

Тема. Методична частина програми соціологічного дослідження (Ч.1. Уточнення об'єкту дослідження (генеральна і вибіркова сукупність).

1. Уточнення об'єкту дослідження. Поняття генеральної та вибіркової сукупності. Репрезентативність вибірки.

2. Об'єм та похибка вибірки.

3. Основні методи побудови вибірки.

Побудовою вихідних гіпотез дослідження завершується розробка методологічної частини програми і наступним кроком є перехід до формулювання методичних складових, а саме:

1. Уточнення об'єкту дослідження, шляхом визначення генеральної та вибіркової сукупності.

2. Визначення методу (методів) збору соціологічної інформації.

3. Побудова логічних схем інструментарію (відповідно до обраного методу).

4. Побудова логічних схем обробки інформації.

1. Уточнення об'єкту дослідження. Поняття генеральної та вибіркової сукупності. Репрезентативність вибірки.

Для того щоб спланувати так званий польовий етап дослідження, який передбачає безпосередньо збір соціологічної інформації необхідно уточнити межі об'єкту, що вивчається. **Об'єктом соціологічного дослідження в широкому сенсі виступає носій тієї чи іншої соціальної проблеми (як правило, різні соціальні групи, наприклад, студенти, працівники певного підприємства, молоді фахівці, глядачі певного телеканалу, користувачі соціальної мережі, громадяни країни тощо).** Отже зазвичай об'єктом прикладного (не плутати з теоретичним) дослідження виступає певна соціальна спільнота, виключенням може бути аналіз документів, де об'єктом стає сукупність документів. Ці спільноти можуть бути як малими (студентська група, працівники малого підприємства, сім'я і т.д.) так і дуже великими (молодь, користувачі соціальних мереж, громадяни країни). Коли об'єктом виступає мала за об'ємом соціальна група її можна дослідити повністю, говорять про суцільне дослідження, але коли соціолог вивчає більші соціальні спільноти їх досить важко охопити, бо організація і проведення суцільного обстеження в даному випадку буде пов'язана з великими матеріальними витратами, тому в таких випадках використовують вибіркові дослідження.

Тож більшість соціологічних досліджень мають вибірковий характер. Це пов'язано, перш за все, з тим, що економічні і тимчасові обмеження не дозволяють проводити суцільні дослідження. Але навіть в тих випадках, коли суцільне дослідження практично може бути реалізоване, часто рентабельніше проводити вибіркоче.

У переважній більшості випадків соціолог використовує той чи інший спосіб виділення з великої сукупності явищ і об'єктів вивчення деяку їх частину в надії, що на цій вибіркової сукупності можуть бути виявлені властивості об'єкта дослідження в цілому, і він зможе поширити свої висновки на всіх носіїв проблеми. Тому від того як в програмі визначається і уточнюється об'єкт дослідження, яким методом користується соціолог, щоб визначити ту частину соціальної спільноти яку він буде безпосередньо досліджувати залежить кінцевий результат.

Отже спочатку соціологу потрібно уточнити об'єкт дослідження, а саме визначити ту соціальну спільноту, що є носієм проблеми та на яку він зможе поширити результати дослідження – це і буде **«генеральна сукупність»**.

Всі члени групи, що цікавить дослідника складають генеральну сукупність, а **«вибірка»** - це певна частина (певна кількість) елементів генеральної сукупності, відібраних по чітко заданому правилу, що дозволяє робити більш-менш точні висновки щодо сукупності в цілому. Навіщо потрібно будувати вибірки? Перш за все, з практичних міркувань, так як вибірка економить сили і ресурси дослідників. **Вибірку також можна назвати мікромоделлю генеральної сукупності, яка служить одним з найбільш економних засобів для перевірки припущень або гіпотез.**

Однією з найважливіших вимог до вибірки є вимога репрезентативності. **Репрезентативність** - здатність вибіркової сукупності за виділеними параметрами (критеріями) відтворювати характеристики генеральної сукупності. Однак абсолютно репрезентативну вибірку яка б повністю збігалась з генеральною сукупністю забезпечити неможливо (завжди будуть певні відхилення, або фактори, що не були враховані), тому слід гарантувати репрезентативність по основних параметрах, що будуть важливі для подальшого аналізу даних.

Перш за все, слід усвідомити, які з наявних відомостей про характеристики генеральної сукупності важливі для цілей дослідження. У більшості випадків критеріями (або параметрами) репрезентативності є статево-віковий і соціально-професійний склад обстежуваних, їх просторова локалізація (місце проживання, роботи, навчання і т.д.).

