

## Лекція 6. ОСНОВИ ВИБІРКОВОГО СПОСТЕРЕЖЕННЯ

### План

1. Причини і умови застосування вибіркового спостереження, його завдання, основні поняття.
2. Формування вибіркової сукупності. Види та схеми відбору. Мала вибірка.
3. Похибки вибірки.

### Основні терміни та поняття

Спостереження. Вибірка. Генеральна сукупність. Одиниця вибірки. Детермінована вибірка. Ймовірна вибірка. Багатоступенева вибірка. Похибка. Повторний відбір. Безповторний відбір.

### Тематичні тези

#### 1. Причини і умови застосування вибіркового спостереження, його завдання, основні поняття

Вибіркове спостереження є найпоширенішим з усіх видів несучого цільного спостереження. Під час проведення вибіркового спостереження вивчаються не всі одиниці досліджуваного об'єкта, а лише деяка відібрана частина цих одиниць. Однак спостереження організоване таким чином, що ця частина відібраних одиниць відображає всю сукупність наче у зменшеному масштабі.

Загальну сукупність однорідних об'єктів (одиниць) з якої здійснюється відбір для дослідження у статистиці називають генеральною сукупністю.

Відібрана для обстеження частина генеральної сукупності має назву вибірки або вибіркової сукупності. На підставі результатів дослідження вибіркової сукупності можна зробити висновки про властивості генеральної сукупності, тобто теоретичні характеристики замінити статистичними.

До теоретичних (шуканих) характеристик належать узагальнюючі показники генеральної сукупності – середній розмір ознаки (Сер.), частка (P), генеральна дисперсія ( $\sigma^2$ ), які називають параметрами.

Статистичними характеристиками (оцінками) є характеристики вибіркової сукупності – середня вибіркова (xсер.в.), вибіркова частка (W), вибіркова дисперсія ( $\sigma^2_{в.}$ ), які називають статистиками.

Отже завдання вибіркового спостереження – за результатами обстеженої частини надати характеристику усій сукупності в цілому.

Переваги вибіркового спостереження перед суцільним:

- практичність (іноді немає можливості чи недоцільно здійснювати суцільне спостереження, наприклад, контроль якості книжкової продукції – підручників, дитячої книги тощо);
- висока точність результатів за рахунок зниження помилок реєстрації (наприклад, соціологічні чи інші наукові дослідження, що вимагають професійних знань суті явища);
- проведення статистичних досліджень, пов'язаних із знищенням зразків (перевірка якості харчової продукції, екологічного стану та інше);

- мінімальні затрати, швидкість дослідження, детальніше вивчення кожної одиниці спостереження.

Ці переваги вибіркового спостереження мають місце лише за умов дотримання наукових принципів його проведення, що надає можливість отримати репрезентативну вибірку, яка відтворює характеристики генеральної сукупності.

Порушення цих принципів, коли спостереженню підлягають одиниці, відібрані на основі суб'єктивного погляду дослідника, призводять до того, що результати такого спостереження мають відношення не до усєї генеральної сукупності, а тільки до тієї її частини, яка була піддана спостереженню.

Вибірка вважається репрезентативною, якщо кожен елемент генеральної сукупності має однакову ймовірність (рівну можливість) потрапити у вибірку і відбір здійснюється строго випадковим чином за схемами, розробленими математичною статистикою. А також, коли обсяг вибіркової сукупності є достатнім.

Вибіркове спостереження здійснюється у такій послідовності:

- обґрунтування мети, складання програми;
- визначення необхідної частки (5%, 25%, 50% чи ін.) одиниць для обстеження та способу їх відбору до вибіркової сукупності;
- формування вибірки;
- узагальнення даних спостереження і визначення їхніх вибірових характеристик (оцінок відповідних характеристик генеральної сукупності);
- обчислення похибок вибірки, перевірка надійності результатів;
- поширення вибірових даних на всю досліджувану сукупність (висновки про параметри генеральної сукупності).

## **2. Формування вибіркової сукупності. Види та схеми відбору. Мала вибірка**

Точність результатів вибіркового спостереження залежить від способу відбору одиниць, варіації ознаки в сукупності і від кількості відібраних одиниць.

Система правил відбору одиниць і способів характеристики досліджуваної сукупності складає зміст вибіркового методу.

