

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЛЬВІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМ. І.ФРАНКА

МЕТОДИЧНІ ВКАЗІВКИ
до лабораторної роботи з курсу
"Статистика"

ВИБІРКОВІ СПОСТЕРЕЖЕННЯ

Львів-2003

Затверджено на засіданні
кафедри статистики

Протокол № 10 від 14 травня 2003 р.

Вибіркові спостереження. Методичні вказівки до лабораторної роботи з курсу "Статистика" для студентів базових напрямків: 0501 "Економіка і підприємництво", 0502 "Менеджмент"/ Укл. О.С.Гринькевич, О.З.Сорочак. - Львів: ЛНУ ім. І. Франка, 2003. - 22 с.

Укладачі

Гринькевич О.С., канд. екон. наук
Сорочак О.З., канд. техн. наук, доц.

**Відповідальний
за випуск**

Матковський С.О., канд. екон. наук, проф.,

Рецензенти

Матковський С.О., канд. екон. наук, проф.,
Войцеховська Ю.В., канд. екон. наук, доц.

1. ЗАГАЛЬНІ ПОЛОЖЕННЯ

Мета даної роботи - набуття студентами практичних навичок формування різних видів вибірок, розрахунку необхідного обсягу вибірки та оцінювання параметрів генеральної сукупності за даними вибірки з використанням для вирішення таких задач „Пакету аналізу даних”, включеного в електронний табличний редактор „Microsoft Excel”.

Виконуючи завдання студенти практично освоюють способи формування основи вибірки, вчать правильно організовувати вибіркові обстеження.

2. ТЕОРЕТИЧНІ ПОЛОЖЕННЯ

Вибіркове обстеження – це найбільш поширений вид несучільного обстеження. Його застосовують при вивченні різноманітних явищ і процесів, наприклад, таких як використання устаткування в промисловості, втрати при збиранні урожаю, попит населення і ступінь його задоволення торгівлею, обстеження сімейних бюджетів населення, вивчення громадської думки тощо.

Вибірковий метод - це система наукових принципів випадкового відбору певної частини сукупності, яка представляла б усю сукупність і характеристики якої слугували б надійною основою статистичного висновку.

Сукупність, з якої відбираються елементи для обстеження, називають *генеральною*, а сукупність, яку безпосередньо обстежують, - *вибірковою*. Статистичні характеристики вибіркової сукупності розглядаються як оцінки відповідних характеристик генеральної сукупності. Оскільки вибіркова сукупність не точно відтворює структуру генеральної, то вибіркові оцінки також не збігаються з характеристиками генеральної сукупності. Розбіжності між ними називають *похибками репрезентативності*. За причинами виникнення ці похибки поділяються на *систематичні (тенденційні)* та *випадкові*. *Систематичні похибки* виникають за умови, що під час формування вибіркової сукупності порушується принцип випадковості відбору (упереджений відбір елементів, недосконала основа вибірки тощо). Для всіх елементів сукупності вони мають односторонній напрям і тому їх ще називають *похибками зміщення*.

Випадкові похибки – це наслідок випадковості відбору елементів сукупності для обстеження.

При організації вибіркового обстеження важливо запобігти виникненню систематичних похибок. Що стосується випадкових похибок, то уникнути їх неможливо, проте на основі теорії вибіркового методу можна визначити їх розмір і по можливості регулювати.

Порівняно із суцільним обстеженням вибіркоче має переваги, оскільки обстеження за однією і тією самою програмою будь-якої частини сукупності потребує менше коштів і часу, ніж обстеження сукупності в цілому. Крім того, дешевше обходиться і статистична обробка інформації вибіркового обстеження. Вибірка застосовується і у випадках, коли суцільне обстеження взагалі неможливе: обстежувана сукупність дуже велика, практично безмежна (наприклад, сукупність ділянок морського дна, або сукупність колосків пшениці на полі); обстеження пов'язане із знищенням або псуванням одиниць, які досліджуються (наприклад, контроль якості хлібобулочних виробів, консервів і т.д.). Поряд з тим при вивченні великих сукупностей вибіркоче обстеження може дати більш точні результати, ніж суцільне, за рахунок зменшення кількості похибок реєстрації.

При достатньо великому обсязі вибіркової сукупності n ймовірність того, що похибка вибірки Δ не перевищує свого граничного значення $\Delta = t\mu$ за теоремою Ляпунова дорівнює інтегралу Лапласа:

$$P(\Delta \leq t\mu) = \frac{1}{\sqrt{2\pi}} \int_{-t}^{+t} e^{-\frac{t^2}{2}} dt = \Phi(t) \quad (1)$$

де μ - середньоквадратична, або стандартна похибка вибірки; t – квантиль розподілу ймовірностей (довірче число для заданого значення ймовірності P), який визначається за табличними значеннями функції $\Phi(t)$.

