

Державна служба статистики України

ЗАТВЕРДЖЕНО

Наказ Державної служби
статистики України

26.02.2018 № 32

**МЕТОДОЛОГІЧНІ ПОЛОЖЕННЯ ЩОДО
ФОРМУВАННЯ ВИБІРКОВИХ СУКУПНОСТЕЙ
ДЛЯ ПРОВЕДЕННЯ ОРГАНАМИ ДЕРЖАВНОЇ СТАТИСТИКИ УКРАЇНИ
ВИБІРКОВИХ ОБСТЕЖЕНЬ НАСЕЛЕННЯ (ДОМОГОСПОДАРСТВ)**

Київ – 2018

Державна служба статистики України

Відповідальний за випуск – директор департаменту обстежень домогосподарств Осипова І. І.

У Методологічних положеннях щодо формування вибірових сукупностей для проведення органами державної статистики України вибірових обстежень населення (домогосподарств) (далі – Методологічні положення) наводяться принципи формування ймовірнісних стратифікованих вибірок, у тому числі багатоступеневих територіальних вибірок, для проведення вибірових обстежень населення (домогосподарств), зокрема умов життя домогосподарств, економічної активності населення та сільськогосподарської діяльності населення в сільській місцевості (далі – вибірові обстеження населення). У Методологічних положеннях містяться опис особливостей формування основи вибірки для стратифікованої багатоступеневої територіальної вибірки, висвітлюються загальні принципи та схема організації робіт з формування вибірок територій та домогосподарств, а також методологія актуалізації основи вибірки.

Методологічні положення підготували фахівці департаменту обстежень домогосподарств Державної служби статистики України:

Осипова І. І. – директор департаменту;

Передерій А. Б. – начальник відділу методологічного забезпечення вибірових обстежень населення;

Кольва Ю. О. – головний спеціаліст-економіст цього відділу, за участю фахівця Інституту демографії та соціальних досліджень ім. М. В. Птухи НАН України д.е.н. Саріогла В. Г.

Методологічні положення схвалені Комісією з питань удосконалення методології та звітної документації Державної служби статистики України (протокол від 16 лютого 2018 року №2).

Державна служба статистики України

- адреса: вул Шота Руставелі, 3, м. Київ, 01601, Україна
- телефони: (044) 287-24-22
- факс: (044) 235-37-39
- електронна пошта: office@ukrstat.gov.ua
- веб-сайт: www.ukrstat.gov.ua

Зміст

	Стор.
Передмова.....	4
I Загальні положення.....	5
II Визначення термінів.....	6
III Загальна схема та принципи організації робіт із формування вибірки.....	6
IV Визначення вимог до вибірки відповідно до цілей вибіркового обстеження.....	8
V Формування основи вибірки з урахуванням підсумків аналізу якості наявних інформаційних джерел.....	11
VI Розрахунок обсягу вибіркової сукупності.....	12
VII Розробка документа "Вимоги до формування вибіркових сукупностей для проведення органами державної статистики вибіркових обстежень населення (домогосподарств)".....	16
VIII Розробка плану (дизайну) вибірки	18
IX Реалізація робіт із формування вибірки.....	22
1. Відбір первинних територіальних одиниць вибірки.....	22
2. Відбір вторинних територіальних одиниць.....	30
3. Відбір домогосподарств.....	31
X Актуалізація основи вибірки.....	33
Додаток 1. Макет документа "Вимоги до формування вибіркових сукупностей для проведення органами державної статистики вибіркових обстежень населення (домогосподарств)".....	36
Додаток 2. Схема ротації домогосподарств в обстеженні економічної активності населення	37
Додаток 3. Приклад розрахунку порога саморепрезентативності.....	39
Додаток 4. Приклад розрахунку мінімального розміру ВТОВ для проведення на окремих мережах респондентів двох вибіркових обстежень населення (домогосподарств).....	40
Додаток 5. Приклад упорядкування територіальних одиниць за розміром та за принципом "географічного серпантину", а також реалізації процедури відбору територіальних одиниць.....	42
Список використаних джерел.....	45

Передмова

Метою розробки Методологічних положень є створення методологічної основи розробки плану (дизайну) вибірок для проведення органами державної статистики вибіркових обстежень населення, розробки відповідних методичних матеріалів і визначення основних етапів організації, проведення та контролю якості робіт із формування вибірок територій та респондентів.

Методологічні положення базуються на рекомендаціях міжнародних організацій, зокрема Євростату, щодо методології та організації проведення вибіркових обстежень населення (домогосподарств), а також на досвіді проведення Держстатом вибіркових обстежень населення у період 1999–2017 років.

Методологічні положення призначені для використання працівниками органів державної статистики, а також можуть бути корисними для фахівців міністерств, інших органів виконавчої влади, наукових організацій та установ, які проводять дослідження з питань методології вибіркових обстежень, інших зацікавлених користувачів.

I. Загальні положення

1. Вибіркові обстеження населення є одним з основних джерел даних для вимірювання багатьох соціально-економічних та демографічних показників, зокрема вони забезпечують отримання інформації щодо соціально-економічних характеристик домогосподарств, необхідних для макроекономічних розрахунків та розрахунків основних показників у галузі сільського господарства, проведення всебічного аналізу рівня життя населення України, моніторингу бідності, комплексної оцінки стану ринку праці у відповідності до міжнародних стандартів та потреб користувачів інформації усіх рівнів.

2. Основним етапом підготовки та проведення будь-якого вибіркового обстеження є визначення принципів формування вибіркової сукупності респондентів (далі – вибірка). Якість планування вибірки та коректність процедури формування значною мірою визначають якість результатів вибіркового обстеження, а також адекватність витрат на його проведення. Ці Методологічні положення розроблені з метою забезпечення належної якості виконання робіт із планування та формування вибірки з урахуванням наявної інформаційної бази для проведення вибірових обстежень населення.

3. Методологічні положення підготовлено з урахуванням рекомендацій міжнародних експертів: Г. Колтона (США) [9, 14], В. Верми [17], положень, викладених у документі "Методологія європейських досліджень робочої сили: Приклад проектування та реалізації" Ф. Гагніарді, В. Верма, Д. Чіпапініні [16], а також викладених у Керівництві "Обстеження бюджетів домогосподарств в ЄС. Методологія та рекомендації з гармонізації" (2003) [13], Регламенті (ЄС) № 1982/2003 від 21 жовтня 2003 року, що впроваджує Регламент (ЄС) № 1177/2003 Європейського Парламенту і Ради стосовно статистики співтовариства щодо доходів і умов життя (EU-SILC) щодо формування вибірки і правил відслідковування [18] та звіті "Обстеження робочої сили в країнах ЄС, країнах-кандидатах. Основні характеристики національних оглядових показників 2015 року" [12].

4. Методологічні положення ґрунтуються на нормах законів України "Про державну статистику" [1], "Про інформацію" [2], наказів Державного комітету статистики України від 29.12.2009 № 498 "Про затвердження структури плану статистичного спостереження та глосарія до нього", від 14.06.2010 № 216 "Про затвердження Принципів діяльності органів державної статистики", від 28.07.2015 № 180 "Про затвердження Концепції забезпечення статистичної конфіденційності", від 15.02.2017 № 41 "Про затвердження Методологічних положень щодо забезпечення статистичної конфіденційності в органах державної статистики".

II. Визначення термінів

1. У цих Методологічних положеннях терміни вживаються у значеннях, визначених глосарієм до плану статистичного спостереження [3], викладених у Статистичному словнику [4], а також у значеннях:

1) вибіркове обстеження населення (домогосподарств) – науково обґрунтований вид несущільного спостереження, при якому обстежуються не всі одиниці сукупності, а лише певним чином відібрана їх частина, а одиницею спостереження є особа чи домогосподарство;

2) вимоги до формування вибіркових сукупностей для проведення органами державної статистики вибіркових обстежень населення (домогосподарств) – основні параметри вибіркової сукупності та вимоги до її формування відповідно до мети дослідження;

3) вторинна територіальна одиниця вибірки (ВТОВ) – територіальна одиниця, яка відібрана на другому ступені формування територіальної вибірки при застосуванні багатоступеневого територіального відбору;

4) первинна територіальна одиниця вибірки (ПТОВ) – територіальна одиниця, яка відібрана на першому ступені формування територіальної вибірки при застосуванні багатоступеневого територіального відбору;

5) стратифікація – процес розподілення генеральної сукупності на однорідні групи (страти);

6) територіальна вибірка – відібрані за певною процедурою територіальні одиниці, в яких проводиться вибіркове обстеження;

7) територіальна одиниця – адміністративно-територіальна одиниця України (наприклад: міста, селища міського типу, райони, сільради тощо) або локальна територія, створена в межах адміністративно-територіальної одиниці (наприклад: переписні відділи, інструкторські дільниці, створені для проведення Всеукраїнського перепису населення; виборчі дільниці; території, що обслуговуються поштовими відділеннями, тощо).

III. Загальна схема та принципи організації робіт із формування вибірки

1. Метою формування вибірки є визначення за спеціально розробленими науково обґрунтованими процедурами частини генеральної сукупності, обстеження якої дозволяє отримати надійну інформацію щодо досліджуваних в обстеженні характеристик генеральної сукупності.

2. Загальний порядок проведення робіт із формування вибірових сукупностей населення, за умови їх функціонування в рамках системи проведення вибірових обстежень населення, складається з таких етапів:

визначення по кожному вибіровому обстеженню вимог до вибірки відповідно до цілей цього державного статистичного спостереження;

формування основи для кожного етапу вибірки з урахуванням підсумків аналізу якості наявних інформаційних джерел;

розрахунок обсягів вибірових сукупностей по кожному обстеженню;

розробка документа "Вимоги до формування вибірових сукупностей для проведення органами державної статистики вибірових обстежень населення (домогосподарств)" (додаток 1);

розробка плану (дизайну) вибірки для забезпечення функціонування системи вибірових обстежень населення;

реалізація робіт із формування вибірок територій та респондентів;

актуалізація основи вибірки.

3. Визначення вимог до вибірки, формування основи вибірки з урахуванням підсумків аналізу якості наявних інформаційних джерел, розробка плану (дизайну) вибірки, розрахунок обсягів вибірових сукупностей, організація робіт, пов'язаних із формуванням вибірки, та розрахунок імовірностей відбору здійснюються на центральному рівні. Роботи із формування основ вибірки та формування вибірових сукупностей територій здійснюються на центральному рівні, із залученням до окремих етапів робіт підрозділів територіальних органів Держстату, до функціональних повноважень яких належать питання організації вибірових обстежень населення. Роботи із актуалізації основи вибірки та формування вибірових сукупностей респондентів здійснюються на регіональному рівні.

4. Вимоги до формування вибірки, визначені окремо для кожного з вибірових обстежень населення, формалізуються у вигляді документа, який узгоджується на центральному рівні всіма зацікавленими сторонами та затверджується в установленому порядку в рамках методики щодо формування вибірових сукупностей для проведення органами державної статистики чергового циклу (наприклад п'ятирічного) вибірових обстежень населення.

Необхідність розробки такого комплексного документа зумовлена інтегрованістю вибірових обстежень населення, які на постійній основі проводять органи державної статистики. Зазначені обстеження взаємопов'язані узгодженим планом (дизайном) вибірок, спільною мережею фахівців з інтерв'ювання, узгодженістю графіків їх роботи по кожному з обстежень.

5. Створення документа "Вимоги до формування вибірових сукупностей для проведення органами державної статистики вибірових обстежень населення (домогосподарств)", як правило, має ітераційний характер, оскільки організаційні та фінансові обмеження можуть призвести до неможливості задоволення всіх вимог, висунутих на початку робіт.

