

**МІНІСТЕРСТВО ВНУТРІШНІХ СПРАВ УКРАЇНИ
ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ВНУТРІШНІХ СПРАВ
КРЕМЕНЧУЦЬКИЙ ЛЬОТНИЙ КОЛЕДЖ**

Циклова комісія економіки та управління

ТЕКСТ ЛЕКЦІЇ

з навчальної дисципліни «Статистика»

обов'язкових компонент

освітньо-професійної програми першого (бакалаврського) рівня вищої освіти

Логістика

за темою № 9 - Вибірковий метод

Харків 2021

ЗАТВЕРДЖЕНО

Науково-методичною радою
Харківського національного
університету внутрішніх справ
Протокол від 23.09.2021 № 8

СХВАЛЕНО

Методичною радою
Кременчуцького льотного
коледжу
Протокол від 22.09.2021 № 2

ПОГОДЖЕНО

Секцією науково-методичної ради
ХНУВС з гуманітарних та соціально-
економічних дисциплін
Протокол від 22.09.2021 № 8

Розглянуто на засіданні циклової комісії економіки та управління, протокол від 31.08.2021 № 1

Розробники: викладач циклової комісії економіки та управління, спеціаліст вищої категорії, викладач-методист Бондарець О.М.

Рецензенти:

1. Доктор економічних наук, професор кафедри бізнес адміністрування, маркетингу і туризму Кременчуцького національного університету ім. М. Остроградського – Дружиніна В.В.
2. Кандидат економічних наук, спеціаліст вищої категорії, викладач-методист, викладач циклової комісії управління та адміністрування КЛК ХНУВС - Пушкар О.І.

План лекції:

1. Суть вибіркового спостереження
2. Вибіркові оцінки середньої та частки
3. Визначення обсягу вибірки
4. Різновиди вибірок

Рекомендована література:

Основна

1. Горкавий В.К. Статистика: підручник. – К.: Алерта, 2020 – 644 с.
2. Карпенко Л. М. Статистика: навчальний посібник. – Одеса: ОРІДУ НАДУ, 2019. – 184 с.
3. Логунова Н. А. Статистика II : підручник. К. : Кондор-Видавництво, 2015. 340 с.
4. Мармоза А. Т. Практикум з теорії статистики : навч. посіб. К. : ЦУЛ, 2013. 484 с.
5. Мармоза А. Т. Теорія статистики : підручник. К. : ЦУЛ, 2013. 592 с.
6. Опря А. Т. Статистика (модульний варіант з програмованою формою контролю знань). К. : ЦУЛ, 2014. 536 с.
7. Теорія статистики : навч. посіб. / М. К. Шапочка, О. М. Маценко. Суми : Університетська книга, 2014. 312 с.

Додаткова

1. Бізнес-статистика : навч. посіб. / С. О. Матковський, О.С. Гринькевич, М. Л. Вдовин, О.М. Вільчинська, О. Р. Марець, О. З. Сорочак. Київ : Алерта, 2016. 281 с.
2. Ковтун Н. В. Теорія статистики : підручник. К. : Знання, 2012. 399 с.
3. Костюк В. О. Прикладна статистика : навч. посіб. Харків : ХНУМГ ім. О. М. Бекетова, 2015. 191 с.
4. Крамченко Л. І. Статистика ринку товарів та послуг : навч. посіб. Вид. 2-ге, переробл. і допов. Львів : Новий світ-2000, 2016. 296 с.
5. Кулинич О. І., Кулинич Р. О. Теорія статистики : підручник. К. : Знання, 2013. 239 с.
6. Моторин Р. М., Чекотовський Е. В. Статистика для економістів : навч. посіб. К. : Знання, 2013. 381 с.
7. Статистика підприємств / С. О. Матковський та ін. Львів : Алерта, 2013. 560 с.
8. Стегній М. І. Статистика : навч. посіб. К. : Кондор, 2012. 306 с.
9. Штагрет А. М. Статистика : навч. посіб. К. : ЦУЛ, 2012. 232 с.