Статево-віковий склад населення – результат дії багатьох факторів, як демографічних (народжуваності, смертності, міграційних процесів), так і тих, що на них впливають – історичних та соціально-економічних, тому цей параметр використовують найчастіше.

Соціально-професійні характеристики – це свідчення роду занять, з чим пов'язані інтереси, особливості режиму праці і відпочинку, багато інших важливих показників діяльності людей.

Просторова локалізація (по території проживання, роботи, навчання і т.д.) важлива і з точки зору особливостей умов цієї діяльності (наприклад, центр і периферія, місто чи село), і з точки зору адресності підсумкових висновків і рекомендацій, які повинні бути «прив'язані» до адміністративних осередків, які мають чіткі межі.

У поєднанні трьох названих параметрів - статево-вікової структури, соціально-професійного складу, просторової локалізації - можна, як правило, бути впевненим, що вибірка буде представницька (репрезентативна) для вирішення багатьох соціальних проблем. Зрозуміло, що це правило має винятки, в залежності від конкретних умов і особливих цілей дослідження (наприклад, в етнічно неоднорідному середовищі варто мати на увазі репрезентацію за критерієм національної приналежності), тож дослідник може варіювати параметри (критерії) репрезентативності, відповідно до мети і завдань дослідження.

2. Об'єм (обсяг) та похибка вибірки.

Чисельність (обсяг) вибірки залежить від рівня однорідності або різномірності досліджуваних об'єктів. Чим більше вони однорідні, тим менша чисельність може забезпечити статистично достовірні висновки. Але ступінь

однорідності соціального об'єкта залежить, по суті, від того, наскільки детально ми маємо намір його досліджувати. Практично будь-який «елементарний» об'єкт при детальному вивченні виявляється надзвичайно складним. Лише в аналізі ми представляємо його як відносно простий, виділяючи ті чи інші його властивості. Чим більш ґрунтовним і детальним буде аналіз, чим більше властивостей даного об'єкта ми маємо намір прийняти до уваги в їх поєднанні, а не ізольовано, тим більше повинен бути обсяг вибірки.

Міра відмінності вибіркової моделі від структури і параметрів генеральної сукупності оцінюється похибкою вибірки, а межі допустимої похибки знову-таки залежать від мети дослідження. Іноді потрібна підвищена надійність, як це має місце в економічних і демографічних обстеженнях, наприклад при переписах населення. Тут істотні помилки обертаються мільйонними витратами матеріальних ресурсів і прорахунками планування, тому похибка вибірки повинна бути мінімальною. Набагато частіше соціологічні обстеження проводяться для з'ясування певних загальних тенденцій, загального орієнтування, відповідно допускаються більші відхилення від параметрів генеральної сукупності.

Дуже корисна наступна приблизна **оцінка надійності результатів вибіркового обстеження**. Підвищена надійність результатів - похибка вибірки (ступінь її відхилення від параметрів генеральної сукупності) до 3%, звичайна - від 3% до 5%, наближена - від 5% до 10%, орієнтовна - від 10 до 20%.

Обсяг вибірки залежить напряду не від збільшення генеральної сукупності, а скоріше від її складності та різномірності, рівня довірчого інтервалу, допустимої помилки, яка, задається доцільною точністю підсумкових узагальнень: від підвищеної до орієнтовної.

В.І. Паніотто наводить такі розрахунки репрезентативної вибірки з допущенням 5% помилки:

Розміри вибірки для різних генеральних сукупностей

Обсяг генеральної сукупності	500	1000	2000	3000	4000	5000	10000	100000
Обсяг вибірки	222	286	333	350	360	370	385	458

Якщо необхідний інший довірчий інтервал або інший показник похибки вибірки можна скористатися калькулятором вибірки (доступний он-лайн на багатьох ресурсах). Для розрахунку на калькуляторі потрібно задати обсяг генеральної сукупності, довірчий інтервал, який вас влаштовує (точність вибірки, довірчим інтервалом будь-якого параметра генеральної сукупності називається випадкова область значень цього параметра, яка з ймовірністю близькою до 1 (надійністю) містить істинне значення цього параметра.) (зазвичай 95%) і % помилки вибірки від 5% до 3%. Наприклад для міста Запоріжжя (759 596 тис. Чол.) (95 і 3 - 1066) (95 і 4 - 600) (95 і 5 - 384).

Отже, обсяг вибіркової сукупності залежить від безлічі факторів, основні з них такі:

1. Витрати на збір інформації, включаючи тимчасові.
2. Прагнення до певної статистичної достовірності результатів, яку сподівається отримати дослідник.

3. Цінність і новизна інформації, одержуваної в результаті опитування.