6. У випадку проведення вибіркового обстеження населення поза межами системи документ "Вимоги до формування вибіркових сукупностей для проведення органами державної статистики вибіркових обстежень населення (домогосподарств)" також має бути розробленим для цього спостереження.

IV. Визначення вимог до вибірки відповідно до цілей вибіркового обстеження

1. Відповідно до цілей по кожному обстеженню мають бути визначені вимоги до вибірки за такими напрямками:

1) вимоги щодо представлення у вибірці територіальних одиниць відповідно до адміністративно-територіального устрою України.

Визначення цих вимог при плануванні вибіркових обстежень населення є обов'язковим.

У випадках обмеженого доступу до окремих населених пунктів або недоцільності з об'єктивних причин проведення в них обстеження можуть бути введені певні обмеження щодо їх уключення до вибірки. Наприклад, до вибірки вибіркових обстежень недоцільно включати населені пункти, розташовані на територіях радіологічного забруднення, віднесених до зони відчуження (І зона) та зони безумовного (обов'язкового) відселення (ІІ зона) внаслідок катастрофи на Чорнобильській АЕС;

2) вимоги до надійності результатів вибіркових обстежень населення.

При плануванні вибіркового обстеження слід визначити надійність, з якою необхідно отримати оцінки основних показників для національного рівня та рівня окремих адміністративно-територіальних одиниць. Для кількісної характеристики надійності оцінок показників використовується коефіцієнт варіації вибіркових оцінок. Для рівня країни, як правило, оцінки показників уважаються надійними при значеннях коефіцієнта варіації не більше 5%, для регіонального рівня – не більше 10%.

Від необхідного рівня надійності оцінок основних показників залежить обсяг вибіркової сукупності. При визначенні надійності слід урахувувати квадратичну залежність між коефіцієнтом варіації та обсягом вибірки. Наприклад, щоб підвищити надійність (тобто зменшити коефіцієнт варіації) утричі, обсяг вибіркової сукупності необхідно збільшити у дев'ять разів (див. розділ VI);

3) вимоги до терміну дії територіальної вибірки.

На етапі планування вибіркового обстеження визначається строк дії територіальної вибірки з урахуванням її потенціалу для забезпечення проведення визначеної кількості запланованих вибіркових обстежень у відібраних територіальних одиницях, а також особливостей організації роботи мережі фахівців з інтерв'ювання. Строк дії територіальної вибірки вказується у роках.

Строк дії територіальної вибірки враховується при визначенні мінімального розміру територіальних одиниць – кількості домогосподарств (осіб), які проживають у них;

4) початкове навантаження фахівця з інтерв'ювання.

Початкове навантаження (кількість одиниць спостереження, які підлягають обстеженню одним фахівцем з інтерв'ювання протягом визначеного на етапі планування проміжку часу) є характеристикою, яка використовується для організації роботи фахівців з інтерв'ювання та формування територіальної вибірки.

При визначенні навантаження слід враховувати всі компоненти витрат часу фахівця з інтерв'ювання, а саме:

- підготовчі роботи (самопідготовка, підготовка до інтерв'ю);
- проведення опитування в домогосподарствах та збір іншої інформації, передбаченої програмою вибіркового обстеження;
- час, необхідний, щоб дістатись до респондента (транспорт, пішки);
- перевірка, кодування (якщо це передбачено функціональними обов'язками в частині окремих вибірових обстежень) та здавання анкет до територіальних органів статистики;
- навчання тощо.

Тривалість проведення інтерв'ю (опитування) визначається в пілотному обстеженні або на підставі фактичних витрат часу (якщо обстеження проводилось раніше).

При формуванні територіальної вибірки початкове навантаження фахівця з інтерв'ювання враховується:

- при визначенні порогу саморепрезентативності;
- при визначенні мінімального розміру територіальної одиниці;

5) вимога щодо врахування очікуваного рівня участі.

Для забезпечення обстеження визначеної кількості одиниць спостереження і зменшення впливу відмов респондентів від участі в обстеженні на результати вибіркового обстеження при формуванні вибірки слід враховувати очікуваний рівень участі. Цей показник використовується для визначення кількості одиниць вибіркового обстеження, яку необхідно відібрати, для обстеження теоретично розрахованої кількості одиниць спостереження.

Рівень участі розраховується за формулою:

$$k_{rr} = c_f / c, \quad (1)$$

де k_{rr} – рівень участі;

c_f – кількість одиниць спостереження (осіб, домогосподарств), які взяли участь в обстеженні;

c – кількість одиниць спостереження, які були відібрані для участі в обстеженні.

Очікуваний рівень участі визначається за даними аналогічних попередніх вибірових обстежень (за їх наявністю), за результатами пілотних обстежень або за експертними оцінками. Очікуваний рівень участі використовується при визначенні обсягу вибіркової сукупності та мінімального розміру територіальної одиниці;

б) схема ротації одиниць вибірки.

Для підвищення якості результатів вибірових обстежень населення, які проводяться на постійній основі, застосовується спеціальна процедура – ротація одиниць вибірки, тобто їх планова заміна за спеціально визначеними схемами. Ротація одиниць вибірки забезпечує покращення якості даних за рахунок зниження навантаження на респондентів, оскільки останні перебувають в обстеженні визначеними схемами час (певну кількість періодів обстеження). Крім того, ротація респондентів забезпечує більш надійне вимірювання ступеня зміни оцінок показників. Застосування цієї процедури регламентується методологією кожного з обстежень.

У вибірових обстеженнях населення застосовуються два типи ротації одиниць вибірки: повна та часткова. Повна ротація передбачає одночасну заміну всіх одиниць вибірки, часткова – заміну частини одиниць відповідно до встановленої схеми.

Для здійснення часткової ротації вибірка сукупність розбивається на групи одиниць вибірки – ротаційні групи. Ротаційні групи відбираються як окремі незалежні підвибірки. Схема ротації одиниць вибірки регламентується методологією певного обстеження.

Приклад схеми ротації домогосподарств при проведенні у 2019–2023 роках обстеження економічної активності населення наведено в додатку 2.

Іншим прикладом може слугувати ротаційний дизайн вибірки в обстеженні доходів та умов життя, яке проводиться в країнах Європейської спільноти, а саме: вибірка на основі низки підвбірок, ідентичних за дизайном та близьких за обсягом. З року в рік одні підвбірки зберігаються, інші відкидаються і замінюються новими підвбірками. У разі ротаційного дизайну на основі чотирьох підвбірок з ротацією однієї підвбірки на рік, одна з підвбірок виключається відразу ж після першого року, друга має зберігатися протягом двох років, третя протягом трьох років, четверта повинна зберігатися протягом чотирьох років. З другого року і надалі одна нова підвбірка вводиться щороку і зберігається протягом чотирьох років [18].

Застосування часткової ротації в цілому збільшує кількість обстежених одиниць вибірки за певний період. У зв'язку з цим ротацію слід урахувати при визначенні кількості одиниць вибірки, яку необхідно відібрати на рік, а також мінімального розміру територіальної одиниці;

7) додаткові вимоги.

Додаткові вимоги до формування вибірки передбачають забезпечення інших особливостей відповідно до програм вибірових обстежень населення або

забезпечення функціонування їх системи. Ці вимоги враховуються при плануванні та реалізації етапів формування вибірки.

У випадку формування вибірки для проведення декількох обстежень може бути визначена вимога щодо обмеження участі в обстеженнях одних і тих самих одиниць вибірки. Наприклад, якщо два обстеження планується проводити в одній територіальній одиниці і визначене обмеження участі в обстеженнях одних і тих самих одиниць вибірки, то мінімальний розмір територіальної одиниці розраховується як сума мінімальних розмірів, необхідних для проведення кожного з цих обстежень.

2. У випадку формування вибірки для проведення декількох обстежень окремо визначаються вимоги до формування вибіркової сукупності кожного обстеження з подальшим їх узгодженням.

V. Формування основи вибірки з урахуванням підсумків аналізу якості наявних інформаційних джерел

1. Визначальною умовою розробки ефективного плану (дизайну) вибірки та формування основи вибірки є дослідження та вибір найбільш придатних інформаційних джерел, які повинні містити інформацію щодо територіального розміщення одиниць відбору та характеристики цих одиниць, необхідні для формування вибіркової сукупності. Інформація, яка використовується при формуванні основи вибірки, повинна задовольняти таким вимогам: повнота, якість, актуальність.

Вимога щодо повноти інформації означає, що необхідні для відбору характеристики є в наявності у визначеному джерелі для всіх без винятку одиниць генеральної сукупності.

Вимога щодо якості означає, що дані, наявні у визначеному інформаційному джерелі, відповідають реальному стану генеральної сукупності, отримані з офіційних джерел та користуються довірою фахівців.

Вимога щодо актуальності означає, що інформація, яка використовується при формуванні основи вибірки, за часом повинна бути максимально наближена до моменту проведення робіт із формування вибірки.

2. Для формування основи вибірки під час здійснення конкретного ступеня відбору повинна використовуватись інформація з одного джерела. Для формування основи багатоступеневої вибірки на різних ступенях відбору може використовуватись інформація з різних джерел.

Основа вибірки для проведення вибірових обстежень населення потенційно може формуватися на основі використання даних з різних джерел, зокрема:

переписів населення (щодо розміщення та кількості постійного населення по кожній адміністративно-територіальній одиниці);

демографічної статистики (щодо розміщення та кількості постійного населення по окремих адміністративно-територіальних одиницях);

сільськогосподарських переписів;

погосподарського обліку;

картографічних матеріалів;

матеріалів проведення виборчих кампаній (щодо розміщення та характеристики виборчих дільниць, утворених на постійній основі на території окремих населених пунктів);

поділу території країни на поштові відділення та кількості населення, що обслуговується ними;

баз даних окремих міністерств і відомств з інформацією про все населення або окремі його групи (реєстр населення, Державний реєстр фізичних осіб тощо).

3. Формування щорічних вибірових сукупностей для вибірових обстежень населення здійснюється з використанням основи вибірки, інформаційна база якої обов'язково повинна бути актуалізована безпосередньо перед початком процедур відбору.

4. На окремих етапах формування вибірової сукупності може виникати необхідність використання персоніфікованої інформації щодо одиниць генеральної сукупності. Необхідність дотримання працівниками органів державної статистики принципу конфіденційності статистичної інформації визначено статтями 17 і 21 Закону України "Про державну статистику" та статтею 21 Закону України "Про інформацію", Концепцією забезпечення статистичної конфіденційності, затвердженою наказом Державного комітету статистики України від 28.07.2015 № 180, Методологічними положеннями щодо забезпечення статистичної конфіденційності в органах державної статистики, затвердженими наказом Державного комітету статистики України від 15.02.2017 № 41.

VI. Розрахунок обсягу вибірової сукупності

1. Одним з визначальних етапів планування вибірового обстеження та формування вибірки є визначення обсягу вибірової сукупності, необхідного для задоволення встановлених вимог. Від обсягу вибірки залежать як якість результатів, так і вартість обстеження. Занадто великий обсяг вибірки призводить до невиправданих витрат та неефективного використання коштів, а недостатній обсяг вибірки – до незадовільної якості результатів обстеження і відповідно також до невиправданості витрат на проведення обстеження.

Оцінка обсягу вибірки здійснюється з урахуванням таких основних факторів [15]:

вимог щодо надійності оцінок основних показників, перелік яких має бути визначений перед початком робіт з планування вибірки. Кількість таких показників, як правило, складає від 1 до 3;

обсягу фінансових коштів, виділених для проведення обстеження;

кількості фахівців з інтерв'ювання, які будуть проводити обстеження, та їх навантаження (кількості респондентів, яких вони мають опитати протягом певного періоду).

У випадку формування вибірки для проведення декількох обстежень оцінка обсягу вибірки здійснюється окремо по кожному з них.