Інформаційні ресурси в Інтернеті

1. Офіційний портал Верховної Ради України: Законодавство України [Електронний ресурс]. – Режим доступу: <https://zakon.rada.gov.ua>
2. Офіційний сайт Державної служби статистики України [Електронний ресурс]. – Режим доступу: www.ukrstat.gov.ua

Текст лекції:

1. Суть вибіркового спостереження

Вибірковим спостереженням називають вид несуттєвого спостереження, за характеристикою відібраної частини одиниць якої судять про всю сукупність.

Розрізняють генеральну і вибіркoву сукупності.

Генеральною сукупністю називають загальну масу одиниць, з якої здійснюють відбір для дослідження.

Вибірковою називають частину генеральної сукупності, яку відібрано для обстеження.

Обсяг генеральної сукупності позначають N , вибіркової - n .

До вибіркового спостереження вдаються у випадках, коли потрібно у стислі строки та з мінімальними затратами праці і коштів одержати кількісні характеристики досліджуваної сукупності, або коли неможна чи недоцільно здійснювати суцільне спостереження.

Відбір одиниць із генеральної сукупності у вибіркoву залежно від умов можна здійснювати по-різному: повторною і неповторною вибіркою.

Повторною називається вибірка, за якою кожна раніше відібрана одиниця повертається до генеральної сукупності і може повторно брати участь у вибірці.

Безповторною називається вибірка, за якою кожна раніше відібрана одиниця не повертається до генеральної сукупності і в подальшій вибірці участі не бере.

Безповторний відбір дає точніші результати.

2. Вибіркові оцінки середньої та частки

У разі вибіркового спостереження крім помилок реєстрації можливі так звані похибки репрезентативності.

Похибками репрезентативності називають розходження між середніми величинами або частками ознаки вибіркової і генеральної сукупностей.

Похибки репрезентативності можуть бути систематичними і випадковими.

Систематичні похибки виникають унаслідок порушення принципів проведення вибіркового спостереження.

Випадкові похибки репрезентативності зумовлені тим, що вибіркова

сукупність не відтворює точно середні і відносні показники генеральної сукупності.

Визначення величини випадкових похибок репрезентативності є одним з головних завдань теорії вибіркового методу.

Для визначення середньої похибки репрезентативності власне випадкової і механічної вибірки застосовують певні формули.

Середня похибка репрезентативності

<i>Спосіб відбору</i>	<i>Визначення середньої</i>	<i>Визначення частки</i>
Повторний	$\mu_x = \sqrt{\frac{\sigma^2}{n}}$	$\mu_p = \sqrt{\frac{w(1-w)}{n}}$
Без повторний	$\mu_x = \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)}$	$\mu_p = \sqrt{\frac{w(1-w)}{n} \left(1 - \frac{n}{N}\right)}$

де $\mu_{x,p}$ — середня похибка репрезентативності;

σ^2 - дисперсія;

n - вибірка сукупність;

N - генеральна сукупність;

W - частка одиниць з однакою ознакою.

Для узагальнюючої характеристики похибки вибірки поряд із середньою розраховують і граничну похибку вибірки.

Граничну похибку репрезентативності обчислюють з певною імовірністю, якій відповідає t - разове значення μ .

$\Delta = t\mu$,

де t - коефіцієнт довіри;

Δ - гранична похибка вибірки.

При ймовірності 0,683 $t=1$

0,954 $t=2$

0,997 $t=3$

0,999 $t=4$

Гранична похибка вибірки дає можливість встановити, в яких межах лежить значення генеральної середньої або частки.

$$\begin{aligned}\bar{X} = \bar{X} \pm \Delta &= \begin{cases} \bar{X}_1 \\ \bar{X}_2 \end{cases} \\ P = w \pm \Delta &= \begin{cases} P_1 \\ P_2 \end{cases}\end{aligned}$$

3. Визначення обсягу вибірки

Під час вибіркового спостереження важливо правильно визначити необхідну чисельність обсягу вибірки, яка з відповідною імовірністю забезпечує встановлену точність результатів спостереження.

Чисельність вибірки n

Спосіб відбору	Визначення середньої	Визначення частки
Повторний	$n = \frac{t^2 \sigma^2}{\Delta^2}$	$n = \frac{t^2 w(1-w)}{\Delta^2}$
Без повторний	$n = \frac{t^2 \sigma^2 N}{\Delta^2 N + t^2 \sigma^2}$	$n = \frac{t^2 w(1-w)N}{\Delta^2 N + t^2 w(1-w)}$

Приклад:

У районі 100 магазинів. Скільки магазинів потрібно відібрати власне випадковим способом для спостереження щоб гранична помилка з ймовірністю 0,997 ($t=3$) не перевищувала ± 10 осіб.

