротшими термінами обробки й оприлюднення інформації. Окремі блоки питань вивчаються один або два рази на рік.

Одиницею спостереження в ОПСП є місцева одиниця (підприємство або його частина), розташована у географічно визначеному місці та залежна тільки від одного підприємства.

В окремих випадках одиницею спостереження також може виступати місцева одиниця за видом економічної діяльності, що об'єднує однорідну діяльність одиниці на рівні групи КВЕД.

Специфіка визначення одиниці статистичного спостереження для ОПСП пов'язана з тим, що обстеженням охоплюються підприємства, установи, організації усіх видів діяльності (за винятком секцій Р „Діяльність домашніх господарств” та Q „Діяльність екстериторіальних організацій”) незалежно від організаційно-правових форм господарювання. Виділення місцевих одиниць за видом економічної діяльності або підприємств, що звітують за інші одиниці, притаманно бюджетним установам. Наприклад, сільські (селищні) ради мають надавати звіти за одиниці кількох видів діяльності: апарат сільської ради (код за КВЕД – 75.11.5), фельдшерсько-акушерський пункт (85.11.1), дитячий садок (80.10.1), будинок культури (92.32.0). Водночас така бюджетна установа, як районний відділ освіти в особі централізованої бухгалтерії, подає звіт за групи установ (юридичних осіб), які самостійно не ведуть облік з питань праці та нарахувань заробітної плати, а саме: шкіл, дошкільних закладів тощо (порядок отримання звітів від бюджетних установ у реальних умовах залежить від процесів формування й розподілу місцевих бюджетів та може змінюватись із часом).

Звітною одиницею або респондентом, відповідно до положень статті 4 Закону України «Про державну статистику», в ОПСП є юридичні особи та відокремлені підрозділи юридичних осіб, які знаходяться на території України (далі – звітні одиниці).

Генеральна сукупність ОПСП формується виключно зі звітних одиниць, що ідентифіковані в ЄДРПОУ.

Відповідно до міжнародної практики ОПСП організовано на основі комбінованого підходу, за яким великі підприємства обстежуються на суцільній основі, а малі – з використанням вибіркового методу. В Україні це обстеження впроваджується як суцільне для сукупності звітних одиниць із кількістю працівників¹ 50 і більше осіб та як вибіркоче для сукупності одиниць із кількістю працівників від 10 до 49 осіб включно.

Таким чином, генеральну сукупність обстеження складають усі юридичні особи та їхні відокремлені підрозділи з середньою кількістю найманих працівників 10 і більше осіб, що ідентифіковані в ЄДРПОУ.

Комбінований підхід забезпечує надійне оцінювання основних показників ОПСП за генеральною сукупністю на національному та регіональному рівнях за секціями та підсекціями КВЕД тощо. Водночас такий підхід сприяє уникненню зайвого навантаження статистичною звітністю тих респондентів, які представляють одиниці генеральної сукупності з кількістю працівників менше 50 осіб.

Сукупність звітних одиниць, що обстежуються на суцільній основі, повинна надавати відомості протягом усього терміну проведення ОПСП. Сукупність звітних одиниць, що обстежуються на основі вибіркового методу, формується з використанням стратифікованого випадково відбору². При формуванні вибірки із її основи з використанням процедури систематичного відбору відбирається встановлена кількість звітних одиниць, які представляють усі регіони України та види економічної діяльності й обстежуються протягом двох років. Обсяг вибірки для щомісячного обстеження становить близько 10 тис. звітних одиниць, або біля 9% генеральної сукупності (розподіл обсягу вибірки за регіонами наведений у додатку А). За розробленою схемою ротації щороку оновлюється 50% вибіркової сукупності звітних одиниць. Для кожної одиниці з генеральної сукупності контролюється ймовірність її включення до вибірки, що забезпечує можливість адекватного поширення результатів обстеження на генеральну сукупність шляхом статистичного зважування.

Основні етапи формування сукупності звітних одиниць для ОПСП наведено на рис. Б.

Дизайн обстеження розроблено для забезпечення надійного оцінювання основних показників за результатами ОПСП за такими кластерами:

- Україна в цілому;
- Україна за видами економічної діяльності (секція та підсекція за КВЕД);
- регіони в цілому;

¹ Для процедур відбору одиниць в якості показника кількості працівників застосовується показник «середня кількість найманих працівників»

² Дизайн вибірки розроблено в рамках проекту TACIS EuropeAid/122761/C/SER/UA

- регіони за видами економічної діяльності (секція та підсекція за КВЕД);
- адміністративно-територіальні одиниці (міста республіканського та обласного значення, райони) в цілому (окремі показники).

Передбачається, що за результатами ОПСП рівень надійності оцінювання основних показників буде відповідати таким критеріям [2]: на національному рівні коефіцієнт варіації оцінок показників не перевищуватиме 5%; за сукупністю звітних одиниць у кожному з регіонів та за кожним із видів економічної діяльності – 10%; за сукупністю звітних одиниць за видами економічної діяльності в межах регіонів та по адміністративно-територіальних одиницях – 15%. Зазначені критерії надійності статистичної інформації відповідають рекомендаціям експертів та міжнародній практиці.

2. ВИЗНАЧЕННЯ СУКУПНОСТІ ЗВІТНИХ ОДИНИЦЬ ДЛЯ ОБСТЕЖЕННЯ

Генеральна сукупність одиниць в ОПСП визначається на основі актуальних даних із таких джерел:

- Реєстр статистичних одиниць у частині Статистичного реєстру підприємств (СРП);
- перелік і характеристики підприємств, установ, організацій та їхніх відокремлених підрозділів, що звітували за формою № 1-ПВ (місячна) “Звіт з праці”;
- перелік і характеристики підприємств, установ, організацій та їхніх відокремлених підрозділів за адміністративними даними Пенсійного фонду України.

Використання зазначених джерел даних забезпечує мінімізацію похибок включення (виключення) одиниць при організації ОПСП.

Роботи з визначення генеральної сукупності виконуються один раз на рік після актуалізації даних СРП відповідно до плану державних статистичних спостережень.

Ураховуючи, що СРП містить інформацію за значним колом одиниць, які самостійно не ведуть облік із питань праці (наприклад, школи, за які звітує районний відділ освіти) та, відповідно, не можуть виступати респондентами в ОПСП, необхідною є реалізація спеціальних процедур обробки даних (редагування даних) при визначенні основи вибірки.

За результатами виконання зазначених робіт формується перелік звітних одиниць з кількістю найманих працівників 10 і більше осіб. Для кожної одиниці із переліку наводяться характеристики, необхідні для адекватного визначення сукупностей одиниць для обстеження в ОПСП (табл. 1).

Таблиця 1

**Інформація, необхідна для формування сукупності звітних одиниць
для ОПСП**

№	Характеристика	Примітка
1	Ідентифікаційний код ЄДРПОУ	
2	Найменування	
3	Код території за КОАТУУ за місцем реєстрації	
4	Код території за КОАТУУ за місцем здійснення діяльності	на цей час отримується за результатами обробки форми № 1-ПВ (місячна), в подальшому з СРП
5	Код організаційно-правової форми господарювання за КОПФГ	
6	Код основного виду економічної діяльності за КВЕД	
7	Код органу державного управління за КОДУ	використовується для реалізації процедур редагування даних при формуванні основи вибірки
8	Правовий статус	юридична особа, без права юридичної особи
9	Ідентифікаційний код юридичної особи	для відокремлених підрозділів
10	Ознака активності за даними Державної податкової адміністрації України у звітному році	
11	Середня кількість найманих працівників	за одним із джерел відповідно до пріоритетності занесення даних про кількість працівників
12	Джерело занесення даних про кількість працівників	

Зі сформованого для організації ОПСП переліку звітних одиниць в окрему групу виділяють звітні одиниці з кількістю працівників 50 осіб і більше. Такі одиниці обстежуються щомісяця на суцільній основі. Якщо протягом періоду перебування певної одиниці з цієї групи в обстеженні кількість працюючих на неї осіб зменшилася до величини, нижчої за 50 осіб, то ця звітна одиниця продовжує обстежуватися до кінця звітного року, після чого виключається з суцільного обстеження і отримує певну ймовірність бути включеною до вибіркової сукупності.