2. Обсяг стратифікованої багатоступеневої вибірки у випадках необхідності отримання надійних оцінок середніх значень розраховується за формулою:

$$n = deff_{\bar{y}} \times \sigma^2 \times \left(\frac{100\%}{CV_{\bar{y}} \times \bar{y}} \right)^2, \quad (2)$$

де n – обсяг вибірки;

σ^2 – дисперсія, яка характеризує варіацію значень ознаки, за якою розраховується показник, по одиницях генеральної сукупності;

\bar{y} – оцінка середнього значення (наприклад, середньомісячні сукупні ресурси домогосподарства на одну особу);

$deff_{\bar{y}}$ – дизайн-ефект при оцінюванні середнього значення \bar{y} . Дизайн-ефект є відносним показником, який кількісно характеризує вплив плану (дизайну) вибірки на надійність оцінок показників вибірових обстежень у порівнянні з вибіркою, побудованою за методом простого випадкового відбору. Для простої випадкової вибірки $deff_{\bar{y}} = 1$.

$CV_{\bar{y}}$ – коефіцієнт варіації вибірових оцінок середнього значення \bar{y} для показника, визначений у вимогах до формування вибірки, який розраховується за формулою:

$$CV_{\bar{y}} = \frac{SE_{\bar{y}}}{\bar{y}} \times 100\%, \quad (3)$$

де $SE_{\bar{y}}$ – стандартна похибка оцінювання середнього значення \bar{y} , яка розраховується на підставі величини оцінки дисперсії σ^2 та обсягу вибірки n за формулою:

$$SE_{\bar{y}} = \sqrt{\frac{\sigma^2}{n}}, \quad (4)$$

Величина σ^2 може бути оцінена за результатами вибірового обстеження за формулою:

$$\sigma^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1}, \quad (5)$$

де y_i – значення ознаки для i -го домогосподарства (наприклад, якщо y – грошовий дохід, то y_i – грошовий дохід певного (i -го) домогосподарства з вибірки).

Значення коефіцієнта варіації встановлюється відповідно до вимог щодо надійності оцінки показника. Величини дизайн-ефектів, дисперсій та оцінок

показників визначаються за даними аналогічних попередніх вибірових обстежень (за наявністю таких даних), за результатами пілотних обстежень або за експертними оцінками. Наприклад, за рекомендаціями міжнародних експертів, які надавали консультаційну допомогу органам державної статистики з питань упровадження вибірових обстежень умов життя домогосподарств та економічної активності населення, значення дизайн-ефекту при оцінюванні середніх значень доходів та витрат домогосподарств на національному рівні приймалося рівним 2,5 з урахуванням досвіду проведення аналогічних обстежень в інших країнах та наявного досвіду проведення цих обстежень в Україні.

3. У випадках необхідності отримання надійних оцінок для показників часток та сумарних значень обсяги вибірок зі складним дизайном розраховуються за формулами, аналогічними формулі (2).

При оцінюванні часток (q) формула для розрахунку обсягу вибірки набуває вигляду:

$$n = deff_q \times q \times (1 - q) \times \left(\frac{100 \%}{CV_q \times q} \right)^2 = deff_q \times \frac{(1 - q)}{q} \times \left(\frac{100 \%}{CV_q} \right)^2, \quad (6)$$

При цьому враховано що дисперсія σ^2 для часток визначається за формулою [9]:

$$\sigma^2 = q \times (1 - q), \quad (7)$$

де q – оцінка очікуваного значення частки (наприклад, частки населення із середньодушовими еквівалентними загальними доходами у місяць нижче середньорічного фактичного прожиткового мінімуму).

При оцінюванні сумарних значень формула для розрахунку обсягу вибірки набуває вигляду:

$$n = deff_Y \times \sigma^2 \times N^2 \times \left(\frac{100 \%}{CV_Y \times \hat{Y}} \right)^2, \quad (8)$$

де N – обсяг генеральної сукупності;

\hat{Y} – оцінка очікуваного сумарного значення показника (наприклад, площі землі домогосподарств). При цьому враховано, що $\hat{Y} = \bar{y} \times N$.

4. У випадках визначення у вимогах декількох показників, для яких необхідно отримати оцінки встановленого рівня надійності, обсяг вибіркової сукупності розраховується для кожного з них за формулами (2), (6) або (8). Максимальне з отриманих значень є загальним підсумковим обсягом вибірки.

5. У випадку визначення у вимогах необхідності отримання надійних оцінок показників для окремих частин вибіркової сукупності, наприклад для регіонального рівня, обсяг вибірки розраховується для кожного регіону окремо.

Сума обсягів вибірок по всіх регіонах є загальним підсумковим обсягом вибірки по Україні.

6. Обсяг вибіркової сукупності коригується, якщо визначені вимоги врахування очікуваного рівня участі.

Розрахунок обсягу вибірки (див., наприклад, формулу (2)) з урахуванням очікуваного рівня участі (k_{rr}) здійснюється за такою формулою:

$$n_c = \frac{n}{k_{rr}}, \quad (9)$$

де n_c – обсяг вибірки з урахуванням очікуваного рівня участі.

7. За умови, що обстеження проводиться на постійній основі (щомісячно, щоквартально, щорічно) і план (дизайн) вибірки не зазнає суттєвих змін, фактичний (діючий) обсяг вибірки такого обстеження може бути скоригований на основі досвіду проведення обстежень у попередніх циклах, стану явищ та процесів, що оцінюються, вимог до проведення обстежень. Наприклад, при визначенні загального обсягу вибірки на наступний рік за обсягами вибірки для регіонів розрахунки здійснюються за формулою:

$$n_{\text{вимог}} = \sum_j (n_{\text{факт},j} \times \frac{CV_{\text{факт},j}^2}{CV_{\text{вимог}}^2}), \quad (10)$$

де $n_{\text{вимог}}$ – теоретичний загальний обсяг вибірки з урахуванням вимог щодо рівня надійності цільового показника на регіональному рівні;

$n_{\text{факт},j}$ – фактичний обсяг вибірки для j -го регіону за результатами обстеження за останній рік попереднього циклу;

$CV_{\text{факт},j}$ – середнє фактичне значення коефіцієнта варіації для показника за три останні роки попереднього циклу проведення обстеження для j -го регіону (використання середнього значення коефіцієнта варіації забезпечує більшу надійність його оцінки);

$CV_{\text{вимог}}$ – значення коефіцієнта варіації, зазначене у вимогах для регіонального рівня ($CV_{\text{вимог}} = 10\%$).

8. При розрахунках загального теоретичного обсягу вибірки для проведення вибірових обстежень домогосподарств може застосовуватись процедура оптимізації дизайну вибірки з метою підвищення надійності оцінювання певних визначених показників на регіональному рівні.

Особливості здійснення процедури оптимізації визначаються окремими методичними документами для кожного вибіркового обстеження населення.

9. Для ефективного розподілу визначених обсягів вибірок за стратами (наприклад: "великі міста регіону", "малі міста регіону" та "райони регіону"), як правило, використовується формула оптимального розміщення Неймана [9]:

$$n_{j,l} = n_j \times \frac{W_{j,l} \times \sigma_{j,l}}{\sum_l W_{j,l} \times \sigma_{j,l}}, \quad (11)$$

де $n_{j,l}$ – обсяг вибірки для l -ої страти j -го регіону ($l=1,2,3$);

n_j – обсяг вибірки по певному обстеженню для j -го регіону;

$W_{j,l}$ – частка загальної кількості постійного населення регіону, що припадає на l -ту страту j -го регіону;

$\sigma_{j,l}$ – оцінка середньоквадратичного відхилення значень показника l -ої страти, j -го регіону.

Такий розподіл обсягу вибірки за стратами забезпечує максимальну надійність оцінювання показника на рівні регіону за умови, що загальний обсяг вибірки для регіону є визначеним.

10. Розрахований для встановленого рівня надійності оцінок показників обсяг вибіркової сукупності порівнюється з фінансовими та інституційними можливостями органів державної статистики щодо забезпечення проведення вибіркового обстеження. З цією метою визначається необхідна кількість фахівців з інтерв'ювання, які мають проводити польовий етап обстеження, інших працівників, які будуть забезпечувати розробку й аналіз результатів, а також розраховується необхідний обсяг фінансових витрат. У випадку недостатності можливостей та коштів для забезпечення проведення обстеження за визначених умов необхідно здійснити коригування вимог, наприклад, у напрямі зменшення надійності оцінок показників. Після цього розраховується новий обсяг вибіркової сукупності і процедура порівняння з можливостями здійснюється повторно. Таким чином, процедура визначення обсягу вибірки є ітераційною, яка завершується після узгодження всіх вимог до вибірки з фінансовими та інституційними можливостями органів державної статистики.

VII. Розробка документа "Вимоги до формування вибірових сукупностей для проведення органами державної статистики вибірових обстежень населення (домогосподарств)"

1. Документ "Вимоги до формування вибірових сукупностей для проведення органами державної статистики вибірових обстежень населення (домогосподарств)" (далі – Вимоги) узагальнює підсумки попередніх етапів визначення та узгодження всіх вимог до вибірки та впливає на подальшу організацію робіт з побудови вибіркової сукупності, проведення обстежень та поширення їх результатів на генеральну сукупність.

Вимоги до формування вибірки офіційно узгоджуються з усіма зацікавленими самостійними структурними підрозділами Держстату.

2. Документ "Вимоги" включає такі основні розділи:

1) мета побудови вибірки;

2) генеральна сукупність;

3) методологія побудови вибірки;

4) одиниці спостереження та одиниці відбору. При визначенні одиниць спостереження та одиниць відбору слід мати на увазі, що можливі випадки, коли ці одиниці не збігаються. Наприклад, при визначенні домогосподарства як одиниці відбору одиницею спостереження відповідно до мети обстеження може бути окрема особа, яка входить до складу домогосподарства;

5) вимоги щодо представлення у вибірці територіальних одиниць відповідно до адміністративно-територіального устрою України. Наприклад, необхідність забезпечення представлення у вибірці всіх регіонів України, а також за необхідністю – інших адміністративно-територіальних одиниць: обласних центрів, районів, об'єднаних місцевих територіальних громад, районів у містах тощо;

6) термін дії територіальної вибірки;

7) тип вибірки, метод та одиниці відбору. Наприклад, необхідність застосування багатоступеневої вибірки з визначенням способу відбору на кожному ступені, зокрема систематичного відбору, з імовірністю, пропорційною розміру, простого випадкового відбору та ін.;

8) кількість територіальних одиниць для відбору для проведення в них вибіркового обстеження одним фахівцем з інтерв'ювання у міських поселеннях та сільській місцевості;

9) обмеження щодо представлення у вибірці окремих груп одиниць спостереження (наприклад, в європейському обстеженні доходів та умов життя (EU SILC): "Особи, які проживають в колективних домогосподарствах і установах, виключаються з цільової сукупності" [18]);

10) перелік основних показників, надійність оцінок яких має бути забезпечена у вибіркового спостереженні;

11) навантаження фахівця з інтерв'ювання;

12) надійність оцінювання основних показників, що вимірюються у вибіркового спостереженні для національного рівня, а також за необхідністю – у

розрізі окремих соціально-економічних груп або для різних територіальних рівнів;

13) очікуваний рівень участі;

14) термін дії вибірки респондентів;

15) наявність і схема ротації одиниць спостереження;

16) обсяг вибірки респондентів з урахуванням визначених вимог (може вказуватися період дії вибірки, наприклад: рік; квартал; місяць);

17) додаткові вимоги до формування вибірки (наприклад, вибірка обстеження умов життя домогосподарств має бути оптимізована для визначених показників бідності; можуть зазначатись вікові обмеження, зокрема у обстеженні економічної активності населення опитуванню підлягають усі особи у віці 15 років і старші, в європейському обстеженні доходів та умов життя – 14 років і старші [18] тощо);

3. У випадку формування вибірки для одночасного проведення декількох вибірових обстежень вищенаведені положення наводяться для усіх, але визначаються для кожного з обстежень окремо.