Із попередніх обстежень відомо, що середнє квадратне відхилення дорівнює ± 20 осіб.

N = 100
маг.
t=3
$\Delta = \pm 10$
$\Sigma = \pm 20$
n - ?

$$n = \frac{t^2 \sigma^2 N}{\Delta^2 N + t^2 \sigma^2} = \frac{3^2 \cdot 20^2 \cdot 100}{10^2 \cdot 100 + 3^2 \cdot 20^2} = 26 \text{ маг.}$$

Отже, для обстеження потрібно відібрати 26 магазинів.

4. Різновиди вибірок

Формування вибірки — не безладний процес. Ця дія виконується за певними правилами. Передусім визначається основа вибірки. У сукупностях, які складаються з «фізичних» елементів, одиниця основи може репрезентувати або окремий елемент сукупності, або певне їх угруповання. Наприклад,

вивчається використання комбайнів. Загальна їх кількість N розподілена за M бригадами, кожна з них має N_j комбайнів. Одиницею основи вибірки може бути комбайн або бригада. Відповідно формується вибіркова сукупність: у першому випадку вибирається n комбайнів із загального їх числа N , у другому — m бригад із загального їх числа M .

Найпростішою основою вибірки є перелік елементів генеральної сукупності, пронумерований від 1 до N . Простими вважаються також набори звітів, анкет, карток тощо.

На практиці досліджувані сукупності мають, як правило, не одну, а низку альтернативних основ для вибірки. Наукове обґрунтування та правильний вибір основи — перша передумова забезпечення репрезентативності результатів вибіркового спостереження.

Від основи вибірки залежить спосіб добору елементів сукупності для обстеження. Найчастіше використовують способи добору: простий випадковий, механічний, розшарований (районований), серійний.

Простий випадковий добір провадиться жеребкуванням або за допомогою таблиць випадкових чисел. Це класичний спосіб формування вибіркової сукупності, який передбачає попередню досить складну підготовку до формування вибірки. Для жеребкування на кожну одиницю генеральної сукупності необхідно заготувати відповідну фішку; при використанні таблиць випадкових чисел усі елементи цієї сукупності мають бути пронумеровані. У великих за обсягом сукупностях така робота здебільшого недоцільна, а часом і неможлива. Тому на практиці застосовуються інші різновиди випадкових вибірок.

Механічний добір. Основа вибірки — упорядкована множина елементів сукупності. Добір елементів здійснюється через рівні інтервали. Крок інтервалу обчислюється діленням обсягу сукупності N на передбачений обсяг вибірки n . Початковий елемент вибірки визначається як випадкове число всередині першого інтервалу, другий елемент залежить від початкового числа й кроку

інтервалу. Так, для частки вибірки $\frac{n}{N} = 0,05$ кроком інтервалу є число $\frac{n}{N} = \frac{1}{0,05} = 20$, тобто у вибірку має потрапити кожний двадцятий елемент. Якщо початковий елемент — випадкове число 7, то другим елементом буде $7 + 20 = 27$, третім — $27 + 20 = 47$ і т. д.

Механічна вибірка порівняно з простою випадковою ефективніша, її простіше здійснити. Проте за наявності циклічних коливань значень ознаки, цикл коливань яких збігається з інтервалом, можливий зсув вибірових оцінок. Похибку механічної вибірки обчислюють за формулою похибки неповторної вибірки.

Вивчаючи безперервні в часі процеси, зокрема технологічні (структури затрат робочого часу, використання виробничого устаткування), проводять **моментні спостереження**. Суть їх — у періодичній фіксації стану процесу на певні моменти часу, які вибирають за схемою випадкової або механічної вибірки (через певні інтервали часу).