Сукупність звітних одиниць з кількістю працівників від 10 до 49 осіб включно обстежується із використанням вибіркового методу. Вибіркова сукупність формується на основі відповідної частини переліку звітних

одиниць, сформованого для організації ОПСП, з використанням процедури стратифікованого випадкового відбору.

Дизайн вибірки забезпечує надійне оцінювання основних показників за результатами щомісячного ОПСП за сукупністю підприємств, установ, організацій та їхніх відокремлених підрозділів із кількістю працюючих від 10 до 49 осіб за такими кластерами:

- Україна в цілому;
- Україна за видами економічної діяльності (секція та підсекція за КВЕД);
- регіони в цілому.

3. ДИЗАЙН ВИБІРКИ

Дизайн вибірки визначається такими особливостями її формування, як стратифікація генеральної сукупності, схема ротації одиниць вибірки, обсяг вибірки та механізми її формування.

3.1. Стратифікація генеральної сукупності

Стратифікація сукупності звітних одиниць із кількістю працівників 10–49 осіб здійснюється за такими ознаками:

– основний вид економічної діяльності за КВЕД: секції (за винятком секцій Р “Діяльність домашніх господарств” і Q “Діяльність екстериторіальних організацій”) та підсекції (усього виділено 29 груп за видами діяльності – 13 секцій та 16 підсекцій);

– територіальна належність за місцем здійснення діяльності відповідно до адміністративно-територіального устрою України (всього 27 регіонів).

Винятком є сукупності звітних одиниць секції В “Рибне господарство” та підсекції DF “Виробництво коксу, продуктів нафтоперероблення та ядерних матеріалів”, які обстежуються на суцільній основі, що зумовлено незначною кількістю звітних одиниць (у 2007 році – 193 одиниці з кількістю працівників 10–49 осіб у секції В та 59 таких одиниць у підсекції DF).

Для забезпечення можливості реалізації ефективної процедури відбору звітних одиниць із кількістю працівників 10–49 осіб за стратами та схеми їх ротації у вибірці всі страти із загальною кількістю одиниць менше ніж 14 об’єднуються в окрему загальну страту (сукупність “малих” страт).

3.2. Схема ротації одиниць вибірки

З метою кращого охоплення обстеженням генеральної сукупності та зниження звітного навантаження на респондентів із кількістю працівників менше 50 осіб реалізується спеціальна схема ротації одиниць вибірки (рис. 1). Відповідно до цієї схеми вибіркова сукупність звітних одиниць

представлена двома рівними за розміром частинами – ротаційними групами, кожна з яких є репрезентативною вибіркою з генеральної сукупності. У кожній страті один раз на рік після актуалізації основи вибірки одна ротаційна група (50% вибіркової сукупності) заміщується новою ротаційною групою. При цьому звітні одиниці, що були відібрані до певної ротаційної групи, обстежуються два роки поспіль. Винятком є звітні одиниці першої ротаційної групи, що обстежуються лише один рік. Реалізована в ОПСП схема ротації забезпечує для одиниць, відібраних для обстеження, після двох років перебування у вибірці п'ятирічні “канікули”, протягом яких вони теоретично не можуть потрапити до вибірки.

		Ротаційна група						
		1	2	3	4	5	6	7
Рік	2010	█	█					
	2011		█	█				
	2012			█	█			
	2013				█	█		
	2014					█	█	
	2015						█	█
	2016	█						█
	2017	█	█					
	2018		█	█				
	2019			█	█			
	2020				█	█		
	2021					█	█	
	2022						█	█

Рис. 1. Схема ротації звітних одиниць у вибірковій сукупності

Використання принципу ротації звітних одиниць у вибірці дозволяє підвищити рівень надійності порівнянь показників у динаміці, не перевантажуючи будь-яку окрему групу одиниць занадто довгим періодом опитування, оскільки, з одного боку, забезпечується оновлення вибіркової сукупності для кожного року спостереження, а з іншого – зберігається незмінною певна її частина (50%). Крім того, використання такої схеми ротації дозволяє більш надійно оцінювати зміни показників між суміжними місяцями, кварталами та роками.

3.3. Визначення обсягу вибіркової сукупності

Основними показниками, що оцінюються за результатами ОПСП, є кількість працівників, фонд оплати праці, фонд робочого часу. Для визначення обсягу вибірки насамперед здійснюється аналіз варіативності відповідних ознак за одиницями. Обсяг вибірки визначається на основі тієї ознаки, яка характеризується найбільшою варіативністю, оскільки для

показника, що розраховується за цією ознакою, слід очікувати найбільшу похибку вибірки i , відповідно, найнижчу надійність. За основну характеристику надійності у цій методиці прийнято коефіцієнт варіації вибірових оцінок CV . Для ОПСП такою найбільш варіативною ознакою є фонд оплати праці. Враховуючи, що він статистично пов'язаний із кількістю працюючих, під час розрахунку обсягу вибірки за визначальний показник прийнято середньорічну заробітну плату працівника підприємства, установи організації (відокремленого підрозділу).

При визначенні загального обсягу вибірки реалізується такий алгоритм розрахунків.

3.3.1. Визначення обсягу вибірки n , що забезпечує надійну оцінку показника середньої заробітної плати для національного рівня

Загальна формула для визначення обсягу вибірки n у випадку, коли критерієм надійності є величина коефіцієнта варіації оцінки показника, може бути представлена у вигляді:

$$n = \frac{deff \cdot \sigma^2 \cdot 100^2}{\bar{Y}^2 \cdot CV(\bar{Y})^2}, \quad (1)$$

де $deff$ – дизайн-ефект; $\bar{Y} = \frac{\sum_{i=1}^N y_i}{N}$ – середнє значення заробітної плати у

генеральній сукупності; y_i – середня заробітна плата по i -й звітній одиниці; N – загальна кількість звітних одиниць у генеральній сукупності;

$\sigma^2 = \frac{\sum_{i=1}^N (y_i - \bar{Y})^2}{N - 1}$ – дисперсія показника середньої заробітної плати за

генеральною сукупністю; $CV(\bar{Y})$ – коефіцієнт варіації вибірових оцінок середнього значення показника Y , який відповідає встановленим вимогам щодо надійності результатів ОПСП на національному рівні.

При здійсненні розрахунків за формулою (1) величини \bar{Y} , σ^2 визначаються не для всієї сукупності звітних одиниць із кількістю працівників 10–49 осіб, а після видалення 0,5% кількості одиниць із початку та з кінця їх переліку, упорядкованого за зростанням величини середньої заробітної плати y_i . Тобто це так звані зрізані значення величин \bar{Y} і σ^2 . Такий підхід забезпечує зменшення впливу нетипових за ознаками, що розглядаються, одиниць на оцінки обсягів вибірки. Урахування нетипових одиниць, виявлених за результатами ОПСП, буде здійснюватися на етапі реалізації процедури оцінювання показників.

3.3.2. Визначення теоретичних обсягів вибірки, що забезпечують надійну оцінку показника середньої заробітної плати за регіонами та за видами економічної діяльності

Для кожного регіону України за формулою (1) розраховуються обсяги вибірок n_i , необхідні для надійного оцінювання заробітної плати. Визначається обсяг вибірки n_I по Україні як сума усіх обсягів, отриманих для кожного регіону:

$$n_I = \sum_i n_i, \quad (2)$$

де i – код регіону за КОАТУУ.

Далі аналогічно для кожного виду економічної діяльності за формулою (1) розраховуються обсяги вибірок n_j , що забезпечують оцінювання заробітної плати за видами економічної діяльності на національному рівні з установленим рівнем надійності. Визначається новий обсяг вибірки n_J по Україні як сума всіх обсягів, отриманих для кожного виду діяльності:

$$n_J = \sum_j n_j, \quad (3)$$

де j – код виду економічної діяльності за КВЕД (секції, підсекції).

