|  |  |
| --- | --- |
| **Лекція 3 Cтатистичне зведення і групування** |  |

1. [Поняття зведення, його задачі і зміст](https://dl.sumdu.edu.ua/textbooks/28071/49805/index.html%22%20%5Cl%20%22p1)
2. [Організація і техніка здійснення зведення](https://dl.sumdu.edu.ua/textbooks/28071/49805/index.html%22%20%5Cl%20%22p2)
3. [Групування, їх сутність, види і принципи формування](https://dl.sumdu.edu.ua/textbooks/28071/49805/index.html%22%20%5Cl%20%22p3)
4. [Вторинне групування](https://dl.sumdu.edu.ua/textbooks/28071/49805/index.html%22%20%5Cl%20%22p4)
5. [Статистичні таблиці, їх види і основні правила побудови](https://dl.sumdu.edu.ua/textbooks/28071/49805/index.html%22%20%5Cl%20%22p5)

**Поняття зведення, його задачі і зміст**

Першим етапом статистичного вивчення тих або інших масових явищ є **статистичний спостереження** (СС), тобто процес збирання різних відомостей про них за певним планом. В результаті СС одержують масу первинних (початкових) даних про окремі одиниці об'єкту, що вивчається, які поки що практично нічого не говорять досліднику, окрім того, що їх багато. Тобто первинні дані, без їх попередньої обробки, не дають можливості зробити які-небудь висновки про об'єкт, що вивчається, або виявити закономірності соціально-економічних явищ. Тому наступний етап статистичного дослідження полягає в систематизації первинних даних і отриманні на цій основі узагальнюючих показників, що характеризують об'єкт, що вивчається, в цілому. Процес систематизації і узагальнення матеріалів статистичного спостереження називається **статистичним зведенням**.

В процесі зведення здійснюється якісний перехід від даних, що характеризують окремі одиниці спостереження, до даних, що характеризують всю сукупність, що вивчається.

***Наприклад*,** за даними про успішність окремих студентів не можна зробити висновок про успішність у вузі. І лише узагальнивши і систематизувавши результати спостережень, можна зробити висновки про успішність в цілому і її тенденціях і закономірностях.

Головним завданням зведення є отримання узагальнюючих статистичних показників, що розкривають сутність соціально-економічних явищ і певні закономірності.

По складності обробки первинних матеріалів зведення є просте і складне.

1. **Просте** - процес підрахунку загальних підсумків по сукупності, що вивчається, без систематизації матеріалів СС.
2. **Складне** - комплекс операцій, що включають групування первинних матеріалів:
	* підрахунок підсумків по кожній групі і сукупності в цілому;
	* представлення результатів групування і зведення у вигляді статистичних таблиць.

Статистичне зведення проводиться за певною програмою, яка розробляється одночасно з планом і програмою СН.

Програма зведення включає наступні етапи:

1. Визначення групувальних ознак і порядку формування груп і підгруп даних.
2. Підрахунок підсумків по кожній групі і сукупності в цілому.
3. Розробка макетів статистичних таблиць, в яких будуть представлені результати зведення.

Всі ці питання повинні розв'язуватися з урахуванням поставленої мети дослідження і особливостей сукупності, що вивчається.

За формою обробки матеріалу і організації зведення буває:

* централізоване;
* децентралізоване.

При централізованому зведенні весь матеріал СС потрапляє до єдиного центрального органу статистики - Держкомстату і там обробляється і систематизується. При децентралізованому зведенні матеріали СС обробляються і систематизуються місцевими органами статистики поетапно, від низу до верху, а до центрального органу статистики представляються результати зведення.

*Наприклад*, дані про виробничо-господарську діяльність підприємств узагальнюються спочатку в межах адміністративних районів і передаються до обласних органів статистики. Там вони узагальнюються вже в межах області і потім результати зведення представляють в Держкомстат.

Вся звітність, як правило, проходить децентралізоване зведення. Централізоване ж зведення більш застосовне для розробки великих спеціальних обстежень.