На етапі підготовки моментних спостережень визначають перелік можливих варіантів стану процесу, наприклад перелік причин простоїв устаткування. Під час обстеження певної сукупності одиниць устаткування, скажімо, верстатів, у визначені моменти часу фіксується, працює r -й верстат чи ні (якщо ні, зазначаються причини простою). Припустимо, що в цеху працюють 10 верстатів і за 8-годинну зміну через кожні півгодини проводилась реєстрація використання цих верстатів. Було зроблено 160 записів ($2 \cdot 8 \cdot 10$), у 144 випадках зазначено, що верстат працював, у 16 — не працював. Частка працюючих верстатів становить 0,9, дисперсія частки — $0,9 \cdot 0,1 = 0,09$. Із

імовірністю 0,954 гранична похибка вибірки $\Delta = 2 \sqrt{\frac{0,09}{160}} = 0,046$, або 4,6%. Отже, частка працюючих верстатів за зміну становила не менш як $90 - 4,6 = 85,4\%$.

Щодо повноти охоплення елементів сукупності, то моментне спостереження суцільне, воно вибіркове впродовж часу, бо охоплює не весь час роботи устаткування, а лише певні моменти. У разі правильної організації моментні обстеження забезпечують досить точні результати швидко і з меншими витратами, ніж при суцільному спостереженні.

Розшарований (районований, типовий) добір — це спосіб формування вибірки з урахуванням структури генеральної сукупності. На відміну від простого випадкового та механічного добору, які проводяться в цілому по генеральній сукупності, розшарований передбачає її попередню структурування й незалежний добір елементів у кожній складовій. Обсягом

розшарованої вибірки є сума частинних вибірок n_j , тобто $n = \sum_j^m n_j$, де m — число складових (груп, типових районів тощо).

Похибку розшарованої вибірки обчислюють, використовуючи середню з групових дисперсій $\overline{\sigma^2}$. Якщо сформовані групи об'єднують «схожі» елементи, а групові середні величини помітно різні, варіація ознаки в групах буде значно меншою, ніж по сукупності. У такому разі $\overline{\sigma^2} < \sigma^2$, а отже, похибка розшарованої вибірки порівняно з простою випадковою чи механічною буде менша:

$$\Delta = t \sqrt{\frac{\overline{\sigma^2}}{n} \left(1 - \frac{n}{N} \right)}.$$

Для того щоб забезпечити більшу точність розшарованої вибірки, слід обґрунтувати ознаку розшарування сукупності, число складових частин m , обсяг частинних вибірок n_j і спосіб добору. Зменшення варіації ознаки при розшаруванні сукупності можливе за умови, що ознака розшарування сукупності корелює з ознакою, характеристики якої оцінюються. Ці ознаки співвідносяться як причина й наслідок.

Відповідно до правила розкладання дисперсій $\overline{\sigma^2} = \sigma^2 - \delta^2$ або $\overline{\sigma^2} = \sigma^2 (1 - \eta^2)$. Отже, розшарування сукупності зменшує похибку вибірки на частку $(1 - \eta^2)$. Чим щільніший зв'язок між ознаками, тим помітніше

зменшення похибки. При $n^2 = 0,50$ похибка вибірки зменшується вдвічі, при $n^2 = 0,66$ — утричі.

У практиці вибірових спостережень застосовують різні способи визначення обсягу вибіркової сукупності n та її складових n_j . Найпростіший з них, коли всі m груп подані однаковою кількістю елементів:

$$n_j = \frac{n}{m}.$$

Проте застосування цього способу обмежене. Якщо чисельності груп у генеральній сукупності N_j дуже різні, може виникнути ситуація, коли $n_j > N_j$.

Найчастіше застосовують пропорційний добір, який передбачає однакове для всіх складових представництво, тобто частки $\frac{n_j}{N_j}$ однакові й обсяг частинної вибірки залежить від обсягу відповідної складової сукупності:

$$n_j = \frac{n_j}{N_j} N_j.$$

Оптимальним щодо мінімізації похибки є добір, пропорційний до середнього квадратичного відхилення:

$$n_j = \frac{N_j \sigma_j}{\sum_{j=1}^m N_j \sigma_j} n.$$

Очевидно, що обсяг вибірки залежить від рівня варіації ознаки в окремих складових генеральної сукупності. Однорідні групи подаються меншим числом елементів, неоднорідні — більшим. Відсутність даних про варіацію ускладнює практичну реалізацію такого способу вибірки.