За теоретичний обсяг вибірки приймається найбільший із розрахованих обсягів вибірки n_I і n_J , тобто: $n' = \max(n_I; n_J)$.

3.3.3. Розподіл теоретичного обсягу вибірки n' за регіонами та видами економічної діяльності

Теоретичний обсяг вибірки n' розподіляється за групами звітних одиниць генеральної сукупності за регіонами з урахуванням кількості звітних одиниць та дисперсії показника середньої заробітної плати у кожній групі за формулою:

$$n'_i = n' \cdot \frac{\sqrt[3]{N_i \cdot \sigma_i}}{\sum_i \sqrt[3]{N_i \cdot \sigma_i}}, \quad (4)$$

де N_i – загальна кількість звітних одиниць у i -му регіоні, σ_i – середньоквадратичне відхилення величини середньої заробітної плати працівників за звітними одиницями i -го регіону.

Теоретичний обсяг вибірки n' розподіляється за групами звітних одиниць генеральної сукупності, сформованих за видами економічної

діяльності, з урахуванням загальної кількості звітних одиниць та дисперсії середньої заробітної плати у кожній групі за формулою, аналогічною (4):

$$n'_j = n' \cdot \frac{\sqrt[3]{N_j \cdot \sigma_j}}{\sum_j \sqrt[3]{N_j \cdot \sigma_j}}, \quad (5)$$

де N_j – загальна кількість звітних одиниць у групі, що відповідає j -му виду економічної діяльності, σ_j – середньоквадратичне відхилення величини середньої заробітної плати працівників за звітними одиницями j -го виду економічної діяльності.

Слід зазначити, що розміщення вибірки за формулами (4), (5) є проміжним варіантом між відомим способом розміщення за формулою Неймана [3] та розміщенням за формулою (1). Обраний спосіб розміщення вибірки забезпечує більш однорідну надійність оцінювання показників у межах окремих видів діяльності порівняно з розміщенням за формулою Неймана і одночасно порівняно з розміщенням без урахування загальної кількості звітних одиниць, зменшує ймовірність реалізації випадків, коли обсяг вибірки по групі перевищує обсяг генеральної сукупності у ній.

Для забезпечення можливості реалізації схеми ротації отримані значення n'_i та n'_j коригуються за тими групами звітних одиниць, де обсяг вибірки перевищує 2/7 від кількості звітних одиниць генеральної сукупності у групі. За цими групами обсяги вибірки приймаються рівними 2/7 від загальної кількості звітних одиниць у групі. Максимальне значення частки відбору звітних одиниць за групами, яке становить 2/7, визначене з урахуванням схеми ротації одиниць у вибірці. Для забезпечення перебування певної ротаційної групи у вибірці протягом двох років із подальшими п'ятирічними “канікулами” для респондентів цієї групи необхідно відібрати сім ротаційних груп, які будуть обстежені протягом семи років, так що кожного року в обстеженні братимуть участь лише дві ротаційні групи (рис. 1).

Скоригований обсяг вибірки по Україні n'' визначається як сума всіх обсягів, отриманих для кожної групи звітних одиниць після коригування теоретичного обсягу. Обсяг вибірки n'' розподіляється за регіонами або видами економічної діяльності за формулами, аналогічними (4), (5).

Передбачається, що обсяг щорічної вибірки n'' буде використовуватись для проведення ОПСП протягом певного періоду часу, який складає за попереднім визначенням 7 років. У разі необхідності дизайн вибірки може бути переглянуто раніше.

Для розподілу скоригованого обсягу вибірки за стратами необхідно врахувати, що, як уже зазначалось, відповідно до схеми ротації вибірка складається з двох ротаційних груп, одна з яких кожного року замінюється на нову, і при цьому на весь семирічний цикл обстеження у кожній страті

необхідно відібрати сім ротаційних груп. Таким чином, необхідно забезпечити можливість відбору зі сформованих страт кожної ротаційної групи окремо. Загальний скоригований обсяг вибірки за сімома ротаційними групами n''_{Σ} становить $n''_{\Sigma} = \frac{7}{2} \cdot n''$.

Для розподілу обсягу вибірки n'' за стратами реалізується така процедура.

1). За всіма стратами встановлюється обсяг вибірки, що становить 4 одиниці. Тим самим приймається, що у жодній ротаційній групі не може бути менше 2 відібраних звітних одиниць.

2). Обсяг вибірки, що залишився нерозподіленим і становить $n'' - S \cdot 4$ (де S – загальна кількість страт), розподіляється за стратами з використанням процедури ітераційної пропорційної апроксимації (далі – процедура ПА), яка реалізується так, щоб забезпечити визначені на попередніх етапах узгоджені розподіли обсягу вибірки за видами економічної діяльності та за регіонами. При цьому страти, в яких обсяг генеральної сукупності дорівнює обсягу вибірки за сімома ротаційними групами в них – $n''_{\Sigma i,j}$, виключаються із процедури ПА.

3.3.4. Реалізація процедури ітераційної пропорційної апроксимації

Ефективним методом узгодження розподілів вибірки за стратами з розподілами за регіонами та видами економічної діяльності є класичний метод ітераційної пропорційної апроксимації або метод відхилення. Головна ідея застосування методу для розміщення вибірки полягає в коригуванні обсягів вибірки за стратами так, щоб забезпечити відповідність сумарних обсягів вибірки за стратами, які відповідають певним кластерам (регіон, вид економічної діяльності), розрахованим розподілам вибірки за цими кластерами по черзі, а потім повторювати весь процес певну кількість разів, доки не буде одержано задовільно близький збіг для всіх розподілів, що контролюються. Доцільність застосування цього підходу зумовлена, головним чином, відсутністю надійних даних щодо дисперсії показників за стратами, що не дозволяє застосувати для розміщення вибірки формули аналогічні (1) або (4).

Вихідними даними для реалізації процедури ПА є розподіли теоретичного обсягу вибірки n'' за регіонами n''_i (де i – код регіону за КОАТУУ) та видами економічної діяльності n''_j (де j – код виду економічної діяльності за КВЕД), а також дані щодо кількості страт і загальної кількості одиниць основи вибірки за кожною стратою.

Процедура ПА реалізується шляхом виконання таких основних етапів:

1). Усі страти, в яких обсяг генеральної сукупності дорівнює обсягу вибірки за сімома ротаційними групами в них ($n''_{\Sigma i,j}$), виключаються з процедури ПА.

2). Розподіли скоригованих обсягів вибірки за регіонами n_i'' та за видами економічної діяльності n_j'' коригуються так, щоб урахувати сумарні обсяги вибірки з виключених страт. У результаті отримуються нові розподіли обсягів вибірки за регіонами n_i''' та видами економічної діяльності n_j''' . При цьому повинні виконуватись умови $n_i'' \geq n_i'''$ і $n_j'' \geq n_j'''$.

3). За кожною стратою s , що задіяна у процедурі ППА, встановлюється початковий обсяг вибірки, який складає 4 звітні одиниці ($n_{i,j}^{(0)}=4$).

4). Розраховується сумарний обсяг вибірки за всіма стратами, що задіяні у процедурі ППА та входять до i -го регіону: $\sum_{s \in i} n_{i,j}^{(0)} = n_i^{(1)}$. Розрахунок здійснюється за всіма регіонами без винятку.

5). За кожною стратою s розраховується коефіцієнт коригування обсягу вибірки страти $q_{i,j}^{(1)} = \frac{n_i''}{n_i^{(1)}}$.

6). За кожною стратою s розраховується новий обсяг вибірки $n_{i,j}^{(1)} = n_{i,j}^{(0)} \cdot q_{i,j}^{(1)}$.

7). Розраховується сумарний обсяг вибірки за всіма стратами s , що входять до j -го виду економічної діяльності: $\sum_{s \in j} n_{i,j}^{(1)} = n_j^{(2)}$. Розрахунок здійснюється за всіма видами діяльності без винятку.