Кожне з цих зведень має свої позитивні сторони. Так, децентралізована дозволяє більш оперативно одержувати зведені відомості для керівних органів окремих адміністративних районів і дає можливість швидше уточнювати ті або інші сумнівні відомості. Централізоване зведення, коли в одному центрі зосереджується маса матеріалу, дозволяє більш ефективно використовувати механізацію, що веде до здешевлення обробки і скорочення термінів зведення. Правда, в даний час широке впровадження швидкодійних персональних ЕОМ дозволяє і на місцях механізувати процес обробки первинних даних і оперативно одержувати необхідні відомості.

**Організація і техніка здійснення зведення**

Перш ніж приступити до зведення, весь первинний матеріал (як вже згадувалося в попередній лекції) піддається перевірці, контролю. Він проводиться для того, щоб в зведення потрапила достовірна, повна, якісна інформація.

**Логічний контроль**перевіряє смислову узгодженість відомостей, зафіксованих у первинному документі.

**Арифметичний контроль** - правильність групових і загальних підсумків.

По техніці або способу виконання зведення може бути ручним і механізованим. Безумовно, домінуюче положення в даний час займає механізоване зведення. Але для невеликих масивів даних може застосовуватися і ручне. В усякому разі, техніку його виконання треба знати і вміти користуватися.

Отже, зведення завжди починається з шифровки матеріалу.  **Шифровка**полягає в тому, що за допомогою умовних знаків (чисел або букв) відмічається картка (формуляр) з відповідями або окремі відповіді (програми спостереження), що свідчить про те, що дана одиниця спостереження віднесена нами до якоїсь з визначених програмою зведення груп або підгруп даних, по якійсь конкретній ознаці.

Так, при шифровці карток перепису населення відповідь на питання про освіту може мати такі варіанти: вища, незакінчена вища, середня, середня спеціальна, неповна середня. Тоді кожний варіант відповіді одержить певний шифр: вища - 1, незакінчена вища - 2, середня - 3 і т.д.

Ця операція проводиться для зручності підрахунків і для того, щоб у всіх спірних випадках одержати правильний одноманітний розв'язок. Після шифровки зведення первинних даних, тобто підрахунок підсумків, вже не представляє методологічних труднощів і здійснюється або ручним, або механізованим способом. Ручне зведення по техніці виконання розрізняється між собою залежно від типу первинного документа (формуляру) - картка, список, бланк і т.д.

Так, якщо первинні дані занесені в списки, або є списками, то зведення проводиться за допомогою запису кожного окремого випадку у вигляді риски на спеціальній робочій таблиці. Така таблиця складається заздалегідь і в ній повинне бути передбачене розбиття даних на певні групи.

**Групування, їх сутність, види і принципи формування**

В процесі обробки і узагальнення первинної інформації просте зведення використовується рідко. Часто виникає необхідність виділення однорідних груп даних за певною ознакою з метою отримання необхідних аналітичних показників. Тобто частіше використовується складне зведення, що включає групування початкової інформації.

**Групування**- це процес розподілу сукупності на однорідні групи за якою-небудь ознакою.

Групування - один з найважливіших етапів статистичного дослідження, центральний момент будь-якого зведення, і один з основних прийомів статистики. Вона дає можливість:

* виявити перехід кількісних змін в якісні.
* виявити закономірності зміни і розвитку соціально - економічних явищ, а також можливість застосування інших статистичних методів.

Ознаки, за якими проводиться розподіл одиниць сукупності на групи, називаються групувальними, або основою групування. Вибір групувальних ознак визначається метою і задачами дослідження.

Як приклад: здійснення ручного групування робітників підприємства за освітою.

Нехай є список робітників з відміткою про рівень їх освіти. Тоді матеріал цього списку можна представити таким чином за допомогою його групування за ознакою освіти:



Для зручності підрахунку окремі випадки у вигляді «рисок» ставлять таким чином: 4 риски - поряд одна з одною, а п'ятий випадок показується горизонтальною рискою, що перекреслює 4 попередні. Якщо ж первинний матеріал розташований на окремих картках, то зведення зручно здійснити, розкладаючи їх на певні групи. Останній етап ручного зведення - підрахунок по сформованих у результаті розкладки групах відповідних підсумків і розрахунок необхідних якісних показників.