Розрізняють такі способи формування вибірових сукупностей:

- власне випадковий (простий випадковий) відбір;
- механічний (систематичний);
- типовий (районований);
- серійний (гніздовий);
- комбінований;
- багатоступінчастий та багатofазний.

Усі види відбору (крім механічного) можуть бути проведені повторним чи неповторним методом, який визначає можливість для відібраної одиниці продовжувати участь в процедурі відбору. Механічний добір завжди є неповторним.

Повторним називають відбір, при якому відібрана одиниця повертається назад у генеральну сукупність і може повторно брати участь у вибірці. При такому доборі шанси (ймовірність) кожної окремої одиниці потрапити у вибірку залишаються постійними.

Безповторним є відбір, у якому відібрана одиниця не повертається до генеральної сукупності і в подальшій вибірці участі не бере. В цьому випадку ймовірність одиниці потрапити до вибірки постійно змінюється – для останніх одиниць вона зростає.

При достатньо великому обсязі генеральної сукупності, оцінки її характеристик, отримані на основі безповторної вибірки, істотно не відрізняються від оцінок, отриманих на основі повторної вибірки.

Випадкова вибірка – це класичний спосіб формування вибіркової сукупності, який передбачає попередню досить складну підготовку до формування вибірки. Дає добрі результати за умов однорідності сукупності.

Простий випадковий добір проводиться жеребкуванням або за допомогою таблиць випадкових чисел без будь-яких елементів системності (без попереднього розчленування генеральної сукупності на окремі групи). Для жеребкування на кожну одиницю генеральної сукупності необхідно заготувати відповідну фішку, шар або картку з інформацією щодо окремої одиниці сукупності. При використанні таблиць випадкових чисел усі елементи сукупності мають бути пронумеровані. Необхідно також встановити чіткі межі генеральної сукупності таким чином, щоб присутність чи відсутність у ній окремих одиниць не викликала сумніву. Так, наприклад, при обстеженні студентів необхідно вказати, чи будуть прийматися до уваги особи, що перебувають в академічній відпустці. У великих за обсягом сукупностях така робота здебільшого недоцільна, а часом і неможлива, тому, незважаючи на простоту власно випадкового відбору, на практиці застосовуються інші різновиди випадкових вибірок, які представляють собою видозмінення простої випадкової вибірки.

Однією з найпоширеніших форм в практиці вибіркового спостереження є механічний відбір. Це послідовний вибір одиниць, який здійснюється за списками чи безпосередньо на місці дослідження через рівні проміжки в розташуванні цих одиниць у генеральній сукупності. Розташування одиниць сукупності буває впорядкованим (ранжированим) або невпорядкованим (випадковим) відносно досліджуваної ознаки.

Наприклад, список співробітників Інституту в порядку зменшення заробітної плати або список студентів за алфавітом. Генеральна сукупність розбивається на рівні за чисельністю одиниць групи (інтервали) і з кожної групи для обстеження вибирається одна одиниця. При цьому задається початок, крок відліку і обсяг вибірки.

Початок відліку залежить від способу розташування одиниць в генеральній сукупності. У випадку їх невпорядкованого розташування він може бути встановлений

довільно або жеребкуванням з сукупності одиниць першого інтервалу. При впорядкованому розташуванні за початок відліку приймають одиницю,

яка розташована всередині першого інтервалу, щоб уникнути систематичної помилки вибірки.

Крок відліку визначається співвідношенням обсягів генеральної та вибіркової сукупностей  $h = N/n$ , наприклад, з генеральної сукупності обсягом 1000 од. Обстеженню підлягає 100 од. ( $1000 / 100 = 10$ ), це означає, що з кожних 10 од. буде обрано для обстеження 1 од. (кожну 10-ту), крок відліку дорівнює 10.

Механічний відбір дуже зручно застосовувати у випадках, коли не можна заздалегідь скласти список одиниць генеральної сукупності, наприклад, у маркетингових дослідженнях можна спостерігати кожного 10-го покупця, що увійшов до магазину, при дослідженні самостійної роботи студентів кожного 5-го відвідувача читального залу бібліотеки та комп'ютерного класу. В такому разі вибірка береться з сукупності, яка формується поступово чи з практично нескінченної сукупності.