8). За кожною стратою s розраховується коефіцієнт коригування обсягу вибірки страти $q_{i,j}^{(2)} = \frac{n_j''}{n_j^{(2)}}$.

9). За кожною стратою s розраховується новий обсяг вибірки $n_{i,j}^{(2)} = n_{i,j}^{(1)} \cdot q_{i,j}^{(2)}$.

Етапи 4–9 повторюються до тих пір, поки коефіцієнти коригування обсягу вибірки за кожною стратою не будуть рівними одиниці з установленим рівнем точності. Як правило, достатнім є виконання умови $q_{i,j}^{(iter)} = 1 \pm 0,05$, де $iter$ – номер останньої реалізованої ітерації. Орієнтовна кількість необхідних ітерацій на практиці становить близько 10.

3.3.5. Коригування обсягу вибірки за стратами з урахуванням очікуваного рівня участі респондентів в обстеженні

Для отримання підсумкової кількості звітних одиниць, яку необхідно відібрати за кожною стратою $n_{i,j}^{(res)}$, кількість звітних одиниць, що відповідають останній реалізованій ітерації $n_{i,j}^{(iter)}$, необхідно скоригувати з урахуванням очікуваного рівня участі в обстеженні відібраних звітних одиниць із кількістю працівників від 10 до 49 осіб включно.

Очікуваний рівень участі зазначених звітних одиниць в ОПСП R_c обчислюється як відношення очікуваної кількості одиниць, що будуть обстежені, до кількості відібраних одиниць:

$$R_c = \frac{\text{кількість обстежених звітних одиниць з кількістю працівників від 10 до 49 осіб}}{\text{кількість відібраних звітних одиниць з кількістю працівників від 10 до 49 осіб}}. \quad (6)$$

Величина R_c може розраховуватись за основними групами одиниць спостереження, що визначаються під час стратифікації генеральної сукупності.

На цьому етапі ОПСП на основі експертних оцінок очікуваний рівень участі встановлено на рівні 0,9 (90%) для всіх груп звітних одиниць. Таким чином, підсумкова кількість звітних одиниць для відбору за кожною стратою становить:

$$n_{i,j}^{(res)} = \frac{n_{i,j}^{(iter)}}{R_c} = \frac{n_{i,j}^{(iter)}}{0,9}. \quad (7)$$

Доцільним є встановлення $n_{i,j}^{(res)}$ кратним 2, що несуттєво змінить загальний обсяг вибірки, але значно спростить процедури її формування та контролю. Для страт, у яких обсяг генеральної сукупності дорівнює обсягу вибірки за сімома ротаційними групами ($n''_{i,j}$), приймається, що підсумковий обсяг вибірки дорівнює скоригованому обсягу вибірки $n_{i,j}^{(res)} = n''_{i,j}$.

4. ВІДБІР ЗВІТНИХ ОДИНИЦЬ

4.1. Відбір звітних одиниць за стратами

Для реалізації процедур відбору звітних одиниць із кількістю працівників 10–49 осіб за стратами перелік одиниць, який складає основу вибірки в кожній страті, упорядковується за зменшенням кількості працюючих та нумерується. Це забезпечує неявну стратифікацію генеральної сукупності за розміром і підвищує якість результатів обстеження. Далі визначається кількість звітних одиниць $n_{i,j,r}$, що складають r -ту ротаційну групу та які необхідно відібрати. Відбір здійснюється із застосуванням процедури систематичного відбору. Процедура відбору складається з трьох основних етапів: розрахунок кроку відбору, розрахунок номера першої звітної одиниці, відбір звітних одиниць для обстеження.

Крок відбору $h_{i,j,r}$ розраховується за формулою:

$$h_{i,j,r} = \frac{N_{i,j}}{n_{i,j,r}}, \quad (7)$$

де $N_{i,j}$ – кількість звітних одиниць у генеральній сукупності у страті.

Розраховане значення кроку відбору для здійснення розрахунків заокруглюється за правилами математики до трьох цифр після коми. Номер першої звітної одиниці $h_{i,j,r}^{(1)}$, що включається до вибіркової сукупності, визначається шляхом множення кроку відбору $h_{i,j,r}$ на спеціально згенероване, рівномірно розподілене від 0 до 1 випадкове число $\pi_{i,j,r}$:

$$h_{i,j,r}^{(1)} = h_{i,j,r} \cdot \pi_{i,j,r}. \quad (8)$$

Подальший відбір звітних одиниць здійснюється шляхом систематичного додавання кроку відбору до попередньо розрахованого значення за формулами:

$$\begin{aligned} h_{i,j,r}^{(2)} &= h_{i,j,r}^{(1)} + h_{i,j,r}; \\ h_{i,j,r}^{(3)} &= h_{i,j,r}^{(2)} + h_{i,j,r}; \\ &\vdots \\ h_{i,j,r}^{(n_{i,j,r})} &= h_{i,j,r}^{(n_{i,j,r}-1)} + h_{i,j,r}. \end{aligned} \quad (9)$$

Для визначення номерів звітних одиниць, що включаються до вибірки, береться тільки ціла частина розрахованих значень незалежно від дробової частини. При відборі звітних одиниць чергової ротаційної групи одиниці, що увійшли до відібраних раніше ротаційних груп, не можуть бути відібраними знову. Якщо номери таких звітних одиниць потрапляють до вибірки при реалізації процедури відбору, то береться наступна за номером одиниця, що не була відібрана до жодної ротаційної групи.

На першому етапі організації ОПСП, відповідно до схеми ротації, відбираються звітні одиниці, що складають першу та другу ротаційні групи.

Далі процедура відбору здійснюватиметься один раз на рік перед заміною чергової ротаційної групи згідно зі схемою ротації після актуалізації основи вибірки.

Імовірність відбору звітних одиниць $P_{i,j,k}$ для r -ї ротаційної групи визначається за формулою:

$$P_{i,j,r} = \frac{n_{i,j,r}}{N_{i,j,r}} = \frac{1}{h_{i,j,r}}. \quad (10)$$

4.2. Вибір звітних одиниць за сукупністю “малих” страт і за стратами, в яких обсяг вибірки дорівнює обсягу генеральної сукупності

Для реалізації процедур відбору звітних одиниць із кількістю працівників 10–49 осіб за сукупністю “малих” страт перелік одиниць, який складає основу вибірки в сукупності “малих” страт, упорядковується за видом економічної діяльності та за зменшенням кількості працівників, після чого звітні одиниці нумеруються наскрізно. Вибір здійснюється із застосуванням процедури систематичного відбору, що описана у п. 4.1.

У стратах, де результатний обсяг вибірки $n_{i,j}^{(res)}$ дорівнює обсягу генеральної сукупності, кількість звітних одиниць у r -й ротаційній групі вибірки складає $1/7$ від $n_{i,j}^{(res)}$. Для реалізації процедур відбору звітних одиниць із кількістю працівників 10–49 осіб за такими стратами перелік одиниць, який складає основу вибірки в кожній страті, упорядковується за зменшенням кількості працюючих. Далі здійснюється наскрізна нумерація звітних одиниць. Вибір реалізується із застосуванням процедури систематичного відбору (п. 4.1).

4.3. Коригування обсягу вибірки для забезпечення оцінювання показників за адміністративно-територіальними одиницями

Після відбору звітних одиниць здійснюється перевірка обсягів сукупності для ОПСП (включаючи звітні одиниці з кількістю працівників 50 і більше) за адміністративно-територіальними одиницями. За тими адміністративно-територіальними одиницями, де обсяг сукупності звітних одиниць для обстеження є меншим, ніж установлений мінімальний обсяг (цей обсяг визначається на основі експертних оцінок з урахуванням якості оцінювання показників), відбирається додаткова кількість звітних одиниць із кількістю 10–49 осіб (додаткова вибірка). Вибір здійснюється, щоб забезпечити за цими адміністративно-територіальними одиницями загальний обсяг сукупності звітних одиниць, що необхідно обстежити, тобто не менший, ніж мінімальний. Звітні одиниці цієї звітної одиниці відбираються з переліку всіх звітних одиниць із кількістю працівників від 10 до 49 осіб включно адміністративно-територіальної одиниці з використанням процедури систематичного відбору, з попереднім упорядкуванням звітних одиниць за видами економічної діяльності та за зменшенням кількості працівників.