В даний час ручне зведення використовується в практичній діяльності рідко. Зараз практично скрізь в зв'язку з упровадженням механізації, автоматизації і ЕОМ застосовується механізоване і автоматизоване зведення первинних даних.

Успіхи в розробці програмного математичного забезпечення ЕОМ, широкий набір стандартних програм контролю, сортування, класифікації і зведення різних даних забезпечили оперативність і дієвість статистики, знизили собівартість обчислювальних робіт.

Зараз у всіх органах статистики (особливо обласного рівня) створені могутні обчислювальні центри, що дозволяють обробляти величезні масиви різної інформації і оперативно видавати необхідні узагальнюючі відомості для аналізу і прийняття управлінсько-господарських рішень.

Автоматизація процесу обробки даних поставила задачу автоматизації шифровки (кодування) первинної інформації.

Були створені різні класифікатори (шифратори) підприємств, галузей, видів продукції, робіт, послуг і т.д.

Автоматизована обробка даних спричинила за собою і автоматизований збір первинної інформації. Нині практично на кожному підприємстві є автоматизована система управління виробництвом (АСУВ), яка включає систему збору, обробки і передачі інформації.

В практиці статистичної роботи існують також загальноприйняті статистичні стандарти розподілу одиниць сукупності на групи. Це так звані класифікації, в яких чітко визначені вимоги й умови формування груп. Наприклад класифікація галузей промисловості, продукції, форм власності, громадян по роду заняття і т.д. необхідність розробки класифікацій обумовлена різноманіттям атрибутивних ознак при вивченні масових соціально - економічних явищ.

При побудові групувань і формуванні однорідних даних, важливим моментом є визначення груп і меж кожної з них, тобто величини групового інтервалу. Кількість груп залежить від варіації групувальної ознаки і обсягу сукупності. Якщо групувальна ознака атрибутивна (її неможливо виразити числом ), то кількість груп визначається числом різновидів ознаки. Наприклад, при групуванні працівників цеху за освітою, ми виділи п'ять груп тому, що саме стільки є різновидів рівня.

Якщо ж групування проводиться за кількісною ознакою (її можна виразити числом), то кількість груп, на які буде розділена сукупність, визначається:

* або експертним шляхом, виходячи з обсягу сукупності (при цьому число груп приймається за звичай не менше 4, і не більше 11 - 12)
* або за формулою, запропонованою американським статистиком Спенсером

**N=1+3,322 lg N**

де n- число груп; N - число одиниць сукупності.

Величина інтервалу в кожній групі визначається таким чином:

* якщо значення ознаки змінюються рівномірно і в невеликих межах, формуються рівні інтервали за формулою:



Де h - величина групового інтервалу;

- максимальне значення досліджуваної ознаки;

- її мінімальне значення в сукупності;

h - число груп, на які буде розбита сукупність.

* якщо ж діапазон значень ознаки, що вивчається дуже великий і змінюється нерівномірно, застосовуються нерівні інтервали. Наприклад, при групуванні міст по числу жителів може бути таке групування: до 9 тис. чол.; 10-19; 20-49; 50-99; 100-499; 500 і більше. Інтервали в кожній з виділених груп в даному випадку нерівні.

Величина інтервалу - це різниця між його верхньою і нижньою межею. Нижньою межею інтервалу вважається число, з якого починається інтервал, а верхньою - яким закінчується інтервал. Розрізняють відкриті і закриті інтервали. Відкритим вважають інтервал, що має одну чітко позначену межу, верхню або нижню. Закритим - у якого позначено обидві межі.

В наведеному групуванні міст за чисельністю жителів інтервали в першій і останньої групах - відкриті, інші - закриті.

В процесі групування часто буває так, що верхня межа одного з інтервалів є нижньою межею наступного і тоді виникає питання в який з інтервалів включати цей граничний показник. В подібних випадках, як правило, чинять таким чином: нижню межу інтервалу вважають включною, а верхню - виключною. Тобто верхня межа в інтервал не включається (окрім, зрозуміло, останньої групи), а нижня завжди включається.