Розмір вибірки зростає:

- при необхідності опублікувати дані для окремих підгруп (розміри підвибірок при цьому підсумовуються, і вибірка в цілому зростає пропорційно числу підгруп);

- при проведенні загальнонаціональних обстежень, коли велика генеральна сукупність;

- якщо вже наявна інформація щодо ключових питань (наприклад, про наміри виборців голосувати за ту чи іншу партію) явно недостатня і ступінь невизначеності значний.

Розмір вибірки зменшується:

- при дослідженні груп, організацій та інших «первинних одиниць відбору», якщо порівняно невелика величина генеральної сукупності, з якої проводиться відбір (наприклад, сукупності співробітників рекламних агентств, школярів, пацієнтів і т.п.);

- при проведенні локальних і регіональних досліджень;

- якщо вже існуюча інформація відносно повна та хоч все ще залишається ступінь невизначеності, він незначний.

3. Основні методи побудови вибірки.

Отже, обсяг (об'єм) вибірки визначено, ми знаємо, скільки людей нам треба опитати. Але як нам відібрати цю кількість з генеральної сукупності, яким методом? Всі вибірки діляться на два великі класи: **випадкові (ймовірнісні) і не випадкові (цільові).**

Випадкова вибірка забезпечує рівність шансів всіх елементів генеральної сукупності потрапити до вибірки. Реалізація цього виду вибірки передбачає наявність списків (карток) і нумерацію всіх елементів генеральної сукупності.

У не випадкових **цільових вибірках** принцип рівності шансів не дотримується.

Існує кілька типів ймовірнісної вибірки, що розрізняються характером вибіркової процедури. **Найбільш поширені проста випадкова, систематична, стратифікована, кластерна і багатоступенева.**

1. Проста ймовірнісна вибірка. Простий випадковий відбір. При простому випадковому відборі кожна одиниця повинна мати рівну можливість бути витягнутої з сукупності в кожному виборі, причому кожен вибір має здійснюватися незалежно від всіх ін. Одиниці спостереження в ймовірнісної вибіркою вибираються випадково.

Цей метод може бути реалізований за принципом лотереї (який через свою надзвичайної трудомісткості практично ніколи не застосовується і в соціології має виключно теоретичне значення) або за допомогою таблиць випадкових чисел. Обидва методи передбачають складання списку і попередню нумерацію всіх елементів генеральної сукупності.

Найбільш очевидні обмеження для використання простий вибірки виникають в разі великого обсягу генеральної сукупності. Перш за все дослідник стикається з проблемами пошуку повної генеральної сукупності (переписи, списки виборців і т. Д.).

Різновидом випадкових вибірок є також: 2. механічна (систематична), коли з генеральної сукупності (наприклад, списковий склад співробітників) через рівні проміжки відбирають необхідну кількість респондентів Систематична вибірка, як і проста випадкова, вимагає повного списку. Техніка здійснення систематичного відбору елементарна: спочатку випадковим чином відбирається перша одиниця, потім

відбору підлягає кожен k-й елемент. Число k в даному випадку називають кроком відбору. Можна, наприклад, відбирати кожен 25-й або кожен 200-й елемент.

Щоб визначити крок відбору, потрібно поділити відомий обсяг генеральної сукупності (N) на передбачуваний обсяг вибірки (n).

Нехай, наприклад, потрібно відібрати 200 чоловік з 20000 власників телефонів:

1) визначимо крок відбору: $N / n = 20000 : 200 = 100$;

2) за допомогою таблиці випадкових чисел знайдемо першу вибірку одну одиницю.

Якщо, скажімо, випав номер «053», то зі списку власників телефонів випишемо того, хто значиться під цим номером;

3) з встановленим кроком відбираємо номери: 153, 253, 353, 453 і т. д. до вичерпання списку.

Випадкова вибірка є ідеальною з точки зору теоретичних міркувань. Однак у реальній соціологічній практиці, коли в якості генеральної сукупності використовується ніяк не перетворений список одиниць, які є об'єктом дослідження, проста випадкова вибірка використовується вкрай рідко. Це викликано рядом причин:

⇒ відсутність достовірних і повних відомостей про генеральну сукупність (найбільш типова);

⇒ громіздкість необхідних для її реалізації процедур;

⇒ економічна і наукова неефективність.

Тому в соціологічній практиці використовуються модифіковані випадкові вибірки. Суть модифікації полягає в тому, що процедура відбору розпадається на кілька послідовних етапів - ступенів. На кожному ступені здійснюється структурування (стратифікація) тієї частини, що необхідно репрезентувати в вибірці.