Імовірність відбору кожної звітної одиниці, відібраної у r -й адміністративно-територіальній одиниці, визначається за формулою:

$$p_{rn,r} = \frac{n_{rn,r}}{N_{rn,r}}, \quad (11)$$

де $n_{m,r}$ – кількість звітних одиниць із кількістю працівників 10–49 осіб, відібраних у g -й адміністративно-територіальній одиниці, $N_{m,r}$ – кількість звітних одиниць із кількістю працівників 10–49 осіб в основі вибірки r - і адміністративно-територіальної одиниці.

Слід ураховувати, що відповідно до реалізованих процедур, кожна звітна одиниця з кількістю працівників 10–49 осіб з основної вибірки характеризується двома ймовірностями включення до вибірки: $P_{i,j,k}$ та $P_{m,r}$. Кожна звітна одиниця з кількістю працівників 10–49 осіб із додаткової вибірки характеризується лише ймовірністю включення до вибірки $P_{m,r}$. Це важливо для реалізації адекватних процедур оцінювання показників.

5. КОНТРОЛЬ ЯКОСТІ ФОРМУВАННЯ ВИБІРКИ

Контроль якості формування вибірки, що здійснюється після реалізації процедур відбору, є обов'язковим етапом робіт з організації ОПСП. Основними процедурами контролю є такі:

1). Порівняння кількості відібраних звітних одиниць по кожній страті та по кожному кластеру (Україна, регіони, види економічної діяльності) з обсягами вибірки, визначеними по відповідних стратах та кластерах на етапі розробки дизайну вибірки. Виявлені відхилення аналізуються та виправляються.

2). Порівняння обсягів вибірки по кожній страті та по кожному кластеру з обсягами генеральної сукупності по відповідних стратах та кластерах і розраховуються частки відбору. Величини часток відбору аналізуються, при цьому особлива увага приділяється їх екстремальним значенням.

3). Аналіз ймовірностей відбору звітних одиниць із кількістю працівників 10–49 осіб. У випадку виявлення нетипово малих ймовірностей здійснюється перевірка процедур відбору одиниць.

4). Розрахунок зважених обсягів вибірки за кластерами: Україна, регіони, види економічної діяльності. При цьому як вага для кожної одиниці вибірки використовується величина, обернена до ймовірності її відбору. Сума статистичних ваг звітних одиниць за кожним кластером повинна дорівнювати кількості одиниць генеральної сукупності за цим кластером.

5). Перевірка загальної кількості звітних одиниць, що підлягає обстеженню в кожній адміністративно-територіальній одиниці. Ця кількість не може бути нижчою, ніж установлена мінімальна кількість звітних одиниць.

6. АКТУАЛІЗАЦІЯ ОСНОВИ ВИБІРКИ

Актуалізація основи вибірки здійснюється один раз на рік після актуалізації СРП. Основною метою актуалізації є адекватне визначення сукупності звітних одиниць для обстеження – переліку всіх звітних одиниць із кількістю працівників 50 осіб і більше та чергової ротаційної групи звітних одиниць із кількістю працівників 10–49 осіб для вибіркового обстеження. Іншою важливою метою актуалізації є визначення поточних параметрів генеральної сукупності, насамперед кількості звітних одиниць за основними кластерами, що необхідно для підвищення рівня надійності результатів поточного ОПСП.

Під час актуалізації поточні дані СРП порівнюються з актуальними даними з інших джерел, зокрема адміністративними даними Пенсійного фонду України, даними інших державних статистичних обстежень підприємств для мінімізації кількості помилок при визначенні сукупності одиниць для обстеження.

Крім того, передбачається, що при актуалізації будуть здійснені такі процедури:

1). Усі респонденти, що в поточному році припинили діяльність, незалежно від того, чи були вони включені до сукупності поточного обстеження, виключаються з основи вибірки при її актуалізації.

2). Респонденти, у яких середня кількість найманих працівників стала меншою за 50 осіб, отримують певну ймовірність бути включеними до вибірки та обстежуватися з початку нового календарного року .

3). Респонденти, у яких середня кількість найманих працівників перевищила 49 осіб, з початку нового календарного року обстежуються на суцільній основі.

4). Респонденти, у яких середня кількість працівників стала меншою за 10 осіб, незалежно від того, чи були вони включені до сукупності одиниць поточного обстеження, виключаються з основи вибірки при її актуалізації.

5). Усі зміни у кількості звітних одиниць в основі вибірки та в обсягах вибірки враховуються при формуванні вибіркової сукупності для поточного обстеження.

Приклад реалізації процедур визначення загального обсягу вибірки

Для визначення загального обсягу вибірки реалізується алгоритм розрахунків, який складається з десяти етапів.

Перший етап.

Визначається обсяг вибірки n , що забезпечує оцінку показника заробітної плати для національного рівня з необхідним рівнем надійності.

Розрахунок обсягу вибірки здійснюється за формулою (1). Припустимо, що для національного рівня коефіцієнт варіації оцінки середньої заробітної плати $CV(\bar{Y})$, як основна характеристика надійності, не повинен перевищувати 2,5%. Величина дизайн-ефекту $deff$ приймається рівною 1,5. Величини середньорічної заробітної плати працівників підприємств, установ, організацій та відокремлених підрозділів із кількістю працівників від 10 до 49 осіб \bar{Y} та дисперсія σ^2 визначаються за генеральною сукупністю, сформованою за результатами державних статистичних спостережень та адміністративними даними за 2007 рік: $\bar{Y} = 10795,5$ грн., $\sigma^2 = 154660512,6$. Обсяг вибірки для надійного оцінювання показника середньої заробітної плати на національному рівні розраховується за формулою (1):

$$n = \frac{deff \cdot \sigma^2 \cdot 100^2}{\bar{Y}^2 \cdot CV(\bar{Y})^2} = \frac{1,5 \cdot 154660512,6 \cdot 100^2}{10795,5^2 \cdot 2,5^2} \approx 3185 \text{ звітних одиниць.}$$

Другий етап.

Визначаються обсяги вибірки, що забезпечують надійні оцінки показника середньої заробітної плати за видами економічної діяльності та за регіонами.

Розрахунок обсягів вибірки здійснюється також за формулою (1). Припустимо, що для надійного оцінювання показника середньої заробітної плати за видами економічної діяльності та за регіонами коефіцієнт варіації не повинен перевищувати 5%. Величина дизайн-ефекту приймається рівною 1,5, як і для національного рівня. Обсяги вибірки розраховуються на основі величин середньорічної заробітної плати працівників підприємств, установ, організацій та відокремлених підрозділів із кількістю працівників 10–49 осіб та дисперсії, визначених за генеральною сукупністю для кожного регіону. Наприклад, для Луганської області $\bar{Y}_{i=44} = 9367,8$ грн., $\sigma_{i=44}^2 = 39547313,5$, тоді обсяг вибірки $n_{i=44}$ дорівнює:

$$n_{i=44} = \frac{deff \cdot \sigma_{i=12}^2 \cdot 100^2}{\bar{Y}_{i=12}^2 \cdot CV(\bar{Y})^2} = \frac{1,5 \cdot 39547313,5 \cdot 100^2}{9367,8^2 \cdot 5^2} \approx 270 \text{ звітних одиниць.}$$

Обсяги вибірки, розраховані для всіх регіонів, наведено у табл. А.1.