Розглянемо приклад побудови групування з рівними інтервалами, оскільки воно найбільш часто застосовується на практиці. Нехай є дані про кількість часу, яку витрачає кожен робітник бригади на виготовлення деталі:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 12,5 | 14,8 | 11,0 | 16,2 | 10,0 |
| 18,4 | 11,6 | 12,0 | 15,7 | 13,2 |
| 15,0 | 17,4 | 13,5 | 16,8 | 17,6 |
| 14,3 | 15,7 | 16,0 | 14,5 | 12,5 |
| 15,5 | 14,0 | 17,5 | 20,0 | 14,3 |

З метою визначення структури робітників за часом, який витрачається на виготовлення деталі, необхідно провести їх групування за витратами часу.

Визначимо величину групового інтервалу - число груп, на які необхідно розбити сукупність робітників за витратами часу. Максимальне значення ознаки, що вивчається, складає 20 хв., максимальне -10. Різниця між ними рівна 10. Підбираємо тепер число груп таким чином, що вийшла їх кількість в межах від 4-х до 11-12. таке число дорівнює 5, тобто



Таким чином, досліджувану сукупність робітників доцільно розбити на 5 груп з величиною інтервалу, рівною 2. Проведемо групування, пам'ятаючи, що нижня межа - включається, верхня - виключається.

Таблиця 1 - Групування робітників за витратами часу

|  |  |  |
| --- | --- | --- |
| **Групування робітників за витратами часу** | **Кількість робітників** | **Кількість робітників в % до підсумку** |
| 10-12 | 3 | 12 |
| 12-14 | 5 | 20 |
| 14-16 | 9 | 36 |
| 16-18 | 6 | 24 |
| 18-20 | 2 | 8 |
| **Разом** | **25** | **100** |

Результати групування дають чітку картину розподілу робітників бригади за часом, який витрачається на виготовлення деталі.

В статистиці групування застосовуються для вирішення різних задач, основними з них є:

1. вивчення структури сукупності за якою-небудь ознакою і структурних змін, що відбуваються в ній. Групування, за допомогою яких розв'язуються такі задачі, називаються структурними. Прикладом структурного групування є наведене вище групування розподілу робітників за витратами часу (табл. 1)
2. виділення соціально - економічних типів явищ з різнорідної сукупності. Групування, що розв'язують задачі такого роду називаються типологічними.

 Наприклад, групування промислових підприємств за формами власності (таблиця 2).

Таблиця 2 - Групування підприємств області за формами власності

|  |  |
| --- | --- |
| **Групи підприємств за формами власності** | **Кількість підприємств** |
| **Всього одиниць** | **В % до підсумку** |
| Державна власність | 260 | 95 |
| Приватна власність | 8 | 3 |
| Змішана власність | 6 | 2 |
| Разом | 274 | 100 |

І структурні і типологічні групування є описовими. Вони характеризують структуру сукупності і відрізняються тільки рівнем якісних відмінностей між групами.

1. виявлення зв'язку між двома ознаками, одна з яких розглядається як результат, інша - як чинник, що впливає на результат. Групування проводиться за факторною ознакою і в кожній групі не розраховується середнє значення результативної ознаки. Якщо із збільшенням факторної ознаки, при переході від групи до групи, відбувається відповідне збільшення або зменшення середнього значення результативної ознаки, значить між ними є зв'язок. Такого роду групування називаються аналітичними. Як приклад аналітичного групування розглянемо чи існує зв'язок між стажем і годинною заробітною платнею десяти робочих бригад.