4. Ураховуючи особливості проведення опитувань населення, вимоги щодо початкового навантаження одного фахівця з інтерв'ювання та очікуваного рівня участі респондентів визначаються в розрізі типів місцевості проживання домогосподарств.

5. У разі проведення фахівцем з інтерв'ювання декількох вибірових обстежень одночасно, але на різних мережах респондентів, під час узгодження вимог початкове навантаження коригується з урахуванням законодавчо встановленої тривалості робочого дня.

6. Макет документа "Вимоги до формування вибірових сукупностей для проведення органами державної статистики вибірових обстежень населення (домогосподарств)" наведений у додатку 1.

VIII. Розробка плану (дизайну) вибірки

1. Після затвердження Вимог здійснюється розробка плану (дизайну) вибірки – організаційно-логічної моделі структури вибірової сукупності та принципів її формування. На цьому етапі також опрацьовуються зміст і послідовність усіх процедур формування вибірки. Результатом цієї роботи має

бути методика формування вибірових сукупностей для проведення вибірових обстежень населення (домогосподарств) протягом певного календарного циклу.

2. При розробці плану (дизайну) вибірки необхідно враховувати обмеження та умови, які визначаються Вимогами.

Разом з тим план (дизайн) вибірки має враховувати не тільки принципи формування вибірки, а й особливості інших етапів проведення вибірового обстеження, зокрема вимоги, обумовлені діючим програмним забезпеченням з обробки даних, наприклад, необхідність парної кількості територіальних одиниць для аналізу надійності оцінок показників у програмних пакетах, які реалізують реплікаційні алгоритми.

Нижче наведені основні методологічні підходи до розробки плану (дизайну) територіальних вибірок для вибірових обстежень населення:

- визначення обмежень щодо охоплення вибіровим обстеженням адміністративно-територіальних одиниць України;
- стратифікація генеральної сукупності;
- визначення типу вибірки і способів відбору на кожному етапі.

3. Визначення обмежень щодо охоплення вибіровим обстеженням адміністративно-територіальних одиниць України.

На початку формування територіальної вибірки визначаються адміністративно-територіальні одиниці України, які не повинні або не можуть бути обстежені з об'єктивних причин. Обмеження щодо територіального охоплення відображаються у Вимогах та враховуються при реалізації процедур відбору й обробки результатів обстеження.

4. Стратифікація генеральної сукупності.

Формування вибірових сукупностей для проведення вибірових обстежень населення передбачає застосування на початковому етапі відбору стратифікації генеральної сукупності. Стратифікація здійснюється з метою адекватного відображення у вибірці основних особливостей адміністративно-територіального поділу країни, а також для забезпечення відбору домогосподарств із більш однорідних за основними характеристиками груп [3, 4, 8, 14, 15].

Для реалізації цієї мети генеральна сукупність розподіляється на страти.

Наприклад, в межах кожного регіону страти можуть відповідати:

- містам і міськрадам із кількістю населення 100 тис. осіб і більше (далі страта "великі міста регіону");

- містам, селищам міського типу, міськрадам і селищним радам із кількістю населення менше 100 тис. осіб (далі страта "малі міста регіону");

- адміністративним районам сільської місцевості (далі страти "райони регіону").

Такий розподіл забезпечує потреби щодо отримання інформації за типами місцевості (міська, сільська), а також надає можливість урахувати різні умови роботи фахівців з інтерв'ювання у різних за розміром та за типами населених пунктах (різні витрати часу на проїзд, різне навантаження), різні інформаційні

джерела, необхідні для відбору територіальних одиниць та одиниць спостереження тощо.

Обсяг вибірки розподіляється між стратами пропорційно їх розміру, наприклад кількості населення. Такий підхід забезпечує самозваженість вибірки (однакову частку відбору по стратах), що зумовлює ефективне вимірювання значної кількості показників вибіркових обстежень населення.

5. Визначення типу вибірки:

1) у відповідності до встановлених вимог до вибірки визначаються особливості її формування: основа вибірки, кількість ступенів відбору, способи відбору на кожному ступені та ін.

За типом вибірка може бути одноступеневою або багатоступеневою.

Формування одноступеневої вибірки здійснюється в один етап – шляхом безпосереднього відбору одиниць генеральної сукупності. При вибіркових обстеженнях населення, які проводяться шляхом відвідування домогосподарств, як правило, недоцільно використовувати одноступеневі вибірки. Це пов'язано з тим, що такий тип вибірки:

потребує основи вибірки, яка містила б перелік усіх одиниць генеральної сукупності (наприклад, повний перелік домогосподарств (населення) із зазначенням їх адрес);

призводить до значного розшарування одиниць спостереження по території, що обстежується, та збільшення витрат часу та фінансових ресурсів, необхідних для проведення обстеження.

У випадку застосування стратифікації генеральної сукупності формування вибірки проводиться всередині кожної страти.

Формування багатоступеневої вибірки здійснюється у кілька послідовних етапів, кількість яких залежить від наявної інформаційної бази, організаційних можливостей органів державної статистики та трудомісткості робіт із формування вибірки. На різних ступенях відбору можуть визначатися різні одиниці відбору, а також використовуватись різні інформаційні джерела;

2) як одиниці відбору можуть визначатися:

об'єкти адміністративно-територіального устрою країни (наприклад, області, міста, селища міського типу, райони, міські, селищні та сільські ради, об'єднані територіальні громади, окремі села);

локальні території, визначені за певним принципом (наприклад, виборчі дільниці; переписні відділи, інструкторські та лічильні дільниці, створені для проведення перепису населення; території, які обслуговуються одним поштовим відділенням, тощо);

домогосподарства;

особи.

Визначення одиниці відбору на конкретному етапі залежить від наявності інформації, необхідної для проведення відбору. У зв'язку з цим на початку робіт слід визначити тип локальних територій, по яких можливо отримати основу

вибірки для проведення одного чи декількох етапів відбору, а на останньому етапі – провести безпосередньо відбір одиниць спостереження.

При розробці етапів формування вибірки визначають мінімально необхідну кількість ступенів відбору, що сприяє:

скороченню обсягів робіт щодо формування вибіркової сукупності;

зменшенню похибки вибірки, оскільки кожен додатковий етап відбору, як правило, впливає на похибку вибірки у бік її збільшення.

На вибір типу вибірки також впливає організаційна структура мережі фахівців з інтерв'ювання. Так, особливістю організації мережі фахівців з інтерв'ювання в органах державної статистики є їх проживання у межах тих міст, у яких вони проводять вибіркові обстеження населення;

3) результати аналізу наявних інформаційних джерел, вимог до формування вибірок для проведення вибіркових обстежень населення та організації роботи мережі фахівців з інтерв'ювання зумовили необхідність застосування багатоступеневої вибірки. Наприклад, в обстеженнях умов життя домогосподарств і економічної активності населення вибірка є тріступеневою, а в обстеженні сільськогосподарської діяльності населення – двіступеневою. Формування тріступеневої вибірки включає такі основні етапи:

відбір одиниць першого ступеня – адміністративно-територіальних одиниць (за прийнятою термінологією – первинних територіальних одиниць вибірки, скорочено – ПТОВ);

відбір одиниць другого ступеня – локальних територій (за прийнятою термінологією – вторинних територіальних одиниць вибірки, скорочено – ВТОВ);

відбір одиниць третього ступеня – домогосподарств (або осіб).

Метою відбору ПТОВ є отримання переліку адміністративно-територіальних одиниць (міських (селищних) рад, міст, селищ міського типу, сільських населених пунктів), у яких будуть проводитись вибіркові обстеження населення. Метою відбору ВТОВ є отримання переліку локальних територій у межах ПТОВ (інструкторських ділянок у міській місцевості тощо), у яких на третьому етапі будуть відбиратися одиниці третього ступеня.

Доцільність формування та відбору ВТОВ зумовлена необхідністю оптимізації роботи фахівців з інтерв'ювання шляхом формування та відбору в межах одиниць першого ступеня, які мають велику за розміром площу, менших локальних територій, по яких існує необхідна для відбору інформація щодо населення;

4) на різних ступенях відбору, відповідно до визначених одиниць відбору та наявних інформаційних джерел, можуть бути різні основи вибірки. Наприклад, при відборі ПТОВ і ВТОВ основу вибірки можуть становити переліки відповідних територіальних одиниць та їх окремі ознаки (наприклад, код відповідно до Класифікатора одиниць адміністративно-територіального устрою України (КОАТУУ) та кількість населення або кількість домогосподарств у цих одиницях); при відборі домогосподарств основою вибірки є повний список їх адрес по відповідній територіальній одиниці;

5) в окремих випадках, коли деякі територіальні одиниці складаються з певної кількості населених пунктів, припустимим є поділ таких територіальних одиниць на окремі частини (групи населених пунктів) з подальшим їх послідовним обстеженням. Ця процедура фактично, є додатковим ступенем відбору, який реалізується в межах окремих територіальних одиниць. Наприклад, сільські/селищні ради або об'єднані територіальні громади можуть включати значну кількість сільських населених пунктів, транспортне сполучення між якими ускладнене. У цьому випадку з метою забезпечення ефективної роботи фахівця з інтерв'ювання, формуються групи сільських населених пунктів з урахуванням їх територіального розташування. Після цього визначається послідовність (випадковим чином) обстеження груп протягом строку дії територіальної вибірки. Відбір домогосподарств здійснюється в межах кожної сформованої групи сільських населених пунктів.

6. Визначення методів відбору.

На різних ступенях відбору, відповідно до цілей вибіркового обстеження та наявних інформаційних джерел, можуть бути різні методи відбору. Дизайн вибірки і методи формування вибірки повинні забезпечувати, щоб кожній особі (домогосподарству) у цільовій сукупності присвоювалась відома ненульова ймовірність вибору [18].

Основні методи здійснення відбору такі:

простий випадковий;

систематичний;

систематичний з імовірністю, пропорційною розміру.

Одиниці вибірки можуть мати однакові ймовірності відбору або різні.

Для вибірових обстежень одиниці вибірки територій мають різні ймовірності відбору, які є пропорційними розміру територіальної одиниці. Такі ймовірності є результатом застосування відбору з імовірністю, пропорційною розміру (probability proportional to size, скорочено – PPS). Як розмір може виступати кількість населення або кількість домогосподарств територіальної одиниці, розмір площі землі домогосподарств тощо.

Основна перевага методу відбору з імовірністю, пропорційною розміру, полягає в тому, що для багатоступеневих вибірок він дозволяє забезпечити однакову загальну ймовірність відбору одиниць вибірки останнього ступеня за умови їх рівномірного розподілу по територіальних одиницях різного розміру.

На окремих етапах формування вибірки після упорядкування одиниць відбору за певним визначеним принципом безпосередньо відбір здійснюється із застосуванням систематичного відбору.

IX. Реалізація робіт із формування вибірки

1. Відбір первинних територіальних одиниць вибірки

1. Реалізація робіт із формування вибірки здійснюється у кілька послідовних етапів із застосуванням процедур, визначених при розробці плану

(дизайну) вибірки. Зокрема, формування вибірки для вибіркових обстежень може здійснюватися у три етапи, а саме: відбір первинних територіальних одиниць вибірки, відбір вторинних територіальних одиниць вибірки, відбір домогосподарств.

2. Етап відбору первинних територіальних одиниць вибірки складається з таких робіт: визначення порога саморепрезентативності територіальних одиниць, визначення мінімального розміру територіальної одиниці, відбір несаморепрезентативних територіальних одиниць, розрахунок імовірностей відбору первинних територіальних одиниць:

1) визначення порогу саморепрезентативності.

При проведенні вибіркового обстеження, з одного боку, кожна одиниця вибіркової сукупності репрезентує певну кількість елементів генеральної сукупності. З іншого боку, фахівець з інтерв'ювання має певне навантаження щодо кількості одиниць вибіркової сукупності, яку він повинен обстежити. Таким чином, визначена кількість одиниць вибіркової сукупності, яку опитує один фахівець з інтерв'ювання, репрезентує визначену кількість елементів генеральної сукупності. Ця кількість елементів генеральної сукупності називається порогом саморепрезентативності, який у разі застосування методу відбору PPS є критерієм визначення територіальних одиниць, що обов'язково мають бути відібрані до територіальної вибірки, тобто включені до територіальної вибірки з імовірністю, що дорівнює одиниці. Такі територіальні одиниці називаються саморепрезентативними. Несаморепрезентативні територіальні одиниці відбираються із застосуванням методу систематичного відбору з імовірністю, пропорційною розміру, та включаються до територіальної вибірки з імовірністю, меншою одиниці.

При формуванні територіальної вибірки для проведення вибіркових обстежень населення кількість територіальних одиниць, яку необхідно відібрати, невідома. У зв'язку з цим поріг саморепрезентативності визначають на основі обсягу вибірки n та встановленої кількості домогосподарств, що обстежуються одним фахівцем з інтерв'ювання (навантаження фахівця з інтерв'ювання), за умови повного завантаження одного фахівця з інтерв'ювання при обстеженні цієї територіальної одиниці. Поріг саморепрезентативності визначається за такою формулою:

$$n_p^{HH} = \frac{N}{n} \times m, \quad (12)$$

де n_p^{HH} – поріг саморепрезентативності (кількість домогосподарств);

N – обсяг генеральної сукупності;

n – обсяг вибіркової сукупності;

m – початкове навантаження фахівця з інтерв'ювання (кількість домогосподарств).

Обсяг генеральної сукупності, обсяг вибірки та навантаження фахівця з інтерв'ювання можуть визначатися як кількістю домогосподарств, так і кількістю

населення. Розмір територіальних одиниць (міст, смт) часто визначається в особах. У цьому випадку для розрахунку порога саморепрезентативності (кількість населення) у формулу 12 слід увести додатковий множник – середній розмір домогосподарства:

$$n_p^{pers} = \frac{N}{n} \times m \times s, \quad (13)$$

де n_p^{pers} – поріг саморепрезентативності (кількість населення (осіб));

s – середній розмір домогосподарства (кількість осіб).

Середній розмір домогосподарства у міських населених пунктах може бути отриманий з підсумків перепису населення або з даних попередніх вибіркового обстежень, а по сільській місцевості – з вищезгаданих джерел, а також із даних форми погосподарського обліку № 1 "Облікова картка об'єкта погосподарського обліку". Доцільність застосування у розрахунках порога саморепрезентативності показника кількості населення пояснюється наявністю даних демографічної статистики щодо кількості населення практично по всіх адміністративно-територіальних одиницях, що дозволяє застосовувати останні для вдосконалення процедур розповсюдження результатів вибіркового обстеження на генеральну сукупність.

Приклад розрахунку порога саморепрезентативності наведено у додатку 3.

У випадку використання однієї територіальної вибірки для проведення декількох вибіркового обстежень розрахунок порогу саморепрезентативності слід проводити для вибіркового обстеження з найбільшим навантаженням фахівця з інтерв'ювання;

2) визначення мінімального розміру територіальної одиниці.

За наявністю вимоги щодо формування вибіркового сукупностей для проведення вибіркового обстеження населення на певній територіальній одиниці протягом кількох років слід визначити мінімальний розмір такої територіальної одиниці.

Мінімальний розмір територіальної одиниці може бути розрахований за двома підходами в залежності від схеми ротації одиниць вибірки: повної щорічної ротації або часткової.

У випадку повної щорічної ротації одиниць вибірки розрахунок базується на навантаженні фахівця з інтерв'ювання і проводиться за формулою:

$$n_{min} = m \times t, \quad (14)$$

де n_{min} – мінімальний розмір територіальної одиниці (кількість домогосподарств або осіб);

m – початкове навантаження фахівця з інтерв'ювання;

t – кількість років перебування територіальної одиниці у вибірці.

У випадку часткової ротації одиниць вибірки мінімальний розмір територіальної одиниці розраховується з урахуванням установленної схеми ротації. За умови використання у вибіркового обстеженні часткової ротації одиниць

вибірки необхідно виділяти два періоди: впровадження схеми ротації (зазвичай перший рік) та підтримка схеми ротації (наступні роки). Це пояснюється тим, що у період упровадження схеми ротації необхідно відібрати більше одиниць вибірки, ніж у наступні періоди.

Мінімальний розмір територіальної одиниці з урахуванням схеми ротації розраховується за формулою:

$$n_{min} = n_1 + n_t, \quad (15)$$

де n_1 – кількість одиниць вибірки (домогосподарств або осіб), необхідних у перший рік запровадження схеми ротації, який розраховується за формулою:

$$n_1 = (q_1 \times z), \quad (16)$$

де q_1 – кількість ротаційних груп (підгруп) у перший рік запровадження схеми ротації;

z – розмір ротаційних груп (підгруп) (домогосподарств або осіб).

Кількість одиниць вибірки (домогосподарств або осіб), необхідних у наступні роки функціонування схеми ротації (n_t), розраховується за формулою:

$$n_t = (q \times z) \times (t - 1), \quad (17)$$

де q – кількість ротаційних груп (підгруп) у наступні роки функціонування схеми ротації протягом певного визначеного циклу (наприклад п'ятирічного).

Для забезпечення надійності функціонування територіальної вибірки протягом визначеної кількості років доцільно збільшити мінімальний розмір територіальної одиниці шляхом додавання резервного року. При повній щорічній ротації одиниць спостереження використовується формула:

$$n_{min} = m \times (t + 1), \quad (18)$$

а при розрахунках мінімального розміру територіальної одиниці з урахуванням схеми ротації – наступна формула:

$$n_{min} = (q_1 \times z) + (q \times z) \times t. \quad (19)$$

Наведена процедура унеможливить негативний вплив змін, які можуть відбутися протягом терміну експлуатації територіальної вибірки, а також дозволить забезпечити більш тривалий термін її експлуатації у випадку можливої затримки введення нової територіальної вибірки.

Приклад розрахунку мінімального розміру територіальної одиниці для проведення на окремих мережах респондентів двох вибірових обстежень населення визначений за двома підходами (в залежності від схеми ротації одиниць вибірки: повної щорічної ротації або часткової), з урахуванням резервного року, наведено в додатку 4.

Якщо у вимогах до формування вибірки був визначений очікуваний рівень участі (k_{rr}), мінімальний розмір територіальної одиниці коригується за формулою:

$$n'_{\min} = \frac{n_{\min}}{k_{rr}}. \quad (20)$$

3) за умови, що обстеження проводиться на постійній основі і план (дизайн) вибірки не зазнає суттєвих змін, розрахований мінімальний розмір територіальної одиниці доцільно скоригувати також на частку домогосподарств, непридатних до опитування.

Коефіцієнт для коригування на частку домогосподарств непридатних до опитування, розраховується як сума середніх значень часток домогосподарств, непридатних для опитування, визначених на основі зведеної інформації щодо кількості зачинених, нежитлових приміщень та приміщень, у яких немає осіб, які проживають постійно, за результатами актуалізації списків адрес домогосподарств у попередньому циклі проведення обстежень;

4) у випадку проведення в одній територіальній одиниці декількох вибірових обстежень мінімальний розмір територіальної одиниці визначається сумою мінімальних розмірів, розрахованих для кожного з цих обстежень.

Територіальні одиниці, відібрані на першому ступені відбору, як правило, мають достатній розмір для проведення вибіркового обстеження протягом кількох років. Проблема визначення мінімального розміру територіальної одиниці більш актуальна для локальних територій, які відбираються на другому ступені відбору (ВТОВ);

5) відбір несаморепрезентативних територіальних одиниць.

У випадках необхідності відбору територіальних одиниць, розмір яких менше порога саморепрезентативності, такі територіальні одиниці визначаються як несаморепрезентативні й уключаються до вибірки з імовірністю, пропорційною їх розміру.

Несаморепрезентативні територіальні одиниці відбираються з дотриманням такої вимоги: одна така одиниця репрезентує групу територіальних одиниць, сума розмірів яких дорівнює порогу саморепрезентативності.

Процедура відбору несаморепрезентативних територіальних одиниць складається з таких основних етапів:

- упорядкування несаморепрезентативних територіальних одиниць;
- визначення кількості територіальних одиниць, які необхідно відібрати;
- визначення кроку відбору та першої територіальної одиниці для відбору;
- відбір наступних територіальних одиниць.

Нижче наведено зміст зазначених етапів.

Упорядкування несаморепрезентативних територіальних одиниць.

У межах кожної страти регіону складається впорядкований певним чином перелік несаморепрезентативних територіальних одиниць. Нижче наведені можливі принципи впорядкування:

- за значенням певної ознаки (характеристики) територіальної одиниці;

за географічним принципом (принцип "географічного серпантину");
за абеткою відповідно до назви територіальної одиниці;
за іншим принципом.

Останні два варіанти можуть бути застосовані при формуванні вибірки методом простого випадкового відбору.

При формуванні територіальної вибірки для проведення вибірових обстежень населення найчастіше застосовується метод систематичного відбору з імовірністю, пропорційною розміру, з попередньою стратифікацією генеральної сукупності, що вимагає впорядкування територіальних одиниць за певним принципом.

Упорядкування за ознакою (наприклад, за кількістю населення) передбачає наявність помітних розбіжностей явищ, що досліджуються, в одиницях різного розміру. Метою цього типу впорядкування є включення до вибірки як одиниць великого, так і одиниць малого розміру. Разом з тим упорядкування за принципом "географічного серпантину" (див. рис. 1) задовольняє вимозі рівномірного з географічної точки зору розміщення по території одиниць вибірки.

Беручи до уваги вищенаведене, при формуванні вибірових сукупностей для проведення вибірових обстежень населення для етапу відбору несаморепрезентативних ПТОВ у межах кожної страти слід скласти перелік територіальних одиниць, ранжований у порядку зменшення кількості населення (або в порядку збільшення кількості населення). Райони слід упорядковувати не за ознакою (наприклад, за кількістю населення), а за принципом "географічного серпантину". Такий підхід забезпечить краще охоплення території регіону (області), а також дозволить урахувати природно-кліматичні особливості різних частин регіону, що важливо, наприклад, для вивчення виробництва сільськогосподарської продукції в домогосподарствах.

Застосування принципу "географічного серпантину" обов'язково передбачає використання картографічного матеріалу. За відсутності карт можуть використовуватися карти-схеми, але на них повинно витримуватися географічне розміщення об'єктів відносно один одного. Важливою умовою є необхідність однакової географічної орієнтації всіх карт у межах області та країни. Ураховуючи, що всі карти, які виробляються спеціалізованими підприємствами, мають орієнтацію "північ-південь", аналогічну орієнтацію слід використовувати при проведенні серпантину. Серпантин слід проводити з лівого верхнього кута на схід, після досягнення краю карти необхідно робити поворот, потім вести серпантин зі сходу на захід паралельно попередній лінії. Обов'язковою вимогою при проведенні серпантину є відсутність "порожніх проходів", тобто при одному проході лінія має пройти щонайменше через одну територіальну одиницю. Територіальні одиниці за порядком проходження лінії включаються до списку.

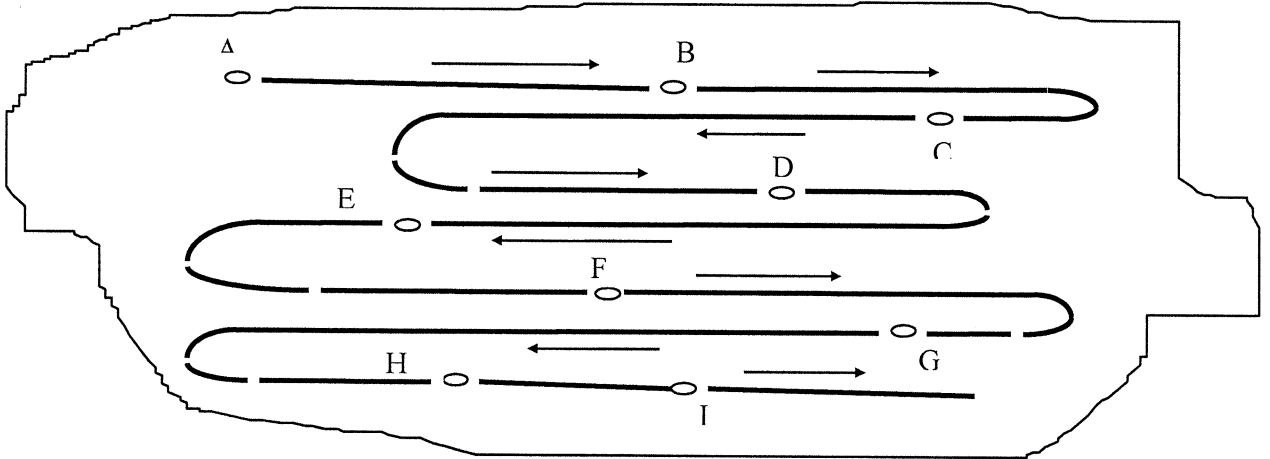


Рис. 1. Впорядкування територіальних одиниць вибірки за принципом "географічного серпантину"

Приклад упорядкування територіальних одиниць за розміром та за принципом "географічного серпантину" наведено у додатку 5.

3. Визначення кількості територіальних одиниць, які необхідно відібрати.

Кількість несаморепрезентативних ПТОВ, які необхідно відібрати у кожній страті, визначається діленням загальної кількості населення всіх несаморепрезентативних територіальних одиниць кожної страти на поріг саморепрезентативності, який визначається за формулою (12) або за формулою (13). Результатом практично в усіх випадках є дробове число, яке необхідно округлити до цілого числа. Кількість несаморепрезентативних територіальних одиниць, яку необхідно відібрати, розраховується за формулою:

$$d = N_{sstr} / n_p^{HH} = N_{sstr}^{pers} / n_p^{pers}, \quad (21)$$

де d – кількість несаморепрезентативних територіальних одиниць, яку необхідно відібрати;

N_{sstr} – кількість домогосподарств у конкретній страті;

N_{sstr}^{pers} – кількість населення у конкретній страті.

Після округлення кількості несаморепрезентативних територіальних одиниць до цілого значення визначається величина інтервалу (крок) відбору;

4. Визначення кроку відбору та першої одиниці для відбору.

Крок відбору – це результат ділення загального значення ознаки, визначеної як основа для ранжування (наприклад, кількості населення) у несаморепрезентативних територіальних одиницях області (району), на число одиниць, які необхідно відібрати.

Якщо несаморепрезентативні територіальні одиниці відранжовані за кількістю населення, крок відбору h розраховується за формулою:

$$h = N_{sstr}^{pers} / d; \quad (22)$$

5. Відбір першої територіальної одиниці вибірки.

Перед початком відбору до переліку несаморепрезентативних територіальних одиниць вносяться дані щодо накопичених значень ознаки, визначеної як основа для ранжування (кількості населення відповідних територіальних одиниць).

Початок відбору (перша відібрана несаморепрезентативна територіальна одиниця ($h^{(1)}$)) визначається шляхом множення кроку відбору на випадкове число:

$$h^{(1)} = h \times rnd, \quad (23)$$

де rnd – випадкове число, визначене для страти (випадкове число, яке набуває значення від 0 до 1 і може бути визначене за допомогою генератора випадкових чисел).

Отриманий результат розрахунку порівнюється з накопиченими значеннями ознаки (наприклад, кількостями населення або розмірів площ землі у переліку територіальних одиниць). Територіальна одиниця, накопиченому значенню ознаки якої відповідає результат, є першою одиницею вибірки;

6. Відбір наступних територіальних одиниць вибірки.

Усі наступні територіальні одиниці визначаються шляхом додавання кроку відбору до накопиченого значення ознаки, розрахованого для відбору попередньої територіальної одиниці:

$$h^{(z)} = h^{(z-1)} + h, \quad z = 1, 2, \dots, d, \quad (24)$$

де d – кількість несаморепрезентативних територіальних одиниць, яку необхідно відібрати;

$h^{(z)}$ – територіальна одиниця вибірки (крім першої).

Приклад розрахунку кроку відбору, визначення першої та наступних одиниць для відбору наведено у додатку 5.

Результати відбору з використанням двох принципів упорядкування територіальних одиниць свідчать, що територіальні одиниці, розмір яких близький до кроку відбору, мають високу ймовірність бути відібраними за кожним з цих принципів упорядкування;

7. Розрахунок імовірностей відбору ПТОВ.

Після відбору ПТОВ розраховується ймовірність їх відбору.

У зв'язку з тим, що при формуванні територіальної вибірки для проведення вибірових обстежень населення найчастіше застосовується метод відбору з імовірностями, пропорційними розміру, нижче наведено методологічні основи до розрахунків імовірностей відбору на прикладі саме цього методу для випадку, якщо характеристикою розміру територіальної одиниці є кількість домогосподарств. У всіх територіальних одиницях проживають в цілому N домогосподарств, з яких необхідно відібрати n домогосподарств, з дотриманням

однакової ймовірності відбору ($f = n / N$) для кожного домогосподарства. Це один з основних принципів, який покладено в основу методу відбору *PPS*.

Наприклад, всі домогосподарства проживають в D територіальних одиницях, з яких треба відібрати d одиниць. Відповідно кількість домогосподарств, які мають бути відібрані в кожній територіальній одиниці вибірки, дорівнює в середньому $b = n / d$.

Якщо M_α є розміром територіальної одиниці α (у цьому випадку – кількістю домогосподарств у територіальній одиниці) і $M = \sum M_\alpha$ є загальним розміром усіх територіальних одиниць D у генеральній сукупності, то ймовірність відбору територіальної одиниці α , пропорційна її розміру, становитиме [11]:

$$P_\alpha = \left(d \times \frac{M_\alpha}{\sum M_\alpha} \right), \quad (25)$$

де P_α – ймовірність відбору первинної територіальної одиниці.

Наявність коефіцієнта d у формулі (25) означає, що територіальні одиниці з розміром M_α , рівним чи більшим певного порогового значення, рівного $\sum M_\alpha / d$, мають ймовірність відбору, рівну чи більшу одиниці. Відповідно ці територіальні одиниці відбираються з 100% визначеністю, тобто з ймовірністю, що дорівнює одиниці.

В окремих випадках на ймовірність відбору мають вплив специфічні вимоги до формування територіальної вибірки. Наприклад, однією з вимог до формування територіальної вибірки для проведення вибіркових обстежень населення може бути проведення обстеження сільськогосподарської діяльності населення в сільській місцевості в усіх адміністративних районах сільської місцевості України. У цьому випадку останні включаються до вибірки з ймовірністю, яка дорівнює одиниці.

2. Відбір вторинних територіальних одиниць вибірки

1. Формування ВТОВ.

Для формування ВТОВ необхідно мати картографічний матеріал (карти, карти-схеми, плани), який повинен містити географічне відображення меж утворень локальних територіальних одиниць, на основі яких формуються ВТОВ. Розмір кожної ВТОВ повинен бути не меншим розрахованого мінімального розміру територіальної одиниці відбору n_{min} . Якщо розмір локальної території більший ніж n_{min} , то вона визначається як ВТОВ; якщо розмір менший ніж n_{min} , то необхідно здійснити об'єднання локальних територій, які географічно знаходяться поруч, поки не буде забезпечено мінімально необхідний розмір (кількість населення (домогосподарств)).

2. Упорядкування ВТОВ.

Упорядкування ВТОВ при проведенні відбору доцільно здійснювати за принципом "географічного серпантину". Застосування цього принципу дозволяє враховувати особливості одиниць спостереження відповідно до їх географічного розміщення. Наприклад, у містах існують райони, які відповідно до специфіки заселення та подальшого формування в побуті мають назви "робочі", "студентські", "привілейовані" тощо. У сільській місцевості, де як ВТОВ виступають сільські населені пункти, а особливості господарювання та розташування господарств населення залежать від природно-кліматичних умов, наявності транспортної інфраструктури тощо, принцип "географічного серпантину" дозволяє краще, ніж упорядкування територіальних одиниць відповідно до кількості населення, врахувати особливості, притаманні умовам життя сільського населення.

3. Відбір ВТОВ.

Відбір ВТОВ здійснюється з використанням аналогічних процедур, які застосовуються при відборі ПТОВ (див. додаток 5).

Розрахунок імовірностей відбору територіальних одиниць другого ступеня. Імовірність відбору ВТОВ (P_{β}) розраховується за формулою:

$$P_{\beta} = \left(b \times \frac{M_{\beta}}{\sum_{\beta} M_{\beta}} \right), \quad (26)$$

де b – кількість ВТОВ, які необхідно відібрати;

M_{β} – кількість домогосподарств у ВТОВ β ;

$\sum_{\beta} M_{\beta}$ – загальна кількість домогосподарств у всіх ВТОВ.

3. Відбір домогосподарств

1. Упорядкування домогосподарств.

Упорядкування домогосподарств здійснюється у межах відібраних ВТОВ. Вибір способу впорядкування залежить від наявної інформаційної бази і має особливості у міських поселеннях та в сільській місцевості. Метод упорядкування повинен бути однаковим у всіх ВТОВ у межах певної страти.

У міських поселеннях, незалежно від джерела інформації щодо адрес домогосподарств (матеріали перепису населення, дані виборчих діляниць, дані поштових відділень тощо), упорядкування здійснюється у такому порядку: спочатку впорядковуються за абеткою назви вулиць, потім – у порядку зростання номерів будинків, потім – у порядку зростання номерів квартир.

У сільській місцевості найбільш доцільним для використання джерелом інформації щодо домогосподарств є форма погосподарського обліку № 1 "Облікова картка об'єкта погосподарського обліку". Для зменшення обсягу робіт із формування вибірки списки адрес домогосподарств доцільно скласти у

порядку зростання номерів об'єктів погосподарського обліку, зазначених у формі погосподарського обліку № 1 "Облікова картка об'єкта погосподарського обліку".

Для формування вибірових сукупностей можуть бути використані й інші джерела інформації щодо адрес домогосподарств. За умови наявності декількох інформаційних джерел доцільно серед них визначити для використання більш повні за обсягом, а також більш актуальні за часом створення.

Перед початком відбору домогосподарств обов'язково здійснюється актуалізація складених списків адрес, яка унеможливорює або мінімізує негативний вплив можливих змін, які відбуваються як у межах ВТОВ (наприклад, знесення будинків, стихійне лихо тощо), так і безпосередньо серед домогосподарств.

Домогосподарствам у впорядкованих списках по кожній ВТОВ зазначаються порядкові номери, починаючи з одиниці.

2. Проведення відбору домогосподарств.

Відбір домогосподарств проводиться способом систематичного відбору номерів домогосподарств з упорядкованих списків адрес. Приклад проведення процедури систематичного відбору наведено при розгляді відбору несаморепрезентативних територіальних одиниць (міст) (додаток 5).

У зв'язку із застосуванням генератора випадкових чисел результати розрахунків кроку відбору, першого та наступних номерів відібраних домогосподарств становлять дробові числа, які слід округлити за правилами арифметики до певної кількості цифр після коми. Ця кількість цифр залежить від кількості домогосподарств, яку необхідно відібрати, і визначається умовою забезпечення всім домогосподарствам з основи вибірки однакової можливості бути відібраними. Для здійснення процедури відбору слід визначення номерів відібраних домогосподарств здійснювати на основі цілої частини згаданого дробового числа незалежно від значення дробової частини (тобто без округлення). Наприклад, при округленні до двох знаків після коми, незалежно від того, становить розрахунковий номер домогосподарства для відбору 74,00; 74,18; 74,50; 74,61 або 74,99, до вибірки включається домогосподарство, яке у списку відповідає номеру 74. Цей підхід унеможливорює при відборі одиниць вибірки вихід за межі списку домогосподарств у випадку проведення округлення за правилами арифметики.

3. Розрахунок імовірностей відбору одиниць третього ступеня.

Останнім етапом формування вибіркової сукупності є розрахунок імовірностей відбору та базових ваг для кожної відібраної одиниці спостереження. Імовірність відбору характеризує шанс конкретного домогосподарства бути відібраним.

Імовірність відбору домогосподарств у ВТОВ розраховується за формулою:

$$P_c = \left(\frac{c_\beta}{M_\beta} \right), \quad (27)$$

де c_β – кількість домогосподарств, які необхідно відібрати у ВТОВ β .

Розрахунок загальної ймовірності відбору для методу триступеневого відбору здійснюється за такою формулою:

$$\left(d \times \frac{M_\alpha}{\sum_\alpha M_\alpha} \right) \times \left(b \times \frac{M_\beta}{\sum_\beta M_\beta} \right) \times \left(\frac{c_\beta}{M_\beta} \right) = n/N = f, \quad (28)$$

де f – частка відбору [11].

Якщо на першому та другому ступенях відбору застосовуються однакові одиниці вимірювання розміру, в ідеальному випадку повинна виконуватися умова $M_\alpha = \sum_\beta M_\beta$, однак ця умова не є необхідною.

Задовільною є умова $M_\alpha \approx \sum_\beta M_\beta$.

Загальна теоретична ймовірність для триступеневого відбору певного i -го домогосподарства (P_i) з вибіркової сукупності розраховується за формулою:

$$P_i = P_a \times P_\beta \times P_c. \quad (29)$$

У випадку якщо у вибіркового обстеженні застосовується двоступеневий відбір (наприклад для обстеження сільськогосподарської діяльності населення), формула для загальної теоретичної ймовірності відбору для певного i -го домогосподарства (P_i) набуває вигляду:

$$P_i = P_a \times P_c. \quad (30)$$

Розрахунки за формулами (27) – (30) можуть здійснюватись як для всієї вибірки, так і в межах кожної страти окремо.

Х. Актуалізація основи вибірки

1. Головною метою реалізації процедури актуалізації основи вибірки є підвищення її якості шляхом унесення змін, які відбулися з моменту її формування або попередньої актуалізації.

2. Діюча в органах державної статистики система організації вибірових обстежень населення передбачає формування територіальної вибірки на термін п'ять років. Відповідно перший та другий етапи формування вибірки також здійснюються з такою ж періодичністю. У зв'язку з цим актуалізацію основ територіальної вибірки першого та другого ступенів відбору протягом зазначених п'яти років проводити недоцільно.

У випадках знесення житлових будинків або їх руйнування через стихійне лихо на території обстежуваної ВТОВ може виникнути необхідність її заміни. Разом з тим для таких випадків також недоцільно робити актуалізацію, оскільки

як для етапів відбору нових міських населених пунктів, так і для етапу відбору ВТОВ слід користуватися діючою основою територіальної вибірки. Це пов'язано з необхідністю застосування для розрахунків імовірностей відбору (формули (25), (26)) єдиної інформаційної бази для усіх адміністративно-територіальних одиниць.

3. Разом з тим у зв'язку зі змінами, які постійно відбуваються на відібраних територіях і призводять до певної невідповідності основи вибірки її реальному стану, актуалізацію списків адрес домогосподарств необхідно здійснювати перед кожним формуванням вибіркової сукупності. Актуалізація списків адрес домогосподарств є однією з основних засад, які забезпечують якісне функціонування вибіркової сукупності домогосподарств протягом усього строку дії територіальної вибірки.

4. Метою процедури актуалізації списків адрес домогосподарств, які є основою для формування нової вибіркової сукупності домогосподарств, є підвищення якості списків шляхом унесення до них змін, які відбулися з моменту складання списків або попередньої їх актуалізації.

5. Інформаційною базою для проведення актуалізації списків адрес домогосподарств є:

списки адрес домогосподарств, які були складені для попереднього відбору;

дані форми № 1 "Облікова картка об'єкта домогосподарського обліку", яка ведеться у сільських населених пунктах;

дані квартирних відділів підприємств, організацій, установ (по відомчих будинках); правлінь кооперативів (по кооперативних будинках); дані по окремих домовласниках та дані об'єднань співвласників багатоквартирних будинків; центрів надання адміністративних послуг, підприємств районних електромереж та управлінь газового господарства тощо;

відомості фахівців з інтерв'ювання щодо змін, які відбулися на територіях, на яких вони проводять вибіркоче обстеження.

6. Актуалізацію здійснюють працівники територіальних органів Держстату, до функціональних повноважень яких належать питання організації вибіркового обстеження населення. До виконання польових етапів робіт з актуалізації списків залучаються фахівці з інтерв'ювання.

7. За результатами актуалізації списків адрес домогосподарств на регіональному рівні формується зведена інформація про стан списків адрес домогосподарств у розрізі ПТОВ/ВТОВ регіону, копія якої надсилається на центральний рівень.

Інформація, зведена за підсумками стану списків адрес домогосподарств у попередньому циклі проведення обстежень, використовується для розрахунку

частки адрес домогосподарств непридатних до опитування для коригування мінімального розміру територіальної одиниці (додаток 4).

8. Інформація, отримана за результатами актуалізації списків адрес домогосподарств, також застосовується для щорічного аналізу якості побудови основи вибірки.

Контроль якості формування та актуалізації основ вибірки є обов'язковим етапом процесу формування вибірових сукупностей респондентів.

Порядок проведення цього контролю визначається методикою формування вибірових сукупностей для проведення певного циклу вибірових обстежень населення.

Директор департаменту
обстежень домогосподарств

І. І. Осипова

Додаток 1
до Методологічних положень
(пункт 2 розділу III)

**Макет документа "Вимоги до формування вибірових сукупностей
для проведення органами державної статистики
вибіркових обстежень населення (домогосподарств)"**

1. Мета побудови вибірки.
2. Генеральна сукупність.
3. Методологія побудови вибірки.
4. Одиниці спостереження.
5. Вимоги щодо представлення у вибірці територіальних одиниць відповідно до адміністративно-територіального устрою України.
6. Термін дії територіальної вибірки.
7. Тип вибірки, метод та одиниці відбору.
8. Кількість територіальних одиниць для відбору для проведення в них вибіркового обстеження одним фахівцем з інтерв'ювання у міських поселеннях та сільській місцевості.
9. Обмеження щодо представлення у вибірці окремих груп одиниць спостереження.
10. Перелік основних показників, надійність оцінок яких має бути забезпечена у вибіркового спостереженні.

		Назва обстеження		Назва обстеження		Назва обстеження	
11.	Навантаження фахівця з інтерв'ювання, домогосподарств:						
	у великих містах						
	у малих містах						
	у сільській місцевості						
12.	Надійність оцінювання основних показників, зазначених у пункті 10, коефіцієнт варіації (CV), %:						
		назва показника	назва показника	назва показника	назва показника	назва показника	назва показника
	на державному рівні						
	на регіональному рівні						
13.	Очікуваний рівень участі:						
	у великих містах						
	у малих містах						
	у сільській місцевості						
14.	Термін дії вибірки респондентів						
15.	Наявність і схема ротації одиниць спостереження						
16.	Обсяг вибірки респондентів						

17. Додаткові вимоги до формування вибірки.

Додаток 2
до Методологічних положень
(пункт 6 розділу IV)

**Схема ротації домогосподарств
в обстеженні економічної активності населення**

Рік	Квартал	Місяць	№ обстеження ¹	Ротаційні підгрупи			
1	2	3	4	5	6	7	8
2019	1	січень	1901	011	021	051	061
		лютий	1902	012	022	052	062
		березень	1903	013	023	053	063
	2	квітень	1904	021	031	061	071
		травень	1905	022	032	062	072
		червень	1906	023	033	063	073
	3	липень	1907	031	041	071	081
		серпень	1908	032	042	072	082
		вересень	1909	033	043	073	083
	4	жовтень	1910	041	051	081	091
		листопад	1911	042	052	082	092
		грудень	1912	043	053	083	093
2020	1	січень	2001	051	061	091	101
		лютий	2002	052	062	092	102
		березень	2003	053	063	093	103
	2	квітень	2004	061	071	101	111
		травень	2005	062	072	102	112
		червень	2006	063	073	103	113
	3	липень	2007	071	081	111	121
		серпень	2008	072	082	112	122
		вересень	2009	073	083	113	123
	4	жовтень	2010	081	091	121	131
		листопад	2011	082	092	122	132
		грудень	2012	083	093	123	133
2021	1	січень	2101	091	101	131	141
		лютий	2102	092	102	132	142
		березень	2103	093	103	133	143
	2	квітень	2104	101	111	141	151
		травень	2105	102	112	142	152
		червень	2106	103	113	143	153
	3	липень	2107	111	121	151	161
		серпень	2108	112	122	152	162
		вересень	2109	113	123	153	163
	4	жовтень	2110	121	131	161	171
		листопад	2111	122	132	162	172
		грудень	2112	123	133	163	173

¹ У графі "Номер обстеження" перші дві цифри позначають рік проведення обстеження, а дві наступні – місяць проведення обстеження.

Продовження додатка 2

1	2	3	4	5	6	7	8
2022	1	січень	2201	131	141	171	181
		лютий	2202	132	142	172	182
		березень	2203	133	143	173	183
	2	квітень	2204	141	151	181	191
		травень	2205	142	152	182	192
		червень	2206	143	153	183	193
	3	липень	2207	151	161	191	201
		серпень	2208	152	162	192	202
		вересень	2209	153	163	193	203
	4	жовтень	2210	161	171	201	211
		листопад	2211	162	172	202	212
		грудень	2212	163	173	203	213
2023	1	січень	2301	171	181	211	221
		лютий	2302	172	182	212	222
		березень	2303	173	183	213	223
	2	квітень	2304	181	191	221	231
		травень	2305	182	192	222	232
		червень	2306	183	193	223	233
	3	липень	2307	191	201	231	241
		серпень	2308	192	202	232	242
		вересень	2309	193	203	233	243
	4	жовтень	2310	201	211	241	251
		листопад	2311	202	212	242	252
		грудень	2312	203	213	243	253

Додаток 3
до Методологічних положень
(пункт 1 розділу IX)

Приклад розрахунку порога саморепрезентативності

Вхідна інформація для розрахунку порога саморепрезентативності для малих міст Вінницької області:

обсяг генеральної сукупності $N = 160\,396$ домогосподарств;
обсяг вибіркової сукупності $n = 102$ домогосподарства;
навантаження фахівця з інтерв'ювання $m = 17$ домогосподарств;
середній розмір домогосподарства $s = 2,47$ особи.

Поріг саморепрезентативності (n_p^{pers}):

$$n_p^{pers} = \frac{N}{n} \times m \times s = \frac{160396}{102} \times 17 \times 2,47 \approx 66\,000 \text{ осіб.}$$

Таким чином, територіальні одиниці (малі міста) з населенням $n \geq 66$ тис. осіб визначаються як саморепрезентативні й відповідно включаються у вибірку з імовірністю, рівною одиниці.

Додаток 4
до Методологічних положень
(пункт 1 розділу IX)

Приклад розрахунку мінімального розміру ВТОВ для проведення на окремих мережах респондентів двох вибіркового обстежень населення (домогосподарств)

Вхідна інформація для розрахунку мінімального розміру ВТОВ:
навантаження фахівця з інтерв'ювання $m = 26$ домогосподарств для обстеження;

кількість років перебування територіальної одиниці у вибірці $t = 5$ років;
коригувальний коефіцієнт (частка домогосподарств, непридатних до опитування) $k_{pp} = 0,12$;

кількість ротаційних груп (підгруп) у перший рік запровадження схеми ротації $q_1 = 27$;

кількість ротаційних груп (підгруп) у наступні роки функціонування схеми ротації $q = 12$;

розмір ротаційних груп/підгруп (домогосподарств або осіб (залежно від інформаційної бази, яка зазначається у Вимогах) $z = 3$

середній розмір домогосподарства у міських поселеннях $s = 2,58$ особи;

Згідно з Вимогами навантаження фахівця з інтерв'ювання має розподілятися між двома ВТОВ, тому для розрахунку мінімального розміру однієї ВТОВ (n_{min}) вводимо додатковий коригувальний коефіцієнт: $k_{part} = 0,5$.

Коригувальний коефіцієнт щодо частки домогосподарств, непридатних до опитування (k_{pp}) для розрахунку мінімального розміру ВТОВ, визначається на основі даних про кількість зачинених або нежитлових приміщень, а також приміщень, у яких немає осіб, які проживають постійно, на основі інформації, зведеної за підсумками стану списків адрес домогосподарств у попередньому циклі проведення обстежень.

Мінімальний розмір ВТОВ для проведення обстеження з повною щорічною ротацією розраховуємо за формулою:

$$n_{min 1} = m \times (t + 1) \times k_{part} / (1 - k_{pp}).$$

Згідно з Вимогами термін дії територіальної вибірки становить 5 років, а з урахуванням резервного року проведення обстеження – 6 років; максимальне навантаження фахівця з інтерв'ювання становить 26 домогосподарств, коригувальні коефіцієнти становлять 0,5 (частка домогосподарств у одній ВТОВ при розподілі між двома ВТОВ) та 0,12 (частка домогосподарств, непридатних до опитування). Таким чином:

$$n_{min 1} = (26 \times 6) \times 0,5 / (1 - 0,12) \approx 89 \text{ (д/г)}.$$

Мінімальний розмір ВТОВ для проведення обстеження з частковою щорічною ротацією розраховуємо за формулою:

$$n_{min 2} = (q_1 \times z) + (q \times z) \times t / (1 - k_{pp}).$$

Продовження додатка 4

Відповідно Вимог термін дії територіальної вибірки становить 5 років, а з урахуванням резервного року проведення обстеження – 6 років; максимально можлива кількість домогосподарств у ротаційній групі становить 3 домогосподарства; для проведення обстеження в перший рік для запровадження схеми ротації відбирається 27 ротаційних підгруп, для проведення обстеження у інші роки щорічно відбирається 12 ротаційних підгруп; коригуючий коефіцієнт становить 0,12 (частка домогосподарств, непридатних до опитування), Таким чином, мінімальний розмір ВТОВ становить:

$$n_{min 2} = ((27 \times 3) \times 1) + ((12 \times 3) \times 5) / (1 - 0,12) \approx 297 \text{ (д/г)}.$$

Отже, мінімальний розмір ВТОВ для проведення на окремих мережах респондентів двох вибірових обстежень населення становитиме:

$$n_{min} = (297 + 89) = 386 \text{ (домогосподарств)},$$

або:

$$n_{min} = 386 \times 2,58 = 996 \text{ (осіб)}.$$

Додаток 5
до Методологічних положень
(пункт 1 розділу IX)

Приклад упорядкування територіальних одиниць за розміром та за принципом "географічного серпантину", а також реалізації процедури відбору територіальних одиниць

Вхідні дані:

У певній страті із загальною кількістю населення 183 тис. осіб заходяться 7 територіальних одиниць:

A (кількість населення – 23 тис. осіб);

B (кількість населення – 31 тис. осіб);

C (кількість населення – 24 тис. осіб);

D (кількість населення – 34 тис. осіб);

E (кількість населення – 27 тис. осіб);

F (кількість населення – 26 тис. осіб);

G (кількість населення – 18 тис. осіб).

Територіальне розміщення цих одиниць наведено на рис. 1.

Упорядкуємо ці територіальні одиниці за двома принципами:

1) ранжування за зменшенням кількості населення (табл. 1);

2) ранжування за принципом "географічного серпантину" (табл. 2).

Таблиця 1. Приклад упорядкування територіальних одиниць за розміром (кількістю населення)

Територіальна одиниця	Кількість населення (тис. осіб)	Накопичена кількість населення (тис. осіб)	Значення кількості населення для відбору (тис. осіб)
D	34	34	
B	31	65	36,46
E	27	92	
F	26	118	97,46
C	24	142	
A	23	165	158,46
G	18	183	
Усього	183	x	

Таблиця 2. Приклад упорядкування територіальних одиниць за принципом "географічного серпантину"

Територіальна одиниця	Кількість населення (тис. осіб)	Накопичена кількість населення (тис. осіб)	Значення кількості населення для відбору (тис. осіб)
A	23	23	
B	31	54	36,46
C	24	78	
D	34	112	97,46
E	27	139	
F	26	165	158,46
G	18	183	
Всього	183	x	

Визначення кількості несаморепрезентативних територіальних одиниць, які необхідно відібрати.

Додаткові вхідні дані:

поріг саморепрезентативності $n_p \approx 66$ тис. осіб;

кількість населення у страті (у вищенаведених несаморепрезентативних територіальних одиницях) $N_{sstr}^{pers} = 183$ тис. осіб.

Кількість несаморепрезентативних територіальних одиниць, яку необхідно відібрати (d):

$$d = N_{sstr}^{pers} / n_p^{pers}$$

$$d = 183000 / 66\ 000 = 2,772 \approx 3.$$

Таким чином, слід відібрати 3 територіальні одиниці із 7.

Визначення кроку відбору.

Відповідно до даних попереднього прикладу ($N_{sstr}^{pers} = 183000$ тис. осіб, $d = 3$) крок відбору становить (h):

$$h = N_{sstr}^{pers} / d = 183000 / 3 = 61 \text{ тис. осіб.}$$

Визначення першої несаморепрезентативної територіальної одиниці для відбору.

Вхідні дані:

випадкове число $rnd = 0,597704$;

крок відбору $h = 61$ тис. осіб.

Значення для відбору першої територіальної одиниці (h_1):

$$h^{(1)} = h \times rnd = 61000 \times 0,597704 = 36459,9 \approx 36,46 \text{ тис. осіб.}$$

Продовження додатка 5

Отримане значення порівнюється з накопиченими кількостями населення у переліку територіальних одиниць. Територіальна одиниця, накопиченому значенню кількості якої відповідає результат, є першою територіальною одиницею вибірки. У таблицях 1 і 2 – це територіальна одиниця В.

Відбір наступних несамопрезентативних територіальних одиниць.

Усі наступні територіальні одиниці визначаються шляхом додавання кроку відбору до накопиченої кількості населення, розрахованої для відбору попередньої територіальної одиниці:

$$h^{(z)} = h^{(z-1)} + h.$$

Приклад відбору територіальних одиниць за даними таблиць 1 та 2:

$$h^{(1)} = 36,46;$$

$$h^{(2)} = h^{(1)} + h = 36,46 + 61 = 97,46;$$

$$h^{(3)} = h^{(2)} + h = 97,46 + 61 = 158,46.$$

Таким чином, у випадку впорядкування територіальних одиниць за розміром (ранжування за зменшенням кількості населення) будуть відібрані територіальні одиниці В, F та А (табл. 1); у випадку упорядкування за принципом "географічного серпантину" – В, D, та F (табл. 2).

Список використаних джерел

1. Закон України "Про державну статистику". [Електронний ресурс]. – Режим доступу: <http://www.ukrstat.gov.ua>
2. Закон України "Про інформацію". [Електронний ресурс]. – Режим доступу: <http://www.ukrstat.gov.ua>
3. Глосарій до плану статистичного спостереження, затверджений наказом Держкомстату від 29 грудня 2009 року № 498 [Електронний ресурс]. – Режим доступу: <http://www.ukrstat.gov.ua>
4. Статистичний словник/[О. Г. Осауленко, О. О. Васечко, М. В. Пугачова та ін.], за ред. д-ра держ. упр., проф., член-кор. НАН України О. Г. Осауленка; НТК статистичних досліджень. – К.: ДП "Інформ.-аналіт. агентство", 2012. – 498 с.
5. Бокун Н. Ч., Чернышова Т. М. Методы выборочных обследований: Учебно-справочное пособие. – Мн.: НИИ статистики Республики Беларусь, 1997. – 416с.
6. Єріна А. М. Організація вибірових обстежень: Навчальний посібник. – К.: КНЕУ, 2004. – 127 с.
7. Йейтс Ф. Выборочный метод в переписях и обследованиях. – М.: Статистика, 1965. – 433 с.
8. Кокрен У. Методы выборочного исследования. - М.: Статистика, 1976. - 440с.
9. Колтон Г. Структура выборки для проведения обследования домашних хозяйств на Украине // Отчет Госкомстата Украины. 1998. – 20 с.
10. Розробка та впровадження методології формування вибіркової сукупності для обстеження бюджетів домашніх господарств, її ротації та розповсюдження результатів на генеральну сукупність / Науковий звіт НДІ статистики Держкомстату України за 1998 рік.
11. Саріогло В. Г. Проблеми статистичного зважування вибірових даних: Монографія. – К.: ІВЦ Держкомстату України, 2005. – 264 с.
12. Labour force survey in the EU, candidate and EFTA countries — Main characteristics of national surveys, 2015 [Електронний ресурс.–Режим доступу: <http://ec.europa.eu/eurostat/documents/3888793/7751652/KS-TC-16-021-EN-N.pdf/8475c2e2-c037-4ba2-9029-93db1ade41fe>].
13. Household Budget Surveys in the EU Methodology and recommendations for harmonisation – 2003 [Електронний ресурс.–Режим доступу: http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-BF-03-003-__-N-EN.pdf]
14. Kalton G. Introduction to Survey Sampling. Newbury Park: SAGE Publications Inc., 1983. – 96 p.
15. Kish L. Survey sampling.- Wiley Classics Library Edition Published, 1995.- 643p.
16. Methodology of European labour force surveys:(2) Sample design and implementation Francesca Gagliardi, Vijay Verma, Giulia Ciampalini Working Paper n. 79, January 2009 [Електронний ресурс.–Режим доступу: https://www.researchgate.net/publication/228877388_Methodology_of_European_labour_force_surveys2_Sample_design_and_implementation].
17. Verma V. Workshop on Labor Force Surveys for CIS Countries and the Baltic States run by the International Labor Office, Bureau of Statistics hosted by State

Statistics Committee of Ukraine (Kyiv, 14-19 September 1998). Part Three: sample design, data evaluation, field operations.

18. Regulation (EC) No 1982/2003 of 21 October 2003 implementing Regulation (EC) No 1177/2003 of the European Parliament and of the Council concerning Community statistics on income and living conditions (EU-SILC) as regards the sampling and tracing rules [Электронный ресурс.—Режим доступа:<http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32003R1982&from=EN>].