Типовий (розшарований, районований) добір орієнтований на забезпечення представництва у вибірці відповідних типових груп генеральної сукупності. Формування вибірки іде з урахуванням структури генеральної сукупності. На відміну від простого випадкового та механічного добору, які проводяться в цілому по генеральній сукупності, типовий передбачає розділення (розшарування) її на якісно однорідні групи за якоюсь важливою типовою ознакою. Наприклад, при обстеженнях населення такими групами можуть бути, райони, соціальні, вікові чи освітні групи. Потім з кожної групи проводять незалежний відбір елементів простим випадковим чи механічним способом пропорційно питомій вазі (обсягу) групи у загальній сукупності.

Перевагою типового відбору є отримання більш достовірних результатів порівняно з попередніми відборами. Оскільки до вибірки потрапляють представники усіх типових груп ймовірність отримати більшу точність вибіркових даних тут вища.

Застосовувати цей метод відбору краще у разі неоднорідності одиниць досліджуваної сукупності.

Серійний добір зручно використовувати у тих випадках, коли одиниці сукупності об'єднані у невеликі групи чи серії, які розглядають як одне ціле.

Серії складаються з одиниць, які пов'язані або територіально (райони, селища), або організаційно (фірми, акціонерні товариства тощо). В якості серій можуть розглядатися упаковки з визначеною кількістю готової продукції, партії товару, студентські групи, творчі колективи та інше.

Вибіркова сукупність серій формується за схемами механічної або простої випадкової неповторної вибірки. Якщо серія потрапила до вибірки, то обстежують усі без винятку елементи серії, тобто проводять суцільне їх обстеження, а результати розповсюджують на всю сукупність. Цей спосіб відбору застосовують при перевірці якості продукції того чи іншого підприємства, роботи відділу, при атестації Вузу (у відібраних групах студентів роблять суцільне вимірювання залишкових знань) тощо.

Застосовуючи серійну вибірку, слід звернути увагу на те, що випадкова похибка цієї вибірки виходить дещо більшою, ніж при інших способах відбору. Щоб цьому запобігти необхідно збільшити обсяг вибірки.

У комбінованих вибірках, які є найбільш поширеними, поєднуються різні способи добору. На практиці можна комбінувати типову і серійну вибірки, коли серії відбирають у встановленому порядку з декількох типових груп. Можлива також комбінація серійного і простого випадкового доборів, при якій окремі одиниці відбирають всередині серії у випадковому порядку. Комбінують також суцільне і вибіркове спостереження, де вибіркове уточнює, підтверджує результати суцільного чи є його пробним варіантом.

Багатоступінчастий відбір характеризується тим, що на усіх ступенях, крім останнього, проводиться лише відбір, а спостереження досліджуваних одиниць відбувається тільки на останньому ступені. Одиниці відбору на попередніх ступенях – це зазвичай організаційні об'єднання одиниць спостереження, їх певні групи.

Отже, спочатку відбувається поступовий вибір з генеральної сукупності великих груп одиниць, потім груп, менших за обсягом і так доти, поки не оберуть ті одиниці, які підлягатимуть обстеженню. Ступенів може бути два, три і більше. На різних ступенях вибірки використовують різні одиниці і різні способи відбору. Наприклад, під час обстежень рівня знань студентів Вузів відбір населених пунктів, у яких є вищі навчальні заклади можна здійснювати типовим способом, Вузів – теж типовим чи випадковим, курсів, груп – серійним чи механічним, студентів – механічним чи випадковим.

При багатофазній вибірці на усіх фазах зберігається одна й та сама одиниця відбору, а кожна фаза відрізняється від попередньої обсягом вибірки та програмою обстеження. З генеральної сукупності формується первинна вибірка, а з первинної – підвибірка і т. д. На кожній наступній фазі обсяг підвибірки зменшується, а програма обстеження розширюється. Наприклад, під час дослідження ставлення студентів до організації навчального процесу в Інституті на 1-й фазі досліджують 25% від загальної чисельності студентів за короткою програмою (три запитання), на 2-й фазі – 15% за ширшою програмою, яка містить ще два додаткових запитання, на 3-й фазі – 5% за повною програмою з семи запитань.

Вибіркові оцінки кожної фази використовуються як додаткова інформація на наступних фазах, що підвищує точність результатів і є важливою властивістю багатофазної вибірки. Похибки багатоступінчастої та багатофазної вибірок розраховують

окремо для кожного ступеня чи фази.