Таблиця 3

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **№ робочого** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| **Стаж, років** | 2,0 | 3,5 | 4,0 | 5,2 | 5,8 | 6,1 | 7,0 | 7,5 | 7,8 | 10,0 |
| **Зарплата, т.грн.** | 1,0 | 1,0 | 1,2 | 1,1 | 1,3 | 1,2 | 1,4 | 1,3 | 1,4 | 1,4 |

Визначимо розмір групового інтервалу і число груп



Проведемо групування:

Таблиця 4

|  |  |  |
| --- | --- | --- |
| **Групи робітників за стажем** | **К-ть робочих** | **Середня зарплата, т.грн.** |
| 2-4 | 2 | 1,0 |
| 4-6 | 3 | 1,2 |
| 6-8 | 4 | 1,3 |
| 8-10 | 1 | 1,4 |
| **Разом** | **10** | **1,2** |

Результати групування свідчать, що між виробничим стажем робітників і їх зарплатою є прямий зв'язок.

Групування, проведені за одною якою-небудь ознакою, називаються простими (монотетичними), або рядами розподілу. Групування за двома і більш ознаками - омбініційними (політетичними). Комбіновані групування застосовуються у випадках, коли однієї груповальної ознаки недостатньо для аналізу явища, що вивчається.

Наприклад, при вивченні складу населення регіону цікаве групування по декількох ознаках, зокрема, за статтю й типу поселень. В цьому випадку спочатку проводять групування населення за статтю, а потім усередині групування ще одне - за типом населених пунктів (табл 5).

Таблиця 5. Склад населення регіону

|  |  |  |
| --- | --- | --- |
| **Розподіл населення за статтю** | **Розподіл населення за типом населених пунктів** | **Разом** |
| **міське** | **сільське** |  |
| Чоловіча | 460 000 | 240 000 | 700 000 |
| Жіноча | 570 000 | 350 000 | 920 000 |
| Всього: | 1030 000 | 590 000 | 1620 000 |

**Вторинне групування**

Особливим видом групувань є так зване **вторинне групування**- операція утворення нових груп на основі раніше проведеного групування.

До такого групування вдаються тоді:

* коли з великого числа спочатку утворених груп потрібно отримати менше число більш крупних;
* коли з метою порівняння потрібно привести в зіставний вигляд по-різному згрупований матеріал.

Найпростішим способом вторинного групування є такий, за якого нові групи утворюються шляхом укрупнення інтервалів. Тобто шляхом об'єднання в одну групу декількох дрібних груп, одержаних при первинному угрупуванні. Або за допомогою перегрупування за питомою вагою окремих груп загалом в їх підсумку.

**Статистичні таблиці, їх види і основні правила побудови**

Існують три способи представлення статистичних даних:

* в тексті,
* в таблицях,
* графічно.

Частіше всього результати обробки статистичних даних оформлюються у вигляді статистичних таблиць. Таблична форма дозволяє висловити матеріал найбільш зручно, компактно, наочно і раціонально. Основна перевага табличної форми викладу статистичних даних полягає в тому, що за допомогою таблиць легше всього здійснювати порівняння, зіставлення і аналіз цифрової інформації.

Слід мати на увазі, що не кожна таблиця, що має цифрову інформацію, є статистичною! Статистичними вважаються тільки ті таблиці, в яких містяться результати (підсумки) зведення первинної статистичної інформації.

Вона являє собою ряд горизонтальних рядків і вертикальних граф, на перетині яких утворюються клітини таблиці. Таблиця містить назву (загальний заголовок), а також верхні заголовки (назви граф) і бічні - назви рядків (таблиця 4).

Статистичну таблицю можна представити як своєрідну логічне речення - в ній розрізняють підмет і присудок.  **Підмет**- те, про що йде мова в таблиці, що характеризується (підприємство, населення, врожайність).

**Присудок**- те, що мовиться про підмет, тобто ті показники, які характеризують підмет (випуск продукції, число працюючих, валовий збір зерна і т.п.).

Таблиця 6 - Приклад макета статистичної таблиці
Таблиця № - Загальний заголовок (назва таблиці)



За побудовою підмета таблиці підрозділяються на: прості, групові і комбінаційні.

**Прості -**це ті, в яких підмет є переліком окремих одиниць сукупності (заводів, районів, об'єктів), а також динамічні таблиці, де підметом є окремі роки, місяцями або іншими періодами часу**Групові**- ті, підмет яких містить групування одиниць спостереження за однією ознакою.