Різновидом розшарованої вибірки є **метод квот**, коли обсяг частинних вибірок n_j визначається завчасно. Цей спосіб поширений при вивченні громадської думки, ринку тощо. Так, при вивченні громадської думки тому, хто має брати інтерв'ю, установлюються квоти, наприклад обстежити двох фермерів-чоловіків віком 30—40 років, трьох мешканців міста віком 20—30 років і т. ін. В який спосіб «заповнити квоти», він вирішує сам. Метод квот не гарантує незсуненості вибірових оцінок.

Серійна вибірка. Одиниця основи вибірки — серія елементів. Серії складаються з одиниць, які пов'язані або територіально (райони, селища), або організаційно (фірми, акціонерні товариства). Вибіркова сукупність серій формується за схемами механічної або простої випадкової вибірки. Дібрана серія розглядається як одне ціле, обстеженню підлягають усі без винятку елементи серії. При обчисленні похибки вибірки враховується міжсерійна варіація:

$$\Delta = t \sqrt{\frac{\delta^2}{m} \left(1 - \frac{m}{M}\right)},$$

де δ^2 — міжсерійна дисперсія; m та M — число серій відповідно у вибірці та генеральній сукупності.

Похибка серійної вибірки буде меншою порівняно з похибкою простої випадкової чи механічної вибірки в тому разі, якщо серії більш-менш однорідні й варіація серійних середніх незначна. Зростання міжсерійної варіації призводить до збільшення похибки вибірки.

Використання того чи іншого способу формування вибіркової сукупності залежить від мети вибіркового обстеження, можливостей його організації та проведення. Іноді поєднуються різні способи добору: механічний і серійний, розшарований і механічний, випадковий і серійний.

Таке поєднання можливе в рамках **багатоступеневої вибірки**. Ступенів може бути два, три й більше. Кожний із них має свою, відмінну від інших основу вибірки. Відповідно поділяються й одиниці вибірки: першого ступеня, другого і т. ін. Повнота охоплення основи й схема добору одиниць на різних ступенях різняться.

Наприклад, сукупність містить K одиниць першого ступеня, які складаються з M одиниць другого ступеня, ті, у свою чергу, об'єднують N_j одиниць третього ступеня. Саме така триступенева вибірка застосовується при організації обстеження домогосподарств. Наприклад, у сільській місцевості одиниці першого ступеня — це райони області; одиниці другого ступеня — селища; одиниці третього ступеня — домогосподарства.

Отже, вибір елементів для безпосереднього обстеження здійснюється на останньому, третьому ступені формування вибіркової сукупності. Частка її відносно до генеральної сукупності залежить від часток вибірки на всіх ступенях. Якщо припустити, що до вибірки потрапив один з десяти районів ($d_1 = 0,10$), у цих районах відібране кожне п'яте селище ($d_2 = 0,20$), а у відібраних селищах обстежується 4% домогосподарств ($d_3 = 0,04$), то частка вибіркової сукупності в генеральній становить:

$$d = d_1 d_2 d_3 = 0,1 \cdot 0,2 \cdot 0,04 = 0,0008,$$

тобто обстеженню підлягає 0,08% домогосподарств.

Багатоступенева вибірка значно зменшує витрати на обстеження й порівняно з іншими вибірками більш ефективна.

Якщо обстежують сукупність за двома й більше ознаками, які різняться варіацією, ефективною є **багатофазна вибірка**. Суть її в тому, що для різних ознак формуються вибіркові сукупності різного обсягу. На відміну від багатоступеневої вибірки багатофазна використовує для всіх ознак одну й ту саму основу вибірки, проте програма обстеження різна.

Вибіркові сукупності формуються поетапно — фазами. З генеральної сукупності утворюється первинна вибірка, а з первинної — підвибірка і т. д. На кожній наступній фазі обсяг підвибірки зменшується, а програма обстеження розширюється. Вибіркові оцінки кожної фази використовуються як додаткова інформація на наступних фазах, що підвищує точність результатів вибіркового обстеження. При організації багатофазної вибірки можливі комбінації різних способів і видів вибірки. Багатофазна вибірка поєднується з багатоступеневою, а також із суцільним спостереженням.