В практиці статистичного дослідження іноді доводиться зустрічатися з невеликими за обсягом так званими малими вибірками. Під малою вибіркою розуміють таке вибіркове спостереження, чисельність одиниць якого варіює в межах від 5 до 30. Хоча загальний принцип вибіркового обстеження (зі збільшенням обсягу вибірки підвищується точність вибіркових даних) лишається в силі, іноді доводиться обмежуватися малою кількістю

спостережень. Часто така необхідність виникає при вибірковій перевірці якості продукції, в науково-дослідницькій роботі та в інших випадках.

### **3. Похибки вибірки**

Між характеристиками вибіркової сукупності та шуканими характеристиками (параметрами) генеральної сукупності, як правило, існують деякі розбіжності, які називають помилками. Загальна величина можливої помилки вибіркової характеристики складається з помилок двоякого роду: помилки реєстрації і помилки репрезентативності.

Помилки реєстрації властиві будь-якому статистичному спостереженню взагалі. Порівняно з суцільними спостереженнями небезпека виникнення помилок реєстрації у вибіркових спостереженнях повинна бути меншою завдяки більш кваліфікованому і ретельному проведенню вибірки.

Помилки репрезентативності виникають від того, що вибірка не точно відтворює склад генеральної сукупності оскільки є випадковою величиною, тому і оцінки характеристик є випадковими величинами, які можуть не співпадати з відповідними параметрами генеральної сукупності.

За причинами виникнення помилки репрезентативності поділяються на систематичні та випадкові.

Систематичні (тенденційні) помилки репрезентативності виникають при порушенні принципу випадковості відбору (упереджений вибір елементів, недосконалий вибір одиниці відбору тощо). Систематичні помилки мають односторонній, тенденційний напрям і призводять до зсунення (викривлення) результатів обстеження в той чи інший бік. Їх ще визначають як помилки зміщення (зсунення). Наприклад: якщо для визначення середньої успішності потоку студентів взяти вибірку студентів, які сидять на останніх партах, то середня успішність вибірки буде меншою за середню успішність всього потоку.

Якщо вибірку сформувати зі студентів, які сидять в перших рядах, то середня успішність у такій вибірці буде вищою від середньої успішності потоку. При такому нерівномірному відборі статистична оцінка успішності потоку буде заздалегідь занижена або завищена.

Систематичних помилок важливо уникати, оскільки непорушність – одна з вимог до будь-якої вибіркової оцінки.

Випадкові помилки репрезентативності зумовлені тим, що розподіл відібраної сукупності одиниць не досить точно відтворює розподіл одиниць генеральної сукупності.

На відміну від систематичних, випадкові помилки властиві вибіркового спостереженню та неминуче виникають і при додержанні принципу випадковості відбору.

Але вони не носять тенденційного характеру і не ведуть до зміщення оцінок.

Випадкові помилки репрезентативності залежать від таких факторів:

- способу відбору;
- обсягу вибірки – обернено пропорційна залежність – чим більше одиниць відібрано, тим менша помилка; але слід звернути увагу, що це

стосується лише репрезентативної вибірки!, у випадку зсуненої – навпаки – збільшення її обсягу збільшує ймовірність отримати помилкові висновки;

- ступеня варіації досліджуваної ознаки у генеральній сукупності – пряма залежність – чим більша варіація ознаки, тим більша у середньому і помилка вибірки.

Визначення можливої і фактично допущеної помилки вибірки (випадкової) має велике значення при застосуванні вибіркового методу. Величина помилки характеризує ступінь надійності результатів вибірки, є необхідною для оцінки параметрів генеральної сукупності.

Для однієї конкретної вибірки можна визначити:

- середню та граничну помилки;
- ймовірність того, що похибка вибірки не перевищить допустимого рівня;
- обсяг вибірки, який забезпечить потрібну точність результатів для взятої ймовірності.

### **?Питання для самоконтролю**

1. Що таке генеральна сукупність.
2. Що розуміють під вибіркою.
3. Які основні параметри генеральної сукупності.
4. Що таке вибіркоче спостереження.
5. Які розрізняють способи формування вибіркових сукупностей.
6. Що таке випадкова вибірка.
7. Які види відбору Ви знаєте.
8. Що розуміють під помилками.
9. Які види помилок Ви знаєте.
10. Вкажіть методи усунення помилок.