**Комбінаційні**- підмет яких містить групування одиниць спостереження за двома і більш ознаками.

Групові і комбінаційні таблиці називаються одним словом - **складні**. В роботі економіста таблиці грають важливу роль. Тому він повинен добре розбиратися в будь-якій складній таблиці, вміти її «читати» і знаходити матеріали для висновків. Крім того, він сам повинен уміти складати таблиці так, щоб вони були зрозумілі всім, кому призначені. Тобто треба уміти грамотно побудувати таблицю.

Складаючи таблиці слід мати на увазі, що таблиця повинна компактно і наочно висловлювати цифровий матеріал. Якщо ж зайва деталізація підмета і присудка ускладнює читання таблиць, то краще замість однієї громіздкої таблиці побудувати декілька більш простих.

На кінець, відзначимо декілька технічних моментів, пов'язаних з побудовою таблиць.

При побудові таблиць необхідно:

1. Щоб кожна таблиця мала докладну назву, що показує яке коло питань виражається або ілюструється таблицею. Тобто чіткість заголовків, їх стислість і виразність.
2. В назві таблиці вказати період часу, за який приведені дані, або момент часу, до якого вони відносяться.
3. В назві вказати одиниці вимірювання, якщо вони однакові для всіх табличних клітин. Якщо неоднакові, то вони вказуються в заголовках граф або рядків. Якщо кожний рядок має свою особливу одиницю вимірювання, то для їх позначення слід відводити спеціальну графу.
4. Бажано давати нумерацію граф. Це полегшує користування таблицею і дозволяє показати спосіб розрахунку цифр в графах. Наприклад, в заголовку 3-ї графи може бути вказано (гр.1 + гр.2).
5. Перша графа призначається для бічних заголовків і тому в нумерацію не входить. Вона позначається буквою «А».
6. Статистична таблиця, як правило, повинна мати підсумки по графах і рядках - «разом» або «всього».
7. Округлення чисел у всіх графах і рядках слід проводити з однаковою точністю.
8. В кожній клітині таблиці повинне стояти яке-небудь число. Якщо число чомусь відсутнє, то це повинне бути пояснено наступним умовним чином:
	* якщо клітина взагалі не підлягає заповненню, то в ній ставлять знак «х»;
	* якщо відомості відсутні, (.), чи «нема від.»;
	* якщо відсутнє саме явище (сам факт), то ставлять тире (-);
	* якщо числове значення менше прийнятої в таблиці точності, то ставлять «0,0» або «0,00».
9. У разі потреби в таблиці робляться зноски та примітки, які треба взяти до уваги при читанні таблиці. Так, якщо поряд з якимсь числом стоїть знак (\*), це значить, що дані є попередніми, або вказати джерело інформації звідки вони взяті.

Правильно і грамотно складена таблиця є важливим засобом викладу статистичного матеріалу і засобом аналізу. Аналіз статистичної таблиці звичайно починають із загального підсумку, що дає характеристику всієї сукупності, потім переходять до вивчення окремих рядків і граф, тобто до оцінки складових частин досліджуваної сукупності.

Таким чином, статистичне зведення - другий етап статистичного дослідження, що є складним процесом наукової обробки первинних статистичних даних. Зведення включає групування даних, розробку системи показників для характеристики типових груп і підгруп, підрахунок по групах і сукупності в цілому, а також представлення результатів у вигляді статистичних таблиць. При вивченні цієї важливої теми необхідно звернути особливу увагу на правила побудови групувань, зокрема на питання визначення розміру групового інтервалу і кількості груп, а також меж інтервалів.

**Питання самоконтролю**

1. Що є другим етапом дослідження?
2. Які види зведення ви знаєте? Дайте їх коротку характеристику.
3. Що таке групування і які задачі вирішуються за допомогою групування?
4. Що розуміють під величиною групового інтервалу і як він визначається?
5. Як визначити число груп, на які буде розбита сукупність, що вивчається?
6. Які інтервали називають відкритими і як вони відрізняються від закритих?
7. Що таке статистична таблиця? Переваги статистичної таблиці і правила їх складання.
8. Розкажіть зміст теми, використовуючи як план запропоновані вище питання.