Згідно таблиці значень функції $\Phi(t)$

при

$$t = 1$$

$$P(\Delta \leq \mu) = 0,683,$$

при

$$t = 2$$

$$P(\Delta \leq 2\mu) = 0,954,$$

при

при $t = 3$ $P(\Delta \leq 3\mu) = 0,9973$,
 $t = 4$ $P(\Delta \leq 4\mu) = 0,99993$.

Наведені дані вказують на те, що практично неможливо отримати похибку вибірки більшу, ніж 3μ .

При малих вибірках ($n \leq 30$) квантиль t визначають за розподілом імовірностей Стюдента (псевдонім англійського математика В. Госета). Малі вибірки не знаходять широкого застосування в економіці, тільки в деяких випадках вони використовуються в технічному нормуванні, контролі якості продукції, вибіркових фотографіях робочого дня тощо.

Найчастіше вживані значення t наведені в наступній таблиці:

Таблиця 1

Ймовірності розподілу похибок вибірки

Значення t	Розподіл		
	нормальний	Стюдента для n	
		8-10	11-15
1,6	0,890	0,928	0,933
1,7	0,911	0,938	0,943
1,8	0,928	0,947	0,952
1,9	0,942	0,955	0,960
2,0	0,954	0,962	0,969
2,1	0,964	0,967	0,972
2,2	0,972	0,972	0,977
2,3	0,979	0,977	0,981
2,4	0,983	0,980	0,984
2,5	0,987	0,983	0,987
3,0	0,997	0,993	0,995

Стандартна похибка вибірки μ є середнім квадратичним відхиленням вибіркових точкових оцінок: вибіркової середньої \tilde{x} або вибіркової частки ω від значень відповідних параметрів генеральної сукупності.

При відборі за схемою поверненої кулі – повторному відборі

$$\mu = \sqrt{\frac{\sigma^2}{n-1}}, \quad (2)$$

при відборі за схемою неповерненої кулі – безповторному

$$\mu = \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)}, \quad (3)$$

де σ^2 - вибіркова дисперсія; n та N – відповідно обсяг вибіркової та генеральної сукупностей.

При практичному використанні наведених формул слід врахувати, що:

- 1) дисперсія частки є добутком часток $\sigma^2 = p(1-p) = pq$;
- 2) у великих за обсягом вибірових сукупностях ($n \geq 30$) поправка $\frac{n}{n-1}$ не вносить істотних змін у розрахунки, а тому враховується

лише у малих вибірках (тобто $\mu = \sqrt{\frac{\sigma^2}{n}}$);

- 3) коригуючий множник для безповторної вибірки $\sqrt{\left(1 - \frac{n}{N}\right)}$ при малих величинах n/N наближається до 1, а тому при 1...5%-ій вибірці ($n/N = 0,01 \dots 0,05$) розрахунок μ ведеться за формулою для повторної вибірки.

Гранична похибка вибірки ($\Delta = t\mu$) - це максимально можлива похибка для прийнятої ймовірності P . Довірче число t вказує, як співвідносяться гранична та стандартна похибки. Отже, розмір граничної похибки середньої величини і частки обчислюється за такими формулами для відбору:

- повторного

$$\Delta_{\bar{x}} = t\sqrt{\frac{\sigma^2}{n}} \quad i \quad \Delta_p = t\sqrt{\frac{pq}{n}}, \quad (4)$$

- безповторного

$$\Delta_{\bar{x}} = t \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)} \quad \text{і} \quad \Delta_p = t \sqrt{\frac{pq}{n} \left(1 - \frac{n}{N}\right)}. \quad (5)$$

Як видно з формул, розмір граничної похибки залежить від варіації ознаки в генеральній сукупності D (дисперсії), яку приймають рівною вибірковій σ^2 , обсягу вибірки n та її частки у генеральній сукупності n/N , а також прийнятого рівня ймовірності, якому відповідає квантиль t .

Виходячи з того, що гранична похибка вибірки для ймовірності P є максимальним відхиленням розміру значень вибіркової оцінки від характеристики генеральної сукупності, можливі межі значень останньої визначають так:

$$\bar{x} = \tilde{x} \pm \Delta_{\bar{x}} \text{ - для середньої величини;}$$

$$p = \omega \pm \Delta_p \text{ - для частки,}$$

де \bar{x} і \tilde{x} – значення середньої величини ознаки відповідно генеральної і вибіркової сукупностей; p і ω - частки елементів відповідно генеральної і вибіркової сукупностей, які мають певні значення ознаки.

Формування репрезентативної вибірки (тобто такої, що правильно характеризує генеральну сукупність) – це не безладний процес; він здійснюється за певними правилами. В залежності від основи вибірки у практиці вибірових обстежень використовують різні способи формування вибірових сукупностей, зокрема: простий випадковий, механічний (систематичний), розшарований (районований, типовий, стратифікований), серійний.