Таблиця А.1

Обсяги вибірки, розраховані для регіонів України

Код за КОАТУУ	Регіони	Середньорічна заробітна плата, грн., \bar{Y}_i	Дисперсія, σ_i^2	Обсяг вибірки, звітних одиниць, n_i
1	АР Крим	9042,8	46139405,5	339
5	Вінницька	9234,1	41395398,0	291
7	Волинська	9143,9	45652199,6	328
12	Дніпропетровська	10836,8	61040392,1	312
14	Донецька	11446,1	63366729,5	290
18	Житомирська	9158,9	39358634,8	282
21	Закарпатська	11414,5	64834306,8	299
23	Запорізька	10245,2	54217944,6	310
26	Івано-Франківська	9597,9	49987925,1	326
32	Київська	10073,6	65005566,3	384
35	Кіровоградська	9696,8	35466143,7	226
44	Луганська	9367,8	39547313,5	270
46	Львівська	9911,7	58669060,2	358
48	Миколаївська	10560,3	53322953,5	287
51	Одеська	8551,4	48259589,0	396
53	Полтавська	10547,2	54585488,2	294
56	Рівненська	9144,6	42541560,8	305
59	Сумська	11225,7	49455142,5	235
61	Тернопільська	8656,1	44925816,0	360
63	Харківська	10889,8	61348369,2	310
65	Херсонська	9927,4	43437939,0	264
68	Хмельницька	8299,5	38756567,0	338
71	Черкаська	10781,9	48846318,4	252
73	Чернівецька	10028,3	51778775,7	309
74	Чернігівська	8907,4	38098248,1	288
80	м. Київ	11815,5	102469295,5	440
85	м. Севастополь	9118,1	50567328,6	365

Для кожного з видів економічної діяльності, за формулою (1) розраховуються обсяги вибірок, що забезпечують надійне оцінювання показника середньої заробітної плати. Величини $CV(\bar{Y})$ та $deff$ приймаються такими, як і для регіонів. Обсяги вибірки розраховуються на основі величин середньорічної заробітної плати працівників підприємств, установ організацій та їхніх відокремлених підрозділів з кількістю працівників 10–49 осіб та дисперсії, визначених за генеральною сукупністю для кожного виду діяльності. Наприклад, для виду економічної діяльності „Будівництво” (секція F), $\bar{Y}_{j=18} = 10228,6$ грн., $\sigma_{j=18}^2 = 57884825,1$, необхідний обсяг вибірки

$n_{j=18}$ дорівнює:

$$n_{j=18} = \frac{deff \cdot \sigma_{j=18}^2 \cdot 100^2}{\bar{Y}_{j=18}^2 \cdot CV(\bar{Y})^2} = \frac{1,5 \cdot 57884825,1 \cdot 100^2}{10228,6^2 \cdot 5^2} \approx 332 \text{ звітні одиниці} .$$

Обсяги вибірки, розраховані для всіх видів економічної діяльності (крім секції В „Рибне господарство” та підсекції DF „Виробництво коксу, продуктів нафтоперероблення та ядерних матеріалів”), наведено у табл. А.2.

Таблиця А.2

Обсяги вибірки, розраховані для видів економічної діяльності

№	Вид діяльності	Код виду діяльності	Середньорічна заробітна плата, грн., \bar{Y}_j	Дисперсія, σ_j^2	Обсяг вибірки, звітних одиниць, n_i
1	2	3	4	5	6
1	Сільське господарство, мисливство, лісове господарство	A	8515,4	26099157,1	216
2	Добування паливно-енергетичних корисних копалин	CA	11065,5	71516783,5	350
3	Добування корисних копалин, крім паливно-енергетичних	CB	9661,8	72652731,3	467
4	Виробництво харчових продуктів, напоїв та тютюнових виробів	DA	7519,7	26077766,6	277
5	Текстильне виробництво; виробництво одягу, хутра та виробів з хутра	DB	6441,3	20681970,0	299
6	Виробництво шкіри, виробів зі шкіри та інших матеріалів	DC	6106,0	31966494,6	514
7	Оброблення деревини та виробництво виробів з деревини	DD	6395,2	15536092,3	228
8	Целюлозно-паперове виробництво; видавнича діяльність	DE	10431,8	56067535,9	309
9	Хімічне виробництво	DG	10158,4	48859136,4	284
10	Виробництво гумових та пластмасових виробів	DH	8499,5	39221318,1	326
11	Виробництво іншої неметалевої мінеральної продукції	DI	8315,7	39369183,1	342
12	Металургійне виробництво та виробництво готових металевих виробів	DJ	9564,0	52246075,8	343
13	Виробництво машин та устаткування	DK	10489,1	53094637,6	290
14	Виробництво електричного, електронного та оптичного устаткування	DL	10744,7	61392008,6	319
15	Виробництво транспортних засобів та устаткування	DM	10267,0	84925193,9	483
16	Інші галузі промисловості	DN	7494,1	30726493,3	328
17	Виробництво електроенергії, газу та води	E	10515,1	43641311,2	237
18	Будівництво	F	10228,6	57884825,1	332
19	Торгівля; ремонт автомобілів, побутових виробів та предметів особистого вжитку	G	10691,9	63046919,7	331
20	Діяльність готелів та ресторанів	H	8351,3	25099602,9	216

21	Діяльність транспорту та зв'язку	I	10548,4	70702553,9	381
Продовження табл.					
1	2	3	4	5	6
22	Фінансова діяльність	J	17223,6	203369321,2	411
23	Операції з нерухомим майном, оренда, інжиніринг та надання послуг підприємцям	K	12533,5	92911745,7	355
24	Державне управління	L	15327,2	153136566,1	391
25	Освіта	M	11216,0	63512318,8	303
26	Охорона здоров'я та надання соціальної допомоги	N	8531,5	45922356,4	379
27	Надання комунальних та індивідуальних послуг; діяльність у сфері культури та спорту	O	9858,5	41218124,9	254

Третій етап.

Визначається теоретичний обсяг вибірки n' .

Розрахунки здійснюються з використанням формул (2) та (3).

За формулою (2) розраховується обсяг вибірки n_I як сума всіх обсягів, отриманих для кожного регіону: $n_I = 8459$ звітних одиниць. Далі за формулою (3) розраховується новий обсяг вибірки n_J як сума всіх обсягів, отриманих для кожного виду діяльності: $n_J = 8965$ звітних одиниць.

За теоретичний обсяг вибірки n' приймається найбільший із розрахованих обсягів вибірки n , n_J , n_I . У цьому випадку $n' = n_J = 8965$ звітних одиниць.

Четвертий етап.

Теоретичний обсяг вибірки n' розподіляється за видами економічної діяльності.

Спочатку теоретичний обсяг вибірки n' розподіляється за видами економічної діяльності на основі формули (5). Це пояснюється тим, що теоретичний обсяг вибірки n' визначено на основі сумарного обсягу вибірки за видами економічної діяльності (як зазначено у розділі 3, дисперсія показника середньої заробітної плати за видами економічної діяльності є відносно більшою, ніж за регіонами).

Розподіл обсягу вибірки n' за видами економічної діяльності здійснюється з урахуванням кількості звітних одиниць та дисперсії показника середньої заробітної плати. Наприклад, для виду економічної діяльності „Будівництво” (секція F) загальна кількість звітних одиниць у генеральній сукупності $N_{j=18}$ складає $N_{j=18} = 9832$ одиниць $\sigma_{j=18}^2 = 57884825,1$.

Знаменник формули (4) у цьому випадку дорівнює $\sum_{j=1}^{27} \sqrt[3]{N_j} \cdot \sigma_j = 2682520,1$. Тоді обсяг вибірки $n'_{j=18}$ для виду економічної діяльності, що розглядається, дорівнює:

$$n'_{j=18} = n' \cdot \frac{\sqrt[3]{N_{j=18}} \cdot \sigma_{j=18}}{\sum_j \sqrt[3]{N_j} \cdot \sigma_j} = 8965 \cdot \frac{\sqrt[3]{9832} \cdot \sqrt{57884825,1}}{2682520,1} \approx 545 \text{ звітних одиниць.}$$

Для підсекцій СА, СВ, DC, DG, DM отримані значення n'_j перевищують 2/7 від кількості звітних одиниць генеральної сукупності, тому для них теоретичні обсяги вибірки зменшуються та приймаються рівними 2/7 від кількості одиниць у генеральній сукупності.