3. Стратифікована вибірка використовується в тих випадках, коли з якихось змістовних міркувань важливо забезпечити представництво ймовірнісної вибірки по якимось конкретним, важливим для дослідницьких цілей, критеріям.

Стратифікацією, строго кажучи, називають процедуру, при якій відбір здійснюють як би з декількох «паралельних» частин (підмножин), заданих на одній і тій же генеральній сукупності. Припустимо, ми знаємо, що в генеральній сукупності 60% жінок і 40% чоловіків. В описаній ситуації бажано заздалегідь забезпечити представленість обох цих груп (страт), зберігши імовірнісний характер відбору. Цього можна домогтися, якщо здійснити якусь незалежну процедуру випадкового відбору для кожної соціальної групи окремо (в нашому прикладі для жінок і чоловіків) і потім об'єднати отримані випадкові підвибірки в одну (зауважте, що для нашого прикладу обсяг підвибірки жінок, буде в 1,5 рази більше обсягу підвибірки чоловіків, оскільки ми заздалегідь знаємо пропорцію їх присутності в генеральній сукупності). Отримана в результаті вибірка буде і стратифікованою (за статтю), і ймовірнісною. **Іноді такий метод формування вибірки ще називають пропорційним.**

При використанні 4. кластерної вибірки з генеральної сукупності відбирають об'єкти, що представляють собою групи (кластери) дрібніших одиниць і вже всередині проводять суцільне опитування. Кластери це природні угруповання одиниць спостереження. Наприклад, генеральна сукупність військовослужбовців природним чином групується по військовим частинам і підрозділам, а сукупність студентів - по університетам, інститутам і коледжам.

5. Гніздовий відбір. Простий і стратифікований випадковий відбір, особливо при інтерв'юванні територіально розподіленої популяції, вимагає значних переїздів і, відповідно, великих фінансових витрат. Гніздовий відбір істотно скорочує ці витрати шляхом розбиття досліджуваної сукупності на групи за географічною ознакою. Ця методика, наприклад, полягає в накладенні масштабної сітки на географічну карту області розселення досліджуваної популяції, випадкового вибору квадратів (гнізд) сітки, які мають цю популяцію, і подальшому обстеженні елементів кожного гнізда. Оскільки в гнізді може виявитися більше одиниць, ніж потрібно для дослідні., Гніздовий відбір часто використовують в якості першого ступеня багатоступеневого відбору. В якості гнізд в соціологічній практиці можуть виступати регіони (адміністративно-територіальні утворення, сфери діяльності (промисловість, сільське господарство, транспорт і т.п.), виборчі або лікарські дільниці, будинки, підприємства, установи, соціально-професійні групи і т.д.

До другої групи не випадкових, цілеспрямованих вибірок незастосовні правила і закони теорії ймовірності.

До даної групи належать методи відбору, які не ґрунтуються на принципі рівності шансів кожної одиниці генеральної сукупності потрапити у вибірку. Умови відбору в вибірках цього типу визначаються дослідником, виходячи з різних міркувань, серед яких часто велику роль відіграють суб'єктивні.

Основні фактори, які визначають природу не випадкового відбору, - а) готовність дослідника мати справу з максимально доступним для нього числом одиниць генеральної сукупності (фактор доступності); б) прагнення максимально гарантувати потрапляння до вибірки тих одиниць, які мають певні властивості (фактор доцільності).

Перший фактор означає, що соціолог готовий змиритися з тим, що рішення про попадання у вибірку належить респонденту. Другий фактор означає, що це рішення належить саме йому.

До такого класу вибірок відносяться стихійна вибірка (першого зустрічного); вибірка основного масиву, при якій опитується 60-70% генеральної сукупності; квотна вибірка, метод «снігової грудки (кулі)»

Серед не випадкових цілеспрямованих вибірок **квотна** на сьогоднішній день є найбільш популярною в опитуваннях громадської думки. Для формування квотної вибірки треба знати параметри генеральної сукупності з потрібних для нас ознак. Всі дані про будь-якій ознаці називаються квотою, а їх окреме числове значення - параметром квоти. При такому типі вибірки респонденти вибираються цілеспрямовано, з дотриманням параметрів квот. Число ознак, дані про яких вибираються в якості квот, зазвичай не перевищує чотирьох.

Метод «снігової кулі» різновид експертної вибірки, при якому респонденти (вони ж виступають в ролі експертів-інформаторів) повідомляють імена інших потенційних респондентів, які володіють необхідними характеристиками, таким чином розмір вибірки-«кулі» зростає доти, доки імена не починають повторюватись. Цінність цього методу особливо велика, коли необхідна участь «рідкісних одиниць» (наприклад, вчених з дуже вузькою спеціалізацією).