Простий випадковий відбір проводиться жеребкуванням або на основі таблиць випадкових чисел. Це класичний спосіб формування вибіркової сукупності і саме на ньому ґрунтується теорія вибіркового методу. Він передбачає попередню досить складну підготовку до формування вибірки. Так, для жереба на кожну одиницю відбору треба заготовити відповідну фішку; при використанні таблиць випадкових чисел всі елементи генеральної сукупності мають бути пронумеровані. У великих за обсягом сукупностях така робота здебільшого недоцільна, а часом і неможлива. Тому на практиці застосовують інші різновиди випадкових вибірок.

Механічний (систематичний) відбір передбачає, що основою вибірки є впорядкована чисельність елементів генеральної сукупності. Відбір елементів здійснюється через однакові інтервали, крок інтервалу

залежить від частки вибірки $D=n/N$ і визначається як $1/D=N/n$. Початковий елемент відбору визначають як випадкове число всередині першого інтервалу жеребкування (іноді розпочинають з середини інтервалу), другий елемент залежить від початкового числа і кроку інтервалу. Так для частки вибірки $D=0,05$ кроком інтервалу є число $1/0,05=20$. Тобто у вибірку має попасти кожний двадцятий елемент. Якщо початковий елемент – випадкове число 3, то другий елемент становить $3+20=23$, третій – $23+20=43$ і т.д.

Механічний відбір особливо зручний в тих випадках, коли вже є списки одиниць сукупності, складені в тому чи іншому порядку, а також в тих випадках, коли мають справу з генеральною сукупністю, чисельність якої відома лише наближено, одиниці якої з'являються поступово на протязі деякого періоду, наприклад, при контролі якості продукції, що виготовляється. Похибка вибірки при механічному відборі визначається за формулою простої випадкової неповторної вибірки. Проте за наявності циклічних коливань значень ознаки, цикл коливань яких збігається з інтервалом, можливе зміщення похибки.

Розширений (районований, типовий, стратифікований) відбір передбачає попередню структурування генеральної сукупності та незалежний відбір елементів у кожній складовій частині. Обсяг

розширеної вибірки – це сума частинних вибірок n_j , тобто $n = \sum_{j=1}^m n_j$, де

m число складових частин (груп, типів, районів, страт тощо).

При обчисленні похибки розширеної вибірки використовують середню з групових дисперсій

$$\overline{\sigma_j^2} = \frac{\sum_{j=1}^m \sigma_j^2 n_j}{\sum_{j=1}^m n_j}. \quad (6)$$

А відповідно до правила складання дисперсій $\overline{\sigma_j^2} = \sigma^2 - \delta^2$, або $\overline{\sigma_j^2} = \sigma^2(1 - \eta^2)$, де $\eta^2 = \delta^2 / \sigma^2$ - кореляційне відношення, яке визначає щільність зв'язку між ознаками, що відображають причину і наслідок.

Таким чином, розшарування (типізація) сукупності зменшує похибку вибірки на частку $1-\eta^2$. Чим щільніший зв'язок між ознаками, тим помітніше зменшення похибки вибірки.

У практиці вибіркових обстежень застосовують різні способи визначення обсягу частинних вибірок n_j . Найпростішим із них, коли всі m груп представлені однаковою кількістю елементів, є такий:

$$n_j = \frac{n}{m}. \quad (7)$$

Проте застосування цього способу обмежене. Якщо чисельності груп N_j дуже відрізняються, може виникнути ситуація, коли $n_j > N_j$ ($N=N_1+N_2+\dots+N_j+\dots+N_m$).

Найчастіше використовують пропорційний відбір, який передбачає однакове для всіх складових частин представництво у вибірці. Обсяг частинних вибірок у цьому випадку визначається як

$$n_j = DN_j, \quad (8)$$

де $D=n/N$ - частка вибірки.

Оптимальним щодо мінімізації похибки вибірки при розшарованому відборі є відбір пропорційний як чисельності групи N_j так і середньому квадратичному відхиленню ознаки, яка вивчається у цій групі σ_j . Обсяг частинних вибірок визначається при цьому за формулою

$$n_j = \frac{\sigma_j N_j}{\sum_{j=1}^m \sigma_j N_j} n. \quad (9)$$

Очевидно, що точні розрахунки оптимального розподілу обсягу вибірки між групами практично неможливі, так як до проведення досліджень про величини σ_j можна мати лише наближені уявлення. Тому на практиці, коли є підстави вважати, що в тій чи іншій групі варіація ознаки, яку вивчають, більша, ніж у решти груп, цій групі „на око” дещо збільшують чисельність вибірки.

Похибку розшарованої вибірки обчислюють за формулою

$$\Delta = t \sqrt{\frac{\sigma_j^2}{n} \left(1 - \frac{n}{N}\right)}. \quad (10)$$

При *серійному відборі* основа вибірки складається з серій елементів сукупності, зв'язаних територіально (райони, селища), організаційно (фірми, акціонерні товариства) тощо. Серії відбираються за схемою механічної або простої випадкової вибірки, обстеженню підлягають всі елементи серії. При обчисленні похибки вибірки враховується міжсерійна варіація:

$$\delta^2 = \frac{\sum_{k=1}^m (\bar{x}_k - \bar{x})^2 n_k}{\sum_{k=1}^m n_k}, \quad (11)$$

де n_k та \bar{x}_k - відповідно обсяг і середня k -тої серії.

