

Н. І. Ніпорко,  
викладач, Чернігівський державний технологічний університет

# ІСТОРІЯ РОЗВИТКУ ВИБІРКОВОГО МЕТОДУ

**У статті здійснена спроба узагальнити інформацію про історію розвитку вибірки та її роль у статистичних дослідженнях. Розглянуто етапи становлення та проблеми застосування імовірнісної вибірки в Україні.**

**In the article realizable attempt to generalize information about history of development of sampling method and its role in statistical surveys. The stage of becoming and problem of application of probabilistic sample are considered in Ukraine.**

*Ключові слова: вибірковий метод, випадкова вибірка, надійність, точність, випадкові спостереження, репрезентативна вибірка.*

*Key words: sampling method, random sample, precision, accuracy, sample survey, representative sample.*

## ВСТУП

З розвитком ринкових відносин стрімко зростають потреби в економічній, соціальній і управлінській інформації. Якщо обробку статистичних даних можна вести з використанням досконалих комп'ютерних програм, то збір первинної інформації, як і раніше, забезпечують інтерв'юєри та працівники обліку, кількість яких весь час збільшується. На сьогоднішній день головне завдання статистики не в зростанні фізичного обсягу зібраних даних, а в кількості і якості інформації, яку можна отримати при тому ж обсязі. В більшості країн світу для забезпечення зростаючих інформаційних потреб використовують вибіркові спостереження, особливо це стосується оцінки показників розвитку підприємств, сільськогосподарського виробництва, доходів і витрат домогосподарств. З теорії ймовірності та математичної статистики відомо, що при застосуванні однакової процедури відбору декілька разів можна отримати різні за складом вибіркової сукупності з близькими значеннями середніх показників. А оскільки сукупність домогосподарств неоднорідна і кількість характеристик, що оцінюються, досить велика, то не всі отримані показники надійні. Ці та багато інших питань викликали сумніви дослідників, особливо практиків, впродовж всієї історії розвитку теорії вибіркового методу.

## ПОСТАНОВКА ЗАДАЧІ

Статистики-науковці неодноразово повертались до етапів зародження вибіркового методу, оскільки виникнення теорії завжди викликано практичною необхідністю. Теоретично доведено, що на підставі тільки одного спостереження можна оцінити показники генеральної сукупності, і не тільки їх значення, а і похибки, з якими вони розраховані. В основі доведення лежить декілька теорем, що носять назву закону великих чисел. По суті, весь

розвиток теорії вибірки є нічим іншим, як спроба оволодіти законом великих чисел, розкрити механізм його дії та навчитися користуватися ним з метою найкращої організації дослідження. В статті здійснена спроба систематизувати та узагальнити інформацію про історичний розвиток вибірки та її роль в статистичних дослідженнях.

## АНАЛІЗ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Вивченню методів випадкового відбору та його застосуванню приділялось багато уваги, починаючи з 18 століття. Дослідження, описані в літературі, охоплювали населення окремих країн та не були пов'язані між собою ні методологією, ні порівняльністю результатів. Теорія і практика випадкового відбору до початку двадцятого століття були відокремлені. В літературі питання історії розвитку вибірки розглядалися Е. Енгелем, А.В. Чаяновим, А.Г. Ковалевським. Багато уваги приділив становленню вибіркового методу М.К. Дружинін. Докладно описав історію вибірових обстежень сімейних бюджетів у країнах Західної Європи, США та СРСР І.Я. Матюха. Своє відображення історія вибірових досліджень знайшла в наукових роботах В.С. Сичової та І.Ф. Девятко.

## ВИКЛАДЕННЯ ОСНОВНОГО МАТЕРІАЛУ

За словами Д. Бернала, прогрес в науці починається із рішення питань, поставлених економічною необхідністю. Потреби господарювання в першу чергу викликали розвиток вибіркового спостереження.

Переписи населення і ресурсів зародилися одночасно з першими формами політичної і соціальної організації. Ще у давнину потребувалась така інформація для здійснення різних управлінських задач, а найважливіше — фіскальної політики та визначення кількості осіб, придатних для військової служби. Перші переписи проводились у стародавньому світі: це Єгипет — 2800—2500 роки

до н.е., Китай — у 2238 році до н.е. В історії зустрічаються згадки про переписи населення в Японії, Індії та інших країнах. У книзі пророка Самуїла був описаний перепис населення в давньому Ізраїлі, здійснений царем Давидом в X ст. до н.е., який закінчився страшною епідемією. У стародавньому Римі та Греції з 435 року регулярно проводились переписи.

З розвитком суспільної організації змінюється мета досліджень, до того ж вони приймають більш обмежений характер. Як правило це: визначення якості товару по окремій партії, оцінка урожаю по обмолоту окремої аграрної общини. В Англії в XIII столітті використовують дослідження окремої сукупності для контролю якості відчеканеної монети.

Історія розвитку людини супроводжується спробами зафіксувати не тільки історичні події, а і стан суспільства, оцінити його та описати в різних вимірах. Аналіз суспільних процесів і явищ філософами, істориками, науковими діячами мав загальний, приблизний характер, оскільки у них не було можливості підтвердити результати конкретними кількісними оцінками масових явищ.

Починаючи з кінця XVI століття, зацікавленість дослідників господарським і суспільним життям зростає. В 1598 році в Англії Джоном Стоу здійснено статистичний опис державних установ. В 1676 році виходить книга У. Петті "Політична арифметика", де вперше згадується про доходи і витрати сімей. Питання аналізу бюджетів робітників у цей час ставили Г. Кінг, Й. Мессі, правда більше з політичною метою. Масові бюджетні обстеження започаткував Давид Девіс, він вивчав причини і стан бідності, використовуючи для збору інформації чітку програму з розробленими опитувальниками.

У кінці XVIII століття та на початку XIX майже всі спостереження за доходами і витратами сімей були пов'язані з економічною і політичною ситуацією в Європі: голод, криза, заворушення, революції. В цей час розвиток ідеї бюджетних спостережень захопив велику кількість учених. Розвивалася і наукова сторона досліджень, які проводились з розробкою планів спостережень та застосуванням нових методів обробки інформації.

Перший міжнародний статистичний конгрес в Брюсселі (1853 р.) поставив питання про одиницю відбору бюджетних спостережень та започаткував спадкоємність у використанні методів і результатів європейських досліджень.

У середині XIX століття бельгійський учений Є. Дукпете і француз Ф. Ле Пле розробили методологічні основи бюджетного спостереження для вивчення рівня життя народних мас. В цей час були сформульовані закони масового споживання Е. Енгеля і Швабе.

З кожним днем перед дослідженнями рівнів доходів і витрат вставало все більше завдань. Вони проводились майже в усіх країнах Європи і США, але внаслідок спонтанного відбору результати спостережень були нерепрезентативні по ряду показників. Ще довгий час практики розробляли в першу чергу способи застосування планомерно організованого типового відбору, тоді як паралельно їм вчені-математики розвивали теорію випадкової вибірки.

Вперше положення теорії ймовірності поєднав з практичними питаннями вибіркового методу М.В. Остроградський [2]. Користуючись апаратом середньої математики він розрахував надійність вибірки армійських то-

варів (5—7% всієї сукупності), відібраних наугад для вивчення їх якості. Майже одночасно, в 1846 році, була опублікована робота Д.П. Журавського про проблеми вибіркового методу з економічної точки зору. Математичну основу вибіркового методу запропонував П.Л. Чебишев в праці "Про середні величини" (1867). За допомогою оригінального методу нерівностей великий математик довів теорему, що обґрунтовує закон великих чисел. Її можна сформулювати так: при достатньо великому обсязі абсолютно випадкової вибірки з ймовірністю близькою до одиниці можна стверджувати, що середня відібраної сукупності буде дуже мало відрізнятись від середньої генеральної сукупності. Пізніше учень П.Л. Чебишева А.М. Ляпунов довів, що у випадку великого числа випадкових величин їх розподіл підкоряється закону нормального розподілу, цей закон тепер називається законом Ляпунова — Лапласа — Гауса і він є основою більшості розрахунків вибіркового методу.

П.С. Лаплас в 1786 році, застосувавши свою урнову модель, використав вибіркового методу для оцінки населення Франції. Йому було відомо населення відібраних районів країни і число щорічних народжень в районах і по всій Франції. Приймаючи відношення народжень до населення за постійне, П.С. Лаплас розрахував необхідний показник всього населення, при цьому на основі своїх формул він вирахував погрішність, яка виявилась незначною. П.С. Лаплас був першим, хто оцінив точність вибіркового методу [2].

К.Ф. Гаус в 1809 році розрахував імовірнісну помилку середнього значення і вивів формулу для розрахунку варіації нормального кривої розподілу.

Застосування вибіркового методу для вивчення випадково розподілених ознак, наприклад, доходу чи розміру сім'ї, були вперше обґрунтовані в роботах норвежця А.Н. Киера, англійців А.Л. Боулі і К. Пірсона, а в Росії — А.І. Чупрова.

Статистики пов'язують початок розвитку наукової теорії вибіркового методу з іменем А.Н. Киера — директора норвежського статистичного бюро. В своїй доповіді Міжнародному статистичному інституту в кінці XIX століття він обґрунтував необхідність впровадження вибіркового спостереження в статистичну практику, але в цей час мова про теорію вибіркового методу на основі теорії ймовірності не йшла. В 1901 році В. Борткевич заявив, що розрахунок ймовірності повинен стати теоретичною основою вибірки, і тільки А.Л. Боулі вирішив задачу випадкових помилок вибіркового спостереження в 1906 році, використавши математичні роботи К. Пірсона і Ф.А. Еджворта. В "Елементах статистики" А.Л. Боулі на прикладах відбору пробних зразків у торгівлі і промисловості зробив висновок, що тільки враховуючи рівність шансів для всіх одиниць сукупності бути включеними у вибірку, можна говорити про правильність організації відбору. В зарубіжній статистичній літературі цей висновок носить назву "Боулієвого методу".

Необхідно сказати, що в природничих науках в XIX столітті було досягнуто значних успіхів у вивченні взаємозв'язків процесів. Наука перейшла від збирання фактів до упорядкування. Було зроблено три відкриття: наука про клітину, закон збереження і перетворення енергії та еволюційна теорія Дарвіна. В фізиці відкрита кінематична теорія газів, що зумовило використання в цій науці методів теорії ймовірності. Стало зрозуміло, що картина

світу надзвичайно складна, всі процеси і явища взаємопов'язані, причинно-наслідкові зв'язки багатогранні.

У роботах статистиків-практиків також з'являються питання про недоцільність використання логічної побудови статистичних досліджень. Ще Кетле говорив, що пояснення статистичних фактів часто затруднюються тим, що складні явища розглядаються як прості. А.А. Кауфман підкреслював, що індукція в вивченні явищ за допомогою статистичних методів не завжди дає тверді надійні результати, оскільки не можливо врахувати вплив всіх взаємозв'язків. Ю.Е. Янсон вважав необхідним звернути увагу на те, що логічні способи Д.С. Милля, які він називав методами дослідження, можуть не спрацювати при статистичному спостереженні внаслідок дії багатогранності взаємозумовлених причин. Найбільш детально ці питання розглянув А.А. Чупров. Він довів, що в практиці статистик зустрічається з явищами, характеристики яких зумовлені великою кількістю причин і наслідків, основою для їх вивчення може бути тільки ймовірність [4].

Незважаючи на розвиток теорії вибіркового спостереження, в практиці математико-статистичні методи не отримали визнання. В цей час особливої уваги заслугоували дослідження російських земських статистиків, які базувались на вивченні типових об'єктів: селянських поселень чи господарств за розміром землеволодіння на двір. Ідея такого відбору належала А.А. Чупрову, він вважав, що подібні дослідження можуть бути вдалимими, якщо вони базуються на даних суцільного спостереження, але математично підтвердити результати в такому випадку не можливо.

Ймовірнісний метод не визнавали не тільки в практиці спостережень, він не отримав підтримки і в Міжнародних органах статистики. Чотири з'їзди, починаючи з 1895 року по 1903, не знімали з програми своїх засідань питання вибіркового спостереження, так і не прийнявши остаточного рішення. Це можна пояснити тим, що, не дивлячись на авторитетність вчених, які розробили і удосконалили методи теорії ймовірності, цей метод викликав недовіру практиків із-за математичної складності розрахунків і неповного їх розуміння, оскільки сумніви і навіть заперечення результатів не мали доказів. Відбір типових одиниць давав результати, які здавалися нагляднішими і значно ближчими до середніх генеральної сукупності.

Тільки в 1925 році Міжнародний статистичний інститут прийняв рішення про застосування вибіркового методу в статистиці, схваливши два основних види вибірки: випадковий і цільовий відбір. Для випадкового відбору це повторна і безповторна вибірка. Під цільовим розуміли відбір районів (округів), репрезентативних у відношенні деяких ознак ("контролів"), відомих як для вибірки, так і для генеральної сукупності [2]. Спосіб "контролів" був запозичений у російській земській статистиці як найбільш надійний, оскільки репрезентативність вибірки перевірялась порівнянням отриманих середніх показників з показниками, розрахованими на основі даних одночасно проведеного суцільного спостереження.

Теорія ймовірності давала можливість оцінити похибку для повторного відбору. Оскільки в статистичних спостереженнях використовується не безкінечна основа, то виникло питання безповторної вибірки. Вперше коригуючий коефіцієнт для окремого випадку дисперсії безповторного відбору (коли доля одиниць сукупності має ознаку, яка вивчається) вивів К. Пірсон в 1899 році. Більш

загальну формулу дисперсії, для середньої кількісної варіюючої ознаки вивів А.А. Чупров. У своїй статті він відмітив, що безповторна вибірка забезпечує більш високу точність — ніж повторна. А.Л. Боулі опублікував цю формулу в 1920 році — як окремий випадок, коли сукупність на практиці обмежена, підкресливши при цьому, що в разі, коли генеральна сукупність безмежно велика, дисперсія безповторного відбору наближується до дисперсії повторного.

У 1920 році з'являється поняття розшарованої вибірки, його застосував А.Л. Боулі, підкреслюючи, що розшарування майже не покращує точність. С.Н. Берштейн доповнює А.Л. Боулі, що розшарована вибірка дає можливість наряду з сукупністю в цілому отримати характеристики окремих регіонів, а це має неабияке практичне значення і сьогодні.

Цікавим для використання в практиці є випадок, розглянутий А.Г. Ковалевським. У соціально-економічних дослідженнях варіація всередині кожного району може бути різною, а отже, доцільно брати обсяг вибірки в конкретному районі пропорційний обсягу варіації ознаки. Вибіркова середня для всієї сукупності повинна розраховуватись як зважена із середніх по кожному району, де вагами виступає загальна чисельність в районі, а не обсяг вибірки.

Районовану (синонім — розшаровану) вибірку А.Г. Ковалевський використовує для теоретичного обґрунтування механічної вибірки, яка не може бути різновидом випадкової і будується на зовсім інших основах. Але практично випадковий відбір майже неможливий, це змушує дослідників звертатись до механічного відбору як форми районованої вибірки. При неоднорідних сукупностях з нерівномірним розподілом ознаки, коли елементи розміщені близько один до одного, механічна вибірка рівносильна районованій з малим обсягом кожного району і відбором у ньому тільки однієї одиниці. У випадку, коли розподіл має випадковий характер, тобто одиниці розміщені так, немов перемішані, або виставлені жеребкуванням, механічна вибірка нічим не поступається випадковій безповторній нерайонованій вибірці і має таку ж точність.

Історія розвитку гніздової (синоніми — серійна, кластерна) вибірки розпочалася в 1887 році, коли комісія статистичного відділу Московського юридичного товариства прийняла резолюцію про доцільність спостережень у формі детального опису господарств та глибокого вивчення конкретних питань. У резолюції підкреслювалось, що господарства і поселення обов'язково повинні бути типовими і відібрані пропорційно окремим типовим групам. А.Н. Кієр на Бернській сесії Міжнародного статистичного інституту в 1895 році доповів про спеціальні прийоми збору інформації, які враховують типовість явищ і нічим не пов'язані з законом великих чисел. Він вважав, що такий підхід, оснований на репрезентативному методі, і дасть можливість на основі спостереження частини одиниць отримати характеристики сукупності в цілому. Це було ніщо інше, як гніздова типова вибірка, де чітко визначався відбір гнізд, але невизначеним залишався відбір одиниць всередині цих осередків.

Теорія і практика вибіркового методу розвивалась російськими статистиками навіть швидшими темпами ніж на Заході. Однак планове керівництво господарством у роки радянської влади, створення єдиної системи народногосподарського обліку сприяло розвитку суцільного

спостереження і обов'язкової державної звітності як необхідного елемента централізованого управління. В таких умовах розвиток вибіркового методу і застосування його на практиці призупинилось. Крім того, імовірнісний відбір викликав недовіру і навіть негативне ставлення зі сторони деяких спеціалістів. У 1954 році наукова нарада по статистиці спростувала тенденції неприйняття закону великих чисел та підкреслила значення математичної статистики в обґрунтуванні вибіркового методу [4]. Як результат вибіркового спостереження отримали в країні широке розповсюдження, вони проводились органами державної статистики, а також окремими науково-дослідними установами. Недоліком на даному етапі розвитку можна назвати недостатню наукову обґрунтованість організації відбору та відсутність в більшості випадків оцінки репрезентативності. Не було в цей час і фундаментальних праць, які б пов'язували теорію вибірки з практикою спостережень.

Першим посібником із застосування вибіркового методу в статистиці стала робота Ф. Йейтса "Вибірковий метод в переписах і обстеженнях", написана у зв'язку з рекомендаціями Підкомісії ООН по статистичній виборці, висунутими на першій сесії в Лейк Саксесі в 1947 році [5]. Як вчений Ф. Йейтс є представником англійської школи математичної статистики. Цінність роботи для практиків полягає в тому, що складні математичні схеми супроводжуються простими практичними числовими прикладами. Проблеми, що виникають в процесі підготовки і реалізації відбору, автор розглядає в тій же послідовності, в якій вони зустрічаються.

У СРСР основними напрямками проведення вибіркового спостережень були дослідження бюджетів сімей та сільськогосподарського виробництва. Методологічні особливості формування вибірки в публікаціях не роз'яснювались. Серед робіт опублікованих в цей час заслуговують на увагу праці А.Г. Волкова, І.Я. Матюхи та І.С. Кільдишева.

Сучасний етап розвитку вибіркового спостережень в Україні, як і в інших країнах СНД, розпочався з моменту розпаду СРСР. Виникла потреба перебудови статистичної системи, що не могла не реагувати на вимоги міжнародних стандартів та економічні умови, що склалися в країні. Більшість спостережень переходять на вибірку основу, як більш економічну, більш оперативну і більш математично обґрунтовану, ніж діюча на той час. І знову, як і в XIX столітті, на вимогу практики з'являється цілий ряд публікацій з питань вибіркового спостереження. Це видання таких авторів, як О.І. Черняк, В.М. Пархоменко, А.М. Єріна, А.В. Головач, С.С. Герасименко, В.Г. Сариогло. На особливу увагу практиків заслуговує монографія О.М. Гладуна "Вибіркові обстеження населення: методологія, методика, практика". Автор створив посібник для організації вибіркового спостережень як органами державної статистики, так і маркетинговими компаніями.

Теорія і практика вибіркового методу сьогодення ставить перед дослідниками питання одержання надійних оцінок показників для "малих територій". Теоретичною основою "малих вибірок" є відкритий В. Госсетом в 1908 році закон розподілу вибірових середніх із нормально розподіленої генеральної сукупності, названий "законом розподілу Стюдента". Він застосовується при будь-якій чисельності вибірки, але тільки у випадку сукупності з нормальним розподілом. Питаннями "малих вибірок" займався Р.А.Фішер, якому належить розробка дисперсійного аналізу.

У статистичній практиці питання "малих територій" зумовлене дослідженням процесів соціального розвитку в регіонах, містах, конкретних округах. Дані отримані на обласному рівні не можуть бути використані для характеристики показників розвитку міст і районів області. В рамках програми "Реформування тарифів та реструктуризації комунальних підприємств в Україні" в 2002 році було підготовлено статистичний бюлетень, який є першою спробою оцінити показники доходів і витрат населення "малих територій" в Україні. Але складність методики оцінювання показників, відсутність програмного забезпечення та висококваліфікованих фахівців на місцях обмежує використання цього методу на регіональному рівні.

## ВИСНОВОК

Впровадження вибіркового методу на основі ймовірнісного відбору в статистичну практику поставило перед дослідниками цілий ряд питань, пов'язаних з проблемою надійності даних. Вибіркова сукупність домогосподарств формується для отримання репрезентативних характеристик на державному рівні. Результати аналізу щодо надійності окремих даних обстеження умов життєдіяльності домогосподарств 2000 року показали, що лише 10% показників щодо витрат населення на продовольчі товари можна вважати надійними, а на непродовольчі товари — тільки 2% [1]. Показники, звичайно, дуже низькі, але жодний інший метод відбору не забезпечує математичного розрахунку оцінки репрезентативності. Регіональні органи статистики, які проводять обстеження, не мають можливості безпосередньо використовувати отримані показники, які в чистому вигляді до того ж викликають недовіру. Практика типологічного відбору створювала видимість відповідності отриманих даних реальності життя, але змінились умови життєдіяльності, зросла диференціація доходів і витрат домогосподарств. Населення, сім'я, домогосподарство — економічні категорії з багатограними характеристиками, що змінюються в часі, врахувати все в типологічному відборі неможливо. Історія вибіркового методу та застосування його в обстеженнях населення ще раз переконує в необхідності ймовірнісного відбору як єдиного, що може забезпечити рівні можливості домогосподарств попасти у вибірку і як наслідок отримати репрезентативні дані по основних показниках та підтвердити їх математичними розрахунками.

## Література:

1. Гладун О.М. Вибіркові обстеження населення: методологія, методика, практика: монографія. — Ніжин: ТОВ "Видