

Лекція 1.

ОСНОВНІ ПОНЯТТЯ ТА КАТЕГОРІЇ ПРИКЛАДНОЇ СТАТИСТИКИ

План

1. Що таке статистика.
2. Як виникла і в яких напрямках розвивалася наука статистика, її предмет, завдання, функції.
3. Найважливіші категорії статистики.
4. В якій послідовності відбувається статистичне дослідження.
5. Метод статистичного спостереження.

1. Слово “статистика” походить від латинського ”status” – стан. Звідси утворилися слова: “stato” – держава; “statista” – знавець держави; “statistica” – сума знань про державу. Спочатку статистика надавала лише словесний опис державних справ і лише з XIX ст. статистичні відомості стали повідомляти в кількісній формі. В сучасному розумінні термін “статистика” вживається в науці з 1749 року. Вперше його запровадив німецький вчений професор філософії та права Готфрид Ахенваль (1719-1772).

Термін “статистика” багатозначний, у науковій і практичній сферах під ним розуміють:

- вид інформації, тобто зібрані статистичні дані (підсумкові цифрові показники) про різноманітні масові явища суспільного життя;
- форму практичної діяльності, спрямовану на збирання, упорядкування, оброблення та інтерпретацію даних (системи статистичних установ, дослідницькі центри тощо);
- суспільну науку – спеціальну наукову дисципліну, яка займається розробкою теоретичних положень і методів, що використовуються статистичною практикою, а також навчальну дисципліну, яку викладають у вищих закладах освіти всіх рівнів для підготовки фахівців найвищої кваліфікації;
- “міру”, “оцінку” або “характеристику”, тобто певні статистичні показники, наприклад, узагальнені показники вибіркової сукупності (середня, частка) тощо.

2. Статистика починалася з масових спостережень: перепису населення (Давній Китай, 2000 р. до н. е.; Давня Греція, 850 р. до н. е.; Рим, 435 р. до н. е.), опису земельних ділянок, робочої сили, ремесел, торгівлі, харчів, (Давній Єгипет, 1500 р. до н. е.). Згадування про статистичні обстеження зустрічаються й у Біблії, у Четвертій книзі Мойсея “Числа” (розповідається про перелік чоловічого населення, здатного носити зброю). Античний світ значно змінив характер господарського обліку. Швидкий розвиток приватної власності призвів до того, що окрім державного обліку, з’явився облік з ініціативи банкірів, торговців, власників майстерень. Середньовіччя в розвитку обліку було кроком назад. Роз’єднаність держав, зростання церковного мракобісся, низький рівень освіченості населення, фантастичні уяви про найпростіше, звичайне та очевидне привели до різкого падіння інтересу досліджувати істину. Однак і з цих часів дійшла до нас унікальна пам’ятка – “Книга страшного суду” – зведення матеріалів поземельного перепису населення Англії, його майна і багато іншого. Розвиток торговельних та міжнародних товарно-грошових відносин, бухгалтерського обліку в епоху Відродження стимулювало розвиток статистики і теорії ймовірностей. В XIV ст. виникли перші страхові товариства в Італії, які підраховували шанси, оцінювали ризик перевезень.

У другій половині XVII століття, в працях таких видатних вчених як Г. Галілей, Р. Декарт, І. Кеплер, Б. Спіноза, Г. Лейбніц, І. Ньютон, були закладені основи сучасної науки. Розвиток таких фундаментальних наук як математика, фізика, астрономія, філософія, біологія допомогли усвідомити значення статистики як засобу соціального пізнання, що призвело до виникнення статистики як науки. Це трапилось майже одночасно наприкінці XVII століття на континенті Європи – в Німеччині та Великобританії.

Наука ця, що зародилася як результат узагальнення вже достатньо розвинутої статистичної практики, в екології і біології почала розвиватися у таких напрямках:

- описова статистика в екології (біології);
- статистична (біологічна) арифметика;
- теорія екологічних (біологічних) явищ і процесів.

Описова статистика веде свій початок від Аристотеля (384-322 рр. до н. е.), однак справжня наукова школа виникла в Німеччині. Її засновником був німецький вчений, лікар і державний діяч Герман Конрінг (1606-1681). Перший спеціальний курс лекцій з статистичної дисципліни (*diciplina statistica*) був прочитаний ним у 1660 р. Його послідовник Г. Ахенваль з 1746 р. спочатку у Марбурзькому, а потім і у Геттингенському університеті почав читати нову учбову дисципліну, яку він назвав статистикою. Це була школа описової статистики. Її представники вважали основним завданням статистики систематизоване описування фактів, що визначають майбутнє держави, а також навчити політичних діячів розуміти причини важливих державних процесів. Досліджували не суспільство, а державу. Переважали словесні характеристики – територія, державний устрій, населення, релігія, зовнішня політика тощо. Школа проіснувала більш ніж 150 років, не змінюючи своїх теоретичних основ. Її заслуга полягає у введенні статистики в число університетських дисциплін. На зміну розумінню статистики лише як опису держави прийшло розуміння її як науки про закономірності масових суспільних явищ.

Статистична арифметика – таку назву отримала статистична школа, що виникла в Англії. Від дослідження відомостей про померлих до розроблення таблиць смертності, від ідеї закону великих чисел до застосування методів усунення випадкових відхилень, від оцінювання явищ в умовах нестачі числових даних до заснування економічної статистики – такий шлях пройшли знамениті дослідники Джон Граунт (1620-1674), Едмунд Галлей (1656-1742) і Вільям Петті (1623-1687) – засновники цієї школи. Вони заклали підвалини політичної арифметики як науки про політичну анатомію, тобто науки про симетрію, структуру і співвідношення частин “політичного тіла”. Політичні арифметики використовували мову чисел, ваги, мір, віддавали перевагу кількісним характеристикам, виявленню закономірностей явищ на основі числових даних. Розвиток математики, зокрема теорії ймовірності, створив теоретичну базу для статистики. Особливий вклад у розвиток цього напрямку вніс бельгійський статистик Адольф Кетле (1796-1874). Теоретична статистика починається з його праць “Розрахунок вірогідних злочинів” (1829), “Листи про теорію ймовірностей” (1830), “Соціальна фізика, або досвід досліджень про розвиток людських здібностей” (1836), у яких встановлено факт існування соціальних законів (закономірностей в моральній поведінці людини), які настільки ж об’єктивні й точні, як і закони природи. І чим більше явищ спостерігалось, тим виразніше вони проявлялись. А. Кетле називав статистику “соціальною фізикою”, яка займається “вивченням життя соціального тіла”, тобто наукою, що вивчає закони суспільної системи за допомогою кількісних методів. В якості фундаментальної основи статистики А. Кетле висував теорію ймовірностей. Він писав: “...цей зв’язок зовсім не випадковий: одна з цих наук своїми обчисленнями досліджує і координує те, що інша здобуває своїми спостереженнями”.

На Русі першими і основними обліково-статистичними джерелами були літописи (IX–XI ст.). Як метод пізнання явищ, вітчизняна статистика почала формуватися у XVIII – XIX ст.

Найбільш прогресивні для цього часу теоретичні основи статистики як самостійної науки були створені Д.П.Журавським (1810-1856). Вершиною його статистичного дослідження стала робота “Статистичний опис Київської губернії” в якій велику увагу він приділив питанню достовірності даних, методу групування, розкрив принцип єдності кількісного і якісного аналізу. Не можна не згадати і такий напрям статистичної практики, як земська статистика. Головним завданням земств було вивчення загальних умов життя народу, бюджету селянських сімей за дуже розгорнутими програмами спостереження, з використанням методів групування. Представниками Чернігівського земства були П. Червинський (1849-1931) та О.Шлікевич (1849-1909), визнаний всіма винахідник особливого типу таблиць – комбінаційних. На початку ХХ ст. відбулося формування основних наукових напрямків вітчизняної статистики, проникнення ідей англо-американської статистики в середовище вітчизняних вчених, зародження й злет стохастичної статистики (від грецького “стохастика” – припускати), а також порівнювальної статистики. Статистичний метод стає універсальним. Представниками цих наукових напрямків були видатні вчені Ю. Янсон, О. І.Чупров, О.О.Чупров, М.Каблуков, А.Кауфман. Становлення статистики як науки в сучасному її розумінні пов’язане з іменами Ф.Гальтона (1822-1911) та К.Пірсона (1857-1936) – засновниками біометрії, кореляційно-регресійного аналізу; П.Чебишева (1821-1894), А.Маркова (1856-1922), О.Ляпунова (1857-1918), О.О.Чупрова (1890-1968) – математичні основи вибіркового методу; Р. Фішера (1890-1968) – теорія статистичного оцінювання та перевірки гіпотез і багатьох-багатьох інших. Отже історія статистики доводить, що статистична наука сформувалася внаслідок теоретичних узагальнень накопиченого людством досвіду обліково-статистичних робіт (практична статистика існує майже 4000 років, в той час як теоретична – лише біля 300), зумовлених потребами управління виробництвом і життям суспільства. Сучасна статистична наука є складною, багатогалузевою системою наукових дисциплін.

Основні складові частини статистики:

- теорія статистики;
- економічна статистика;
- галузеві статистики;
- екологічна (біологічна) статистика;
- соціальна статистика.

Предметом статистики є вивчення розмірів і кількісних співвідношень між якісно визначеними масовими явищами, а також закономірностей їх формування, розвитку, взаємозв’язку у конкретних умовах місця та часу.

Слід звернути увагу на такі принципові особливості статистичної характеристики досліджуваних явищ:

кількісну розмірність явищ неможливо розглядати без їх якісної визначеності (наприклад: кількість книжок виданих у 1997 р. – 6308, з них українською мовою – 3140, російською – 2364 тощо);

статистика досліджує не поодинокі, а масові явища.

Відповідно статистичні показники не відносяться до кожного окремого випадку, а завжди являють собою результат узагальнення даних за масою випадків. Завдання статистичного дослідження полягає в узагальненні даних і у виявленні закономірностей явищ в конкретних умовах місця та часу, які проявляють себе лише у великій масі явищ через подолання властивої одиничним елементам випадковості.

Головна функція статистики – постачання інформації, її математична обробка та оцінка достовірності.

3. Життя настільки складне і різноманітне, що майже завжди можна знайти факти, приклади, які б або підтвердили, або спростували одне і теж твердження. Тому, щоб охарактеризувати масове суспільне явище в цілому, необхідно розглянути всю чи дуже велику масу окремих явищ та процесів, що його стосуються. Таку масу називають статистичною сукупністю, об'єктом статистичного дослідження.

Для вивчення цього об'єкта в статистиці застосовують ряд важливих понять і категорій:

- статистична сукупність;
- одиниця статистичної сукупності;
- одиниця статистичного спостереження;
- ознака;
- варіація;
- статистична закономірність,
- статистичний показник.

Статистична сукупність – це впорядкована множина існуючих елементів (однорідних у певному відношенні), поєднаних спільними, притаманними кожному елементу множини, властивостями, умовами та причинами існування і розвитку. Статистична сукупність зовсім не обов'язково вміщує велику кількість одиниць, вона може бути і досить маленькою, наприклад, обсяг сукупності малої вибірки може складати іноді 5-10 одиниць. Важливо пам'ятати, що статистична сукупність складається з реально існуючих матеріальних об'єктів, які підлягають дослідженню. Наприклад: в якості такої множини об'єктів можуть виступати навчальні заклади, підприємства, сім'ї, студенти, громадяни будьякої країни тощо.

Одиниця статистичної сукупності – це окремий елемент множини, який має всі властивості, що притаманні досліджуваній сукупності. Одиниця сукупності є носієм ознак, значення яких підлягають реєстрації. Наприклад: студенти інституту – це сукупність, кожен студент – її елемент, який має спільні з іншими елементами ознаки: стать, вік, спеціальність, екзаменаційна оцінка тощо. Одиниця статистичного спостереження – це первинна одиниця, від якої одержують інформацію. Зверніть увагу на те, що одиниці сукупності та одиниці спостереження можуть збігатися, наприклад, при вивченні суспільної думки населення це – кожний респондент – людина, яка висловлює свою думку. Але найчастіше обидві одиниці не збігаються. Так, у разі дослідження ефективності роботи друкарського обладнання одиницею спостереження є власник обладнання, а одиницею сукупності – окремий верстат.

Ознака – це статистичний еквівалент властивостей (рис, якостей), що притаманні елементам (одиницям) сукупності і відбивають їх сутність, характер та особливість. Наприклад: рівень акредитації, факультети, відділення, спеціалізації, термін та форми навчання – ознаки закладів освіти.

За характером вираження ознаки поділяються на:

атрибутивні (описові, якісні) – не мають кількісного вираження, а лише словесне. Не піддаються прямому числовому визначенню; їх можна виразити тільки числом одиниць в сукупності або реєстрацією наявності чи відсутності властивостей, наприклад: стать, професія, ім'я та інші;

кількісні (варіаційні) – мають числове вираження. На відміну від якісних вони можуть бути вимірюваними, тобто визначеними шляхом вимірів, зважувань і підрахунків (ціни, чисельність, вага, зріст, стаж тощо).

Варіація – це зміна значень конкретної ознаки при переході від однієї одиниці сукупності до іншої, інакше кажучи, це окреме значення (індивідуальний рівень) ознаки кожного елемента (одиниці) сукупності, наприклад, “стать” – чоловіча, жіноча, “оцінка” – 2,

3, 4, 5 тощо. Саме наявність варіації зумовлює необхідність статистики! Рівень значення ознаки у окремих елементах може вимірюватися за допомогою номінальної шкали (шкали якісних величин, найменувань), порядкової (рангової) шкали, метричної шкали (кількісних фізичних, економічних величин).

Статистична закономірність – це кількісна закономірність змін у просторі та часі масових явищ і процесів суспільного життя, які складаються з множини елементів (одиниць сукупності). Це форма прояву причинного зв'язку, виражена у послідовності, регулярності, повторюваності подій, якщо причини (умови), що викликали ці події не змінюються або мають незначні зміни. Характерною особливістю статистичної закономірності є те, що вона властива не окремим одиницям сукупності, а усій сукупності в цілому і виявляється лише в масових явищах при значній кількості одиниць сукупності. Пояснити це можна так: формування масового процесу відбувається під впливом основних причин (загальних) і причин випадкових (індивідуальних). Якщо кількість подій достатньо велика, вплив випадкових причин усувається, взаємно урівноважується (закон великих чисел). Наприклад, успішне навчання окремих студентів залежить від загальних для усіх студентів причин (організація навчального процесу, наявність в достатній кількості підручників, комп'ютерів та спецобладнання), а також від індивідуальних (бажання, здібності, стан здоров'я, матеріальні умови). Якщо дослідити рівень успішності багатьох студентів у нормальних (середніх) умовах навчання, то виявиться, що навчаються вони переважно добре, оскільки загальні причини є основними.

Статистичний показник – це узагальнююча кількісна характеристика властивостей сукупності загалом чи її частин зокрема щодо конкретних умов місця і часу. Статистичні показники поділяються на абсолютні, відносні та середні величини. Якісний зміст показника визначається суттю явища і знаходить своє відображення у назві. Кількісну сторону представляють число та розмірність (вимірник). Наприклад: середня заробітна плата працівників освіти України за січень 1999 року становила 168 грн. На відміну від значень ознак значення статистичних показників отримують розрахунковим шляхом. Це може бути простий підрахунок одиниць сукупності (кількість студентів академічної групи – 30 чол.), сума їх значень ознаки (з них: дівчат – 20 чол., хлопців – 10 чол.), порівняння двох або декількох величин (дівчат у 2 рази більше) чи більш складні розрахунки (частка дівчат у групі –67%). Важливо правильно визначити зміст статистичних показників та методи їх побудови.

4. Будь-яке статистичне дослідження послідовно проходить 3 етапи:

□ перший етап – збирання первинного статистичного матеріалу реєстрацією фактів чи опитуванням респондентів (метод статистичного спостереження);

□ на другому етапі зібрані дані підлягають первинній обробці, систематизації та групуванню – від характеристик окремих елементів переходять до узагальнюючих показників у формі абсолютних, відносних чи середніх величин (методи зведення та групувань, табличний та графічний методи);

□ третій етап передбачає аналіз отриманих під час зведення та групування матеріалів для одержання обґрунтованих висновків про стан явищ, що вивчаються, та закономірностей їх розвитку (методи вивчення варіації, диференціації та сталості, тенденцій розвитку, прогнозування, вивчення взаємозв'язків тощо).

Вказана послідовність є лише загальною, а конкретний зміст того чи іншого етапу залежить від мети дослідження та характеру даних. Передумовою використання статистичних методів має бути визначення і розуміння суті явища, що вивчається, його властивостей, особливостей конкретних обставин.

5. **Статистичне спостереження** є першим кроком статистичного дослідження і полягає в планомірному, науково-організованому збиранні даних. Формування якісної інформаційної бази – це фундамент статистичного дослідження, оскільки використання лише об'єктивної та достатньо повної інформації на подальших етапах дослідження надає можливість отримати правильні, обґрунтовані висновки про характер і закономірності досліджуваного процесу. Але не всі вихідні дані можна покласти в основу узагальнень і висновків, тому статистичні дані, що придатні для цього повинні бути:

□ вірогідними і точними – статистичні дані мають доказову силу лише тоді, коли вони правдиві і достовірні;

□ повними, не випадковими чи уривчастими; дістають їх реєстрацією значень ознак усіх одиниць сукупності за необхідний період чи на певний момент часу;

□ однотиповими, порівнюваними – для їх узагальнення і зіставленості у часі і просторі.

Слід звернути увагу на те, що статистичні дані від інших даних відрізняє масовість! Лише завдяки переходу від окремих фактів до масових можна визначити загальну закономірність, позбавлену впливу випадкових причин.

Процес проведення статистичного спостереження здійснюється в три етапи:

□ підготовка спостереження (складається план: визначаються мета, об'єкт, одиниці дослідження, способи збирання даних);

□ проведення масового збору інформації (реєстрація установлених фактів; саме забезпечення реєстрації відрізняє статистичне спостереження від спостереження);

□ формування бази даних (контроль та нагромадження даних спостереження, а також їх збереження).

Мета спостереження – основний очікуваний результат статистичного дослідження. Вона повинна мати чітке формулювання. Невизначеність поставленої мети може призвести до зібрання зайвих або неповних даних. Мета спостереження визначає його об'єкт – деяку сукупність явищ, що підлягають обстеженню. Чітке визначення суті, істотних ознак та меж об'єкта дозволяє запобігти різному тлумаченню результатів обстеження. Наприклад, для дослідження рівня безробіття, необхідно визначити які категорії населення можна віднести до безробітних. Якщо вважати за таких усіх хто займається пошуками роботи, то рівень безробіття буде високим, якщо ж безробітною визначити людину, яка протягом 3-х місяців не має роботи, то рівень безробіття буде низьким. Уявлення про об'єкт дослідження можна отримати лише при наявності інформації, яка характеризує його окремі складові. Тому в ході спостереження виокремлюють одиницю спостереження – первинну, звітну одиницю, від якої одержують інформацію та одиницю сукупності – первинний елемент об'єкта статистичного спостереження, який підлягає обстеженню і реєстрації. Іноді ці одиниці збігаються, іноді ні. Для досягнення більшої точності результатів дослідження та можливості застосування широкого спектра засобів аналізу, під час реєстрації слід врахувати, що ознаки за формою вираження можуть бути як кількісними, так і атрибутивними; тому, добираючи шкалу їх вимірювання, перевагу слід віддати не тільки більш інформативним ознакам (номінальній шкалі), а й ознакам із ширшими можливостями статистичної обробки (порядковій та метричній шкалам). Щоб спостереження дало вірогідні та своєчасні дані, необхідно вирішити питання часу та періоду спостереження. Час спостереження (об'єктивний час) – це час, до якого належать дані спостереження. Період спостереження (суб'єктивний час) – час, протягом якого реєструються дані. Зібрана інформація відображується в обліковому документі – статистичному формулярі у вигляді анкети, картки, звіту, опитувального листа.

При статистичному дослідженні використовують три форми спостереження:

□ статистична звітність; статистичний реєстр;

□ спеціально організоване спостереження (охоплює ті явища і процеси, які знаходяться поза звітністю – екологічні дослідження, результати моніторингу, опитування тощо).

Види статистичного спостереження розрізняють:

□ за повнотою охоплення одиниць сукупності це – суцільне спостереження (реєструються усі без винятку одиниці сукупності, наприклад, переписи);

□ несучільне спостереження (реєструється певна частина одиниць сукупності); залежно від завдань дослідження і характеру об'єкта несучільне спостереження може здійснюватися способом основного масиву, монографічним, вибіркоким, анкетним, через моніторинг; при цьому має забезпечуватися репрезентативність відібраної частини елементів, тобто їх здатність відтворювати властивості усієї сукупності елементів (висновки роблять за відібраною частиною про ціле!); найпоширенішим є вибіркоче спостереження;

□ за часом реєстрації даних це – поточне (систематична реєстрація фактів щодо явищ у міру їх виникнення, наприклад, робота транспорту, запис актів громадського стану);

□ періодичне (проводиться через певні, як правило, рівні проміжки часу, наприклад, переписи);

□ разове (виконується за необхідністю, наприклад, вивчення думки читачів часопису);

□ за способом одержання відомостей (первинних статистичних даних) це – безпосередній облік (дослідник особисто реєструє факти, наприклад, метеорологічні спостереження);

□ документальний облік (джерелом відомостей є відповідні документи);

□ опитування (відомості фіксуються зі слів респондентів); може здійснюватися експедиційним способом, через самореєстрацію та кореспондентським способом).

Різні види та форми статистичного спостереження можуть бути поєднані, взаємно доповнювати одне одного. Але слід взяти до уваги, що анкетний, кореспондентський способи, самореєстрація не забезпечують відповідної якості матеріалу в силу незацікавленості частини респондентів в опитуванні (повертається не більше 80% анкет), неможливості безпосередньої перевірки достовірності одержаних відповідей. Ці методи найбільш поширені у соціальних дослідженнях, при вивченні громадської думки, коли не вимагається висока точність, а потрібно лише зорієнтуватися в тому чи іншому питанні. Високу точність результатів дослідження гарантують вибіркочий, експедиційний методи за умов їх правильної організації.

Моніторинг – це не зовсім традиційне статистичне спостереження, проте воно є важливим джерелом статистичних оперативних даних для прийняття рішення.

Помилки спостереження – це розбіжності між даними спостережень та реальними даними. Залежно від причини виникнення розрізняють помилки:

□ репрезентативності – властиві тільки несучільному спостереженню і виникають в результаті некоректного формування відібраної сукупності, яка недостатньо повно відтворює склад усієї досліджуваної сукупності;

□ реєстрації – виникають при будь-якому спостереженні внаслідок неправильного встановлення фактів або невірного їх запису і бувають випадковими (описки, незнання, неухважність) або систематичними (навмисне викривлення фактів, приховування, некомпетентність, неосвіченість). Зверніть увагу, що систематичні помилки є більш небезпечними. Вони приводять до викривлення загальних підсумків, оскільки завжди мають однакову тенденцію чи до збільшення, чи до зменшення значень показників по кожній одиниці спостереження. Тому величина показника по сукупності в цілому включатиме накопичену помилку і взаємно не урівноважується. Наприклад: округлення віку респондентами, замість 49 чи 51 кажуть 50. Випадкові помилки виникають внаслідок дії випадкових факторів, наприклад, цифри переставлені місцями. Вони можуть відхиляти дані

спостереження в бік збільшення або зменшення. Проте деякою мірою взаємознищуються у відповідності з дією закону великих чисел і не є небезпечними. Помилки, допущені на етапі збору даних не, можуть бути виправлені на 2-му та 3-му етапах спостереження.

Тому ще на стадії спостереження необхідно застосовувати засоби контролю даних, а саме:

- зовнішній – контроль повноти даних, здійснюється візуально при перевірці наявності даних за всіма одиницями та позиціями;

- арифметичний – базується на використанні кількісних зв'язків між значеннями різних показників і полягає у перевірці (перерахунку) усіх узагальнених показників; з його допомогою можна встановити розмір помилки та виправити її;

- логічний – базується на сумісності даних і полягає у зіставленні відповідей респондентів за їх логічним зв'язком, наприклад, порівнянні віку з сімейним станом, освітою; виду діяльності з джерелом засобів існування тощо.

Такий контроль лише встановлює наявність помилки, а не її розмір. Для виправлення помилок, що встановлені в результаті логічного та арифметичного контролю, треба повторно звернутися до джерела вихідних даних.