Похибка серійної вибірки визначається за формулою:

$$\Delta = t \sqrt{\frac{\delta^2}{s} \left(1 - \frac{s}{S}\right)}, \quad (12)$$

де S – загальна кількість серій у генеральній сукупності; s – число відібраних серій для обстеження. Ця похибка буде меншою порівняно з похибкою простої випадкової чи механічної вибірки в тому разі, якщо серії більш-менш однорідні і варіація серійних середніх незначна.

Застосування того чи іншого способу формування вибіркової сукупності залежить від мети вибіркового обстеження, можливостей його організації і проведення. Найбільш поширеними є комбіновані вибірки, які поєднують різні способи відбору: механічний і серійний, розшарований і механічний, простий випадковий і серійний. Поєднання способів відбору забезпечує високу репрезентативність результатів з найменшими трудовими і грошовими витратами на організацію та проведення досліджень.

Як же визначити необхідний обсяг вибірки?

Визначення необхідного обсягу вибірки – це перша задача, з якою стикається дослідник, що організовує вибіркоче обстеження. Практично

при цьому йому точно відомі тільки мета вибірки і величина похибки Δ , яку він вважав би несуттєвою. Сам по собі розрахунок чисельності вибірки на основі допустимої величини похибки нескладний, так як, задавшись довірчою ймовірністю, з якою необхідно забезпечити певну точність Δ , можна перейти до величини стандартної похибки відбірки μ , а потім і до чисельності вибірки.

Проектуючи вибіркові обстеження, визначають мінімально достатній обсяг вибірки, за якого вибіркові оцінки репрезентували б основні властивості генеральної сукупності з заданою точністю.

При *повторному* відборі $\mu = \sqrt{\frac{\sigma^2}{n}}$, а $\Delta = t\mu$. Визначивши $\mu = \Delta/t$, підставимо його значення у формулу стандартної похибки вибірки і виразимо з неї n :

$$n = \frac{t^2 \sigma^2}{\Delta^2}, \text{ аналогічно для частки } n = \frac{t^2 pq}{\Delta^2}. \quad (13)$$

При *безповторному* відборі $\mu = \sqrt{\frac{\sigma^2}{n} \left(1 - \frac{n}{N}\right)}$. Замінивши μ на Δ/t у формулі стандартної похибки вибірки і виразивши з неї n , отримаємо – для середньої величини

$$n = \frac{t^2 \sigma^2 N}{\Delta^2 N + t^2 \sigma^2}, \quad (14)$$

– для частки

$$n = \frac{t^2 pq N}{\Delta^2 N + t^2 pq}. \quad (15)$$

Для визначення обсягу вибірки n використовують оцінки дисперсій σ^2 аналогічних пробних обстежень. Якщо такі обстеження відсутні, то можна скористатися співвідношенням

$$\sigma = \frac{x_{\max} - x_{\min}}{6} = \frac{R}{6}, \quad (16)$$

а для частки взяти найбільше значення дисперсії $\sigma^2 = pq = 0,5 \cdot 0,5 = 0,25$.

Якщо в основу розрахунку n покласти відносну граничну похибку вибірки $V_{\Delta} = \frac{\Delta}{x} = \frac{t\mu}{x}$, яку ще можна виразити як $V_{\Delta} = \frac{tV_{\sigma}}{\sqrt{n}}$, де $V_{\sigma} = \frac{\sigma}{x}$ - коефіцієнт варіації ознаки x , то формули обсягу вибірки відповідно модифікуються:

– для середньої величини

$$n = \frac{t^2 V_{\sigma}^2}{V_{\Delta}^2}, \quad (17)$$

– для частки

$$n = \frac{t^2 q}{V_{\Delta}^2 p}, \text{ де } q = 1-p. \quad (18)$$

У практиці вибірових обстежень одночасно вивчають декілька ознак. Якщо бажану ступінь точності визначити для кожної ознаки окремо, то результатом розрахунків буде низка значень обсягу вибірки. З метою їх узгодження використовують або максимальний обсяг n (тоді решта ознак оцінюється „надто точно”), або обсяг головної ознаки.

Отримані в результаті вибіркового спостереження оцінки поширюються на генеральну сукупність. Якщо вибіркоче спостереження проводиться з метою визначення розміру ознаки у генеральній сукупності, і відома лише чисельність одиниць цієї сукупності, то застосовується *прямий перерахунок* даних вибірки на всю сукупність. Наприклад, якщо середній товарооборот одного магазину у простій випадковій вибірці 350 тис. грн., а генеральна сукупність складається з 1500 однорідних магазинів, то можна стверджувати, що загальний товарооборот всієї сукупності магазинів буде складати приблизно 525 млн. грн. (350 тис. грн. \times 1500) Однак гранична похибка вибірки вказує на можливість відхилення генеральної середньої від вибіркової середньої. Припустимо, що з імовірністю 0,954 у даному прикладі гранична похибка становить 15 тис. грн. Вона визначає межі генеральної середньої, тобто інтервал довір'я: $335 \leq \bar{X} \leq 365$. Враховуючи цей інтервал, з імовірністю 0,954 можна стверджувати, що загальний товарооборот у генеральній

сукупності магазинів буде знаходитись в інтервалі від 502,5 млн. грн. до 547,5 млн. грн.

Вибіркове спостереження може проводитись з метою уточнення результатів суцільного спостереження. У цьому випадку характеристики суцільного спостереження можуть бути точнішими завдяки тому, що вони коректуються за результатами додатково проведеного вибіркового спостереження. Уточнення даних суцільного спостереження за даними вибірки застосовується, зокрема, при переписах сільськогосподарських тварин.

Якщо, наприклад, суцільні дані перепису тварин у господарствах, що потрапили у вибірку, показали 150 голів коней, а вибіркові дані по тих же господарствах - 153, то можна визначити процент "неповного обліку" (недообліку) тварин при суцільному спостереженні. Він становитиме $\frac{3}{150}$, або 2%. Отже, необхідно поширити цей процент недообліку на всю обстежену під час суцільного спостереження сукупність. Припустимо, що в результаті перепису у даному регіоні під час суцільної реєстрації було записано 3500 коней. Слід зробити поправку на 2%, і чисельність коней в даному регіоні буде становити 3570 голів.

Таким чином, поширення вибірових даних на всю генеральну сукупність проводиться методом прямого перерахунку і методом коефіцієнтів. Прямий перерахунок використовується у випадку, коли розмір генеральної сукупності визначається лише на основі характеристик вибіркової сукупності. Метод коефіцієнтів застосовується тоді, коли вибіркове спостереження проводиться з метою уточнення результатів суцільного спостереження і на основі цього визначаються параметри розміри генеральної сукупності.

Поширити вибіркові дані на всю сукупність у другому випадку можна і за допомогою знаходження рівняння зв'язку між вибірковими даними і даними суцільного спостереження. Знайшовши таке рівняння і підставивши в нього відповідні вибіркові оцінки, отримаємо необхідні характеристики генеральної сукупності. Але таке поширення пов'язане з великою кількістю розрахунків і на практиці застосовується рідко.

3. ПОРЯДОК ВИКОНАННЯ РОБОТИ

Робота виконується в лабораторії ПЕОМ з використанням електронних таблиць „Microsoft Excel”. Комп’ютер вмикається тільки з дозволу викладача. Вихідні дані про генеральну сукупність містяться у файлі „Вибіркові обстеження.xls” на листі „Генеральна_сукупність” Для формування вибірових сукупностей методом простого випадкового і механічного відборів використовується інструмент Вибірка (Sampling). Інструмент Вибірка є елементом надбудови Аналіз даних (Data Analysis) і розміщується в меню Сервіс (Tools). Якщо ця надбудова відсутня в меню Сервіс (Tools), виберіть команду Надбудови (Add-Ins) і вкажіть Аналіз даних (Data Analysis) в списку надбудов, та додайте цей компонент.

Після того як визначений обсяг вибірки для простого випадкового і механічного відборів, виберіть в меню Сервіс (Tools) команду Аналіз даних (Data Analysis). Відкриється вікно діалогу, яке показано на рис. 1.

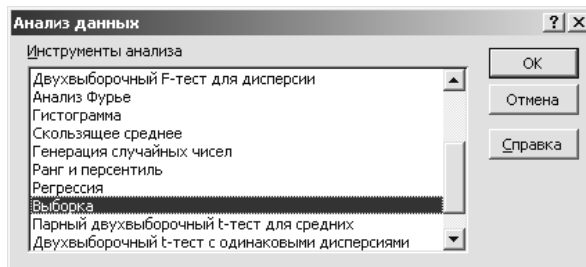


Рис. 1. Вікно діалогу Аналіз даних для постановки задачі.

В цьому вікні слід вказати необхідний інструмент, а саме Вибірка (Sampling), та натиснути ОК. На рис. 2. показано, як виглядає вікно інструменту Вибірка (Sampling), який дозволяє вилучити випадкову чи періодичну підмножину (вбіркову сукупність) із вихідної множини чисел (генеральної сукупності). У комірці вхідного інтервалу необхідно вказати посилання на діапазон порядкових номерів магазинів $\$A\$2:\$A\2291 , у методі відбору відмітити „Випадковий” та вказати обсяг вибірки (число вибірок). В параметрах виводу рекомендується вибрати “вихідний інтервал” - $\$G\2 та натиснути ОК.

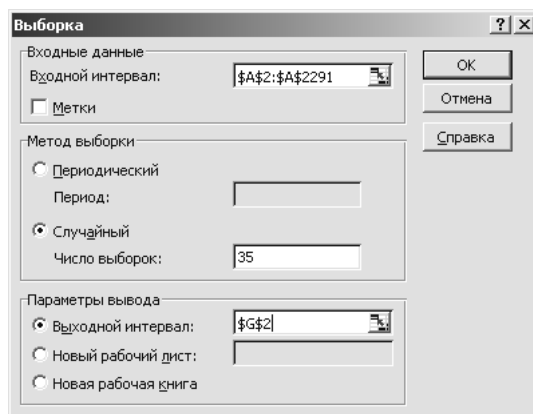


Рис. 2. Ввід параметрів у вікні інструменту Вибірка.

Вилучені номери магазинів скопіювати та вставити на лист „Вибірка”, починаючи з комірки A2. Після чого формули записані в комірках B2:D2:

B2=ИНДЕКС(Генеральна_сукупність!\$A\$2:\$D\$2291;A2;2);

C2=ИНДЕКС(Генеральна_сукупність!\$A\$2:\$D\$2291;A2;3);

D2=ИНДЕКС(Генеральна_сукупність!\$A\$2:\$D\$2291;A2;4);

скопіювати донизу на весь діапазон вибіркової сукупності, щоб вибрати значення „кількості зайнятих”, „товарообороту” і „витрат обігу” для порядкових номерів магазинів, які попали у вибірку.

Щоб позбутись залежності даних від генератора випадкових чисел, який при перерахунку листа буде кожен раз змінювати значення в комірках, що містять пов’язані з ним формули, необхідно сформовану вибірку сукупність скопіювати і вставити на лист „Випадковий_відбір”, вибравши в меню параметрів вставки параметр Значення (Values). Всі подальші розрахунки пов’язані з обробкою даних вибірки проводити на даному листі „Microsoft Excel”.

При механічному відборі процедура проведення вибірки аналогічна попередній. Відмінність полягає тільки в тому, що у вікні інструменту Вибірка зображеному на рис. 2, у методі відбору відмітити „Періодичний” та вказати період вибірки. Сформовану вибірку сукупність скопіювати і вставити на лист „Механічний_відбір”, вибравши в меню параметрів вставки параметр Значення (Values). Всі подальші розрахунки пов’язані з

обробкою даних вибірки проводити на даному листі „Microsoft Excel”.

Для проведення розширеного відбору генеральну сукупність слід структурувати (стратифікувати). Для цього такі дані генеральної сукупності, як „товарооборот”, „витрати обігу” та „кількість зайнятих” слід скопіювати і вставити на лист „Стратифікація_ген_сукупн”, вибравши в меню параметрів вставки параметр Значення (Values).

Утворення груп рекомендується здійснювати після сортування сукупності за зростанням товарообороту з допомогою інструменту Автофільтр (AutoFilter), який знаходиться в підменю Фільтр (Filter) меню Дані (Data). Для цього слід виділити будь-яку комірку в списку, вибрати інструмент Автофільтр (AutoFilter) в меню, а Excel введе кнопки зі стрілками (кнопки автофільтру) поряд з кожним заголовком стовпця. Клік мишкою на кнопці зі стрілкою поряд із заголовком стовпця розкриває список значень, які можна використовувати для встановлення умов відбору стрічок. Утворення груп магазинів за розміром товарообороту зручно здійснювати за допомогою Автофільтра користувача (Custom AutoFilter). Для цього розкривши список автофільтра стовпця „Товарооборот, тис. грн.” слід вибрати в ньому пункт Умова (Custom). З’явиться вікно діалогу, подібне до того, що показано на рис. 3.

Списки, що розкриваються, у лівій частині цього вікна діалогу містять оператори порівняння (рівно, не рівно, більше і т. д.), а списки в правій частині дозволяють вибрати чи ввести значення, які повинні використовуватися в умовах відбору. Зверніть увагу, що на рис. 3 встановлений перемикач І (And), а не Або (Or).

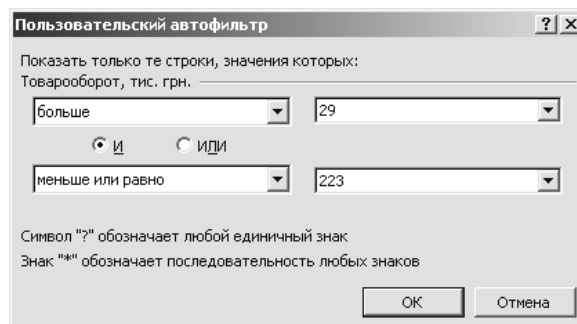


Рис. 3. Вікно діалогу Автофільтр користувача (Custom AutoFilter).

Після виділення кожної групи слід підбити групові та загальний підсумки. Наступним кроком є відключення інструменту Автофільтр (AutoFilter), аналогічно до процедури його запуску.

Для зручності роботи зі згрупованими даними на листі „Стратифікація_ген_сукупн” варто створити структуру. Щоб автоматично створити структуру для листа, виділіть одну комірку та в меню Дані (Data) виберіть команду Група і структура (Group And Outline) і потім - Створення структури (Auto Outline). Структурування робочого листа дозволяє легко виділяти тільки ті комірки, які знаходяться на одному ієрархічному рівні, що полегшує аналіз даних (рис. 4).

		A	B	C	D	E	F	G
		Групи підприємств за обсягом товарообороту, тис. грн.	Товарооборот, тис. грн.	Кількість підприємств	Витрати обігу, тис. грн.	Середні витрати обігу, тис. грн.	Кількість зайнятих, осіб	Середня кількість зайнятих, осіб
+	326	29 - 223	60750,19	324	7542,33	23,28	4643	14,33
+	1194	223 - 417	272702,64	867	18963,03	21,87	17269	19,92
+	2188	417 - 611	488832,18	993	12358,02	12,45	25091	25,27
+	2295	611 - 805	69974,97	106	1672,98	15,78	3326	31,38
-	2296	Разом	892260	2290	40536	17,70	50330	21,98

Рис. 4. Робочий лист „Стратифікація_ген_сукупн” після структурування.

Комірки групових підсумків містять, наприклад, наступні формули: B326=СУММ(B2:B325); C326=СЧЁТ(B2:B325); D326=СУММ(D2:D325); E326=D326/C326; F326=СУММ(F2:F325); G326=F326/C326, а комірки загального підсумку - B2296=СУММ(B2295;B2188;B1194;B326); C2296=СУММ(C2295;C2188;C1194;C326); D2296=СУММ(D2295;D2188;D1194;D326); E2296=D2296/C2296; F2296=СУММ(F2295;F2188;F1194;F326); G2296=F2296/C2296.

Кнопки Показати деталі (Show Detail) (кнопка зі знаком „+” на рис. 4) дозволяють розгорнути структуру, а кнопки Сховати деталі (Hide Detail) (кнопка зі знаком „-” на рис. 4) ховають нижчі ієрархічні рівні даних. Альтернативними кнопками, які виконують аналогічні функції є кнопки рівнів ієрархії, розташовані у лівому верхньому кутку структурованого робочого листа.

Формування розширеної вибірки здійснюється на основі проведених попередньо розрахунків, за формулою (9), обсягу частинних

вибірок n_j . Вибір елементів з кожної групи структурованої генеральної сукупності здійснюється за допомогою інструменту Вибірка (Sampling) простим випадковим способом. Сформовану вибірку сукупність слід скопіювати і вставити на лист „Розшарований_відбір”, а всі подальші розрахунки пов’язані з обробкою даних вибірки проводити на даному листі „Microsoft Excel”.

Даний розділ методичних вказівок містить тільки загальні відомості про роботу з надбудовою Аналіз даних (Data Analysis). Більш докладну інформацію можна знайти в спеціальній літературі [3].

4. ЗМІСТ ЗВІТУ З ЛАБОРАТОРНОЇ РОБОТИ

1. Мета роботи.
2. Основні теоретичні положення та розрахункові формули.
3. Вихідні дані згідно заданого варіанту завдання.
4. Привести у звіті всі проведені розрахунки та їх результати (роздрук рішення).
5. Висновки по роботі.

5. ЗАВДАННЯ ДЛЯ ВИКОНАННЯ РОБОТИ

За даними перепису підприємств роздрібної торгівлі у регіоні сформовано генеральну сукупність таких підприємств і представлено електронним файлом “Вибіркові обстеження” у табличному редакторі “Microsoft Excel”. З метою оцінки у майбутньому обсягів роздрібного товарообороту всіх підприємств за наведеними даними необхідно:

- 1. Простим випадковим способом** за даними файлу “Вибіркові обстеження” сформувавши вибірку, яка б з імовірністю 0,954 забезпечувала 10% відносну граничну похибку щодо обсягу роздрібного товарообороту. Результати відбору представити у табличному редакторі “Microsoft Excel” на окремому листі з назвою „Випадковий_відбір” у вигляді робочої таблиці. У таблиці зазначити порядковий номер відібраного підприємства, кількість зайнятих осіб, обсяг роздрібного товарообороту і відповідні витрати обігу. Нижче навести формули, за якими розраховувався необхідний обсяг вибірки.

- 2. Механічним способом** за даними файлу “Вибіркові обстеження” сформувати вибірку для обстеження середнього роздрібного товарообороту підприємств торгівлі, яка б з імовірністю 0,954 забезпечувала граничну похибку вибірки 20 тис. грн. Результати відбору представити у табличному редакторі “Microsoft Excel” на окремому листі з назвою „Механічний_відбір” у вигляді робочої таблиці. У таблиці зазначити порядковий номер відібраного підприємства, кількість зайнятих осіб, обсяг роздрібного товарообороту і відповідні витрати обігу. Нижче навести формули, за якими розраховувався необхідний обсяг вибірки.
- 3. Розширеним способом** за даними файлу “Вибіркові обстеження”, попередньо структурувавши генеральну сукупність, сформувати стратифіковану вибірку для обстеження середнього роздрібного товарообороту підприємств торгівлі, яка б з імовірністю 0,954 забезпечувала граничну похибку вибірки 20 тис. грн. Результати відбору представити у табличному редакторі “Microsoft Excel” на окремому листі з назвою „Розширений_відбір” у вигляді робочої таблиці. У таблиці зазначити порядковий номер відібраного підприємства, кількість зайнятих осіб, обсяг роздрібного товарообороту і відповідні витрати обігу. Нижче навести формули, за якими розраховувався необхідний обсяг вибірки.

У ході стратифікації (структурування) генеральної сукупності пропонується утворити від 3 до 5 груп. Критерієм якості проведеного групування є виконання наступної умови: $\sigma_j N_j \approx const$, де σ_j – середньо-квадратичне відхилення ознаки групування у j -тій групі; N_j - кількість підприємств торгівлі у j -тій групі;

За даними кожної вибірки розрахувати:

- загальний обсяг товарообороту за даними вибірки;
- середній обсяг товарообороту в розрахунку на одне підприємство;
- граничну похибку середньої та довірчий інтервал;
- поширити результати вибірових оцінок на генеральну сукупність з метою оцінки загального обсягу товарообороту всіх підприємств;
- порівняти отримані результати з відповідними показниками генеральної сукупності. Визначити абсолютну і відносну похибки вибірки та зробити висновки.

Обчислення провести з точністю до 0,00, результати розрахунків та висновки навести на робочому листі.

- 4. Порівняти результати** вибіркового оцінювання середнього значення роздрібного товарообороту за допомогою різних способів формування вибірки і зробити висновок про переваги і недоліки кожного із цих способів.

6. КОНТРОЛЬНІ ПИТАННЯ

1. Переваги та недоліки вибіркового обстеження великих статистичних сукупностей.
2. Як називається статистична сукупність з якої відбираються елементи для вибіркового обстеження?
3. Що таке похибки репрезентативності та як вони поділяються за причинами виникнення?
4. Поясніть, що таке стандартна і гранична похибки вибірки та взаємозв'язок між цими показниками.
5. Від яких параметрів залежить розмір граничної похибки вибірки?
6. Що таке довірчий інтервал вибіркової оцінки та як його визначають?
7. Які способи формування вибірових сукупностей використовують при несущільних обстеженнях?
8. Яким чином основа вибірки впливає на вибір способу формування вибірових сукупностей?
9. У яких випадках доцільно застосовувати механічний відбір елементів із генеральної сукупності?
10. Що спільного між розшарованим та серійним відборами і в чому полягає різниця між ними?
11. Чому похибки розшарованої та серійної вибірок є меншими ніж випадкової і механічної?
12. Які застосовують способи визначення обсягу частинних вибірок n_j при розшарованому відборі елементів генеральної сукупності у практиці вибірових обстежень?
13. Виходячи з чого визначають необхідний обсяг вибірки?
14. Яким чином модифікуються формули обсягу вибірки, якщо в основу розрахунку n покласти відносну граничну похибку вибірки?
15. Яким приймають обсяг вибірки, якщо за допомогою вибіркового обстеження одночасно вивчають декілька ознак?

7. СПИСОК ЛІТЕРАТУРИ

1. В.П.Боровиков, И.П.Боровиков. STATISTICA - Статистический анализ и обработка данных в среде Windows. Издание 2-е, стереотипное - М.: Информационно-издательский дом "Филинь", 1998. - 608 с.
2. Статистика. Підруч. для вузів за ред Герасименка С.С. – К.: Вища шк., 2000. –468 с.
3. Додж М., Стинсон К. Эффективная работа с Microsoft Excel 2000 - СПб: Издательство "Питер", 2000. – 1056 с.
4. Єріна А.М., Пальян З.О. Теорія статистики: Практикум. - К.: - Товариство "Знання", КОО, 1997. - 325 с.
5. Захожай В.Б., Попов І.І., Коваленко О.В. Практикум з основ статистики: Навч. посіб. – К.: МАУП, 2001. – 176 с.
6. Статистика підприємництва: Навч. посібник / П.Г.Вашків, П.І.Пастер, В.П.Сторожук, Є.І.Ткач; за ред. П.Г.Вашківа, В.П.Сторожука. – К.: "Слобожанщина", 1999. – 600 с.
7. Фещур Р.В., Барвінський А.Ф., Кічор В.П.. Статистика: Навчальний посібник. – Львів: "Інтелект-Захід", 2001. – 276 с.

НАВЧАЛЬНЕ ВИДАННЯ

ВИБІРКОВІ СПОСТЕРЕЖЕННЯ

МЕТОДИЧНІ ВКАЗІВКИ

до лабораторної роботи
з курсу "Статистика"

для студентів базового напрямку 6.0502 "Економіка і підприємництво"
спеціальності 7.050105 "Економічна статистика"

Укладачі

Гринькевич Ольга Степанівна
Сорочак Олег Зіновійович