П'ятий етап.

Визначається новий обсяг вибірки n'' .

За формулою (2) розраховується новий обсяг вибірки n'' як сума всіх обсягів, отриманих для кожного виду діяльності. Для прикладу, що розглядається, $n'' = 8520$ звітних одиниць. Обсяги вибірки, розраховані для всіх видів економічної діяльності, наведено у табл. А.3.

Таблиця А.3

Розподіл обсягу вибірки за видами економічної діяльності

№	Вид діяльності	Код виду діяльності	Кількість звітних одиниць у генеральній сукупності, N_j	Дисперсія, σ_j^2	Обсяг вибірки, звітних одиниць, n''_j
1	2	3	4	5	6
1	Сільське господарство, мисливство, лісове господарство	A	7622	26099157,1	336
2	Добування паливно-енергетичних корисних копалин	CA	38	71516783,5	10
3	Добування корисних копалин, крім паливно-енергетичних	CB	105	72652731,3	30
4	Виробництво харчових продуктів, напоїв та тютюнових виробів	DA	2281	26077766,6	225
5	Текстильне виробництво; виробництво одягу, хутра та виробів з хутра	DB	1337	20681970,0	167
6	Виробництво шкіри, виробів зі шкіри та інших матеріалів	DC	73	31966494,6	20
7	Оброблення деревини та виробництво виробів з деревини	DD	959	15536092,3	130
8	Целюлозно-паперове виробництво; видавнича діяльність	DE	1912	56067535,9	311
9	Хімічне виробництво	DG	293	48859136,4	82

Продовження табл.

1	2	3	4	5	6
10	Виробництво гумових та пластмасових виробів	DH	669	39221318,1	183
11	Виробництво іншої неметалевої мінеральної продукції	DI	1048	39369183,1	213
12	Металургійне виробництво та виробництво готових металевих виробів	DJ	1304	52246075,8	264
13	Виробництво машин та устаткування	DK	1496	53094637,6	279
14	Виробництво електричного, електронного та оптичного устаткування	DL	1091	61392008,6	270
15	Виробництво транспортних засобів та устаткування	DM	226	84925193,9	64
16	Інші галузі промисловості	DN	1067	30726493,3	189
17	Виробництво електроенергії, газу та води	E	918	43641311,2	215
18	Будівництво	F	9832	57884825,1	545
19	Торгівля; ремонт автомобілів, побутових виробів та предметів особистого вжитку	G	19669	63046919,7	716
20	Діяльність готелів та ресторанів	H	2849	25099602,9	237
21	Діяльність транспорту та зв'язку	I	4570	70702553,9	466
22	Фінансова діяльність	J	2211	203369321,2	621
23	Операції з нерухомим майном, оренда, інжиніринг та надання послуг підприємцям	K	15102	92911745,7	796
24	Державне управління	L	16184	153136566,1	1046
25	Освіта	M	2643	63512318,8	368
26	Охорона здоров'я та надання соціальної допомоги	N	3267	45922356,4	336
27	Надання комунальних та індивідуальних послуг; діяльність у сфері культури та спорту	O	6533	41218124,9	401

Шостий етап.

Обсяг вибірки n'' розподіляється за регіонами.

Обсяг вибірки n'' розподіляється за регіонами за формулою (4). Наприклад, для Луганської області загальна кількість звітних одиниць у генеральній сукупності $N_{i=44}$ становить $N_{i=44} = 3793$ звітні одиниці, $\sigma_{i=44}^2 = 39547313,5$. Знаменник формули (4) у цьому випадку дорівнює $\sum_{i=1}^{27} \sqrt[3]{N_i} \cdot \sigma_i = 2958483,7$. Тоді обсяг вибірки $n''_{i=44}$ для регіону, що розглядається, дорівнює:

$$n''_{i=44} = n'' \cdot \frac{\sqrt[3]{N_{i=12}} \cdot \sigma_{i=12}}{\sum_i \sqrt[3]{N_i} \cdot \sigma_i} = 8520 \cdot \frac{\sqrt[3]{3793} \cdot \sqrt{39547313,5}}{2958483,7} \approx 282 \text{ звітні одиниці.}$$

Розподіл обсягу вибірки n'' за регіонами наведено у табл. А.4.

Таблиця А.4

Розподіл обсягу вибірки за регіонами України

Код за КОАТУУ	Регіони	Кількість звітних одиниць у генеральній сукупності, N_i	Дисперсія, σ_i^2	Обсяг вибірки звітних одиниць, n''_i
1	АР Крим	3955	46139405,5	309
5	Вінницька	3436	41395398,0	280
7	Волинська	2177	45652199,6	252
12	Дніпропетровська	6934	61040392,1	429
14	Донецька	7638	63366729,5	451
18	Житомирська	2972	39358634,8	260
21	Закарпатська	1960	64834306,8	290
23	Запорізька	3872	54217944,6	333
26	Івано-Франківська	2382	49987925,1	272
32	Київська	4383	65005566,3	380
35	Кіровоградська	2236	35466143,7	224
44	Луганська	3793	39547313,5	282
46	Львівська	6149	58669060,2	404
48	Миколаївська	2659	53322953,5	291
51	Одеська	5574	48259589,0	355
53	Полтавська	3363	54585488,2	319
56	Рівненська	2133	42541560,8	242
59	Сумська	2525	49455142,5	276
61	Тернопільська	2278	44925816,0	254
63	Харківська	6494	61348369,2	421
65	Херсонська	2358	43437939,0	253
68	Хмельницька	2958	38756567,0	257
71	Черкаська	2868	48846318,4	286
73	Чернівецька	1446	51778775,7	234
74	Чернігівська	2480	38098248,1	241
80	м. Київ	15342	102469295,5	724
85	м. Севастополь	934	50567328,6	200
	Усього	105299		8520

Сьомий етап.

Обсяг вибірки n'' розподіляється за стратами.

Розподіл обсягу вибірки n'' за стратами здійснюється з використанням процедури ітераційної пропорційної апроксимації.

Восьмий етап.

Коригування обсягів вибірки у кожній страті $n''_{i,j}^{(iter)}$ для врахування очікуваного рівня участі респондентів в обстеженні.

На основі експертних оцінок передбачається, що очікуваний рівень участі становитиме близько 0,9 (90% відібраних звітних одиниць візьмуть участь у обстеженні). Таким чином, за формулою (7) по кожній страті коригуються обсяги вибірки $n_{i,j}^{(iter)}$. Отримані після коригування непарні величини $n_{i,j}^{(res)}$ збільшуються на одиницю для забезпечення кратності двом. Далі за формулами, аналогічними (2) та (3), розраховується загальний обсяг вибірки. Для прикладу, що розглядається, $n^{(res)} = 9372$ звітних одиниць.

Дев'ятий етап.

Відбір звітних одиниць за стратами

Для реалізації процедур відбору звітних одиниць з кількістю працівників 10–49 осіб за стратами перелік одиниць, який складає основу вибірки в кожній страті, упорядковується за зменшенням кількості працівників та нумерується. Визначається кількість звітних одиниць

$n_{i,j,r} = \frac{1}{2} n_{i,j}^{(res)}$, що складають r -ту ротаційну групу та які необхідно відібрати.

Відбір здійснюється із застосуванням процедури систематичного відбору.

Десятий етап.

Коригування загального обсягу вибірки для забезпечення оцінювання показників за адміністративно-територіальними одиницями.

Після відбору звітних одиниць здійснюється перевірка обсягів сукупності для ОПСП (включаючи одиниці з кількістю працівників 50 і більше) за адміністративно-територіальними одиницями.

За результатами перевірки обсягів сукупності звітних одиниць для ОПСП за адміністративно-територіальними одиницями встановлено, що майже у 11% з них обсяги сукупностей є меншими, ніж установлений мінімальний обсяг (на цьому етапі цей обсяг складає 18 звітних одиниць). Для цих адміністративно-територіальних одиниць відбирається додаткове число звітних одиниць із кількістю працівників 10–49 осіб, щоб забезпечити обсяги сукупності для ОПСП не менші, ніж 18 звітних одиниць.

Наприклад, у Київській області у Поліському районі налічується 4 звітні одиниці з кількістю працівників 50 осіб і більше. Для вибіркового обстеження відібрано 2 звітні одиниці. Таким чином, загальна кількість звітних одиниць для ОПСП у цьому районі дорівнює 6. Для забезпечення мінімальної кількості звітних одиниць у районі відбирається ще 12 одиниць серед тих, що не були включені до вибірки на попередніх етапах її формування.

У табл. А.5 наведено розподіл сукупності звітних одиниць ОПСП за регіонами.

Розподіл сукупності звітних одиниць ОПСП за регіонами

Регіон	Вибіркове обстеження	Суцільне обстеження			Кількість звітних одиниць з кількістю працівників ≥ 10 осіб для ОПСП
	Загальна кількість звітних одиниць для відбору	Кількість звітних одиниць підсекції DF	Кількість звітних одиниць секції В	Кількість звітних одиниць з кількістю працівників ≥ 50 осіб	
АР Крим	391	6	18	1869	2284
Вінницька	327	2	10	1671	2010
Волинська	309	0	6	1017	1332
Дніпропетровська	499	6	13	3140	3658
Донецька	525	6	6	3846	4383
Житомирська	306	1	2	1244	1553
Закарпатська	349	0	1	953	1303
Запорізька	391	2	13	1702	2108
Івано-Франківська	324	1	6	926	1257
Київська	442	0	12	1880	2334
Кіровоградська	274	0	7	1035	1316
Луганська	328	0	0	2180	2508
Львівська	471	4	4	2371	2850
Миколаївська	342	2	6	1131	1481
Одеська	416	0	22	2228	2666
Полтавська	373	6	5	1625	2009
Рівненська	291	0	1	1033	1325
Сумська	325	3	3	1197	1528
Тернопільська	300	0	1	978	1279
Харківська	488	4	4	2666	3162
Херсонська	306	1	9	994	1310
Хмельницька	303	0	0	1299	1602
Черкаська	338	0	14	1387	1739
Чернівецька	284	1	2	636	923
Чернігівська	283	2	3	1178	1466
м. Київ	728	12	4	5175	5919
м. Севастополь	255	0	21	408	684
Усього	9968	59	193	45769	55989
Частка одиниць для обстеження, %	9,4%	100,0%	100,0%	100,0%	36,8%

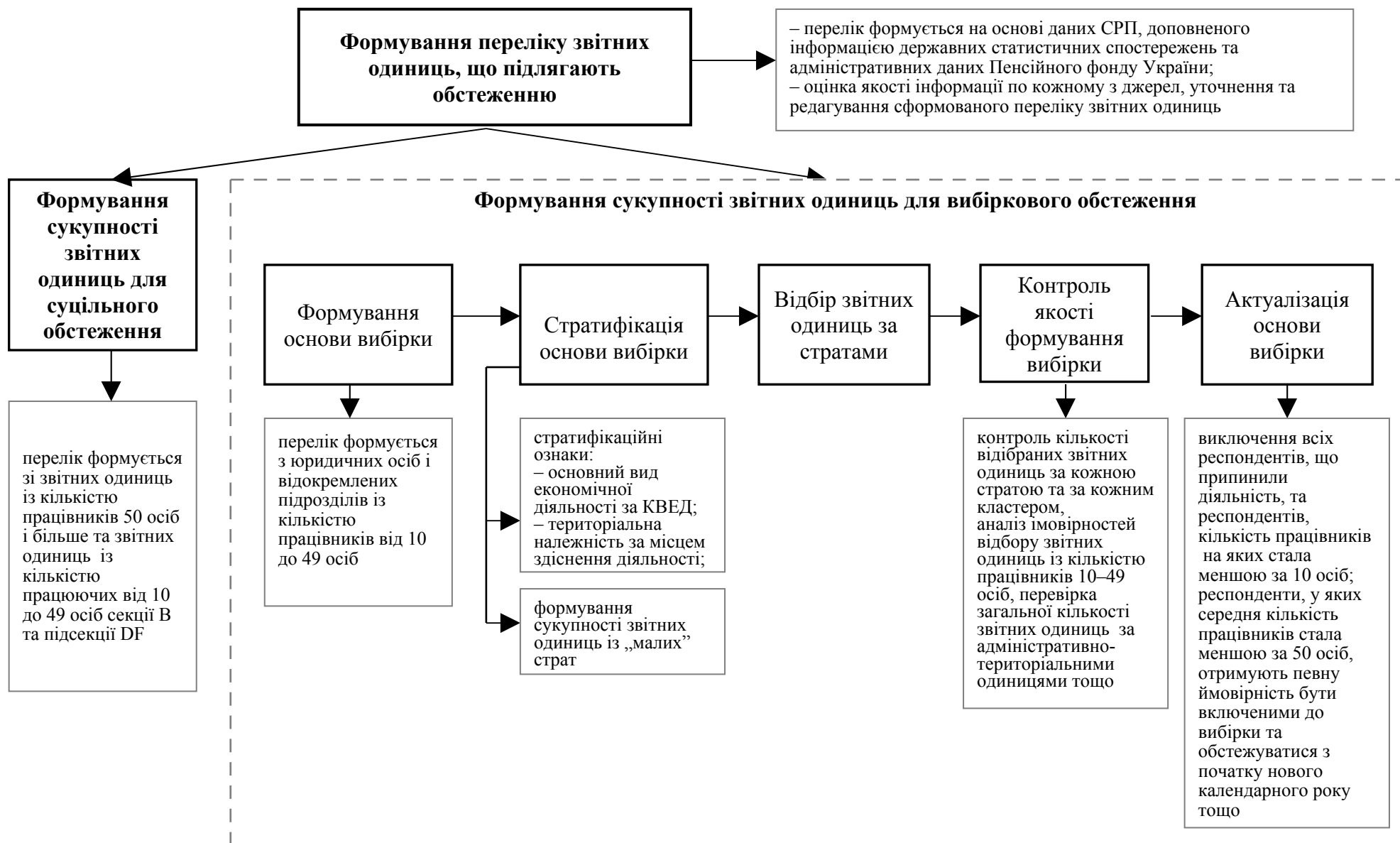


Рис. Б. Основні етапи формування сукупності звітних одиниць для ОПСП

ПЕРЕЛІК ДЖЕРЕЛ

1. Вибіркове спостереження : Термінологічний словник / [уклад. О. І. Черняк, Є. М. Жуйкова, О. В. Гончар та ін.] ; під наук. кер. О. О. Васечко — К. : НТК статистичних досліджень, 2004. — 140 с.
2. Методологічні положення з підготовки стандартних звітів щодо якості результатів державних вибірових обстежень населення (домогосподарств). — К. : Державний комітет статистики України, 2008. — 92 с.
3. Кокрен У. Методы выборочного исследования / Кокрен У. — М. : Статистика, 1976. — 440 с.
4. Kish L. Survey sampling / Kish L. — New York : John Wiley & Sons, 1995. — 643 p.
5. Єріна А. М. Організація вибірових обстежень : [навч. посіб.] / Єріна А. М. — К. : КНЕУ, 2004. — 127 с.
6. Technical Notes to Establishment Survey Data Published in Employment and Earnings [Електронний ресурс]. — Режим доступу : www.bls.gov/sae/790stderr.htm — Назва з титул. екрана.
7. Smith P. Developments in business survey methodology in the Office for National Statistics, 1994–2000. / Smith P., Pont M., Jones T. // The Statistician. — 2003. — 52, Part 3. — P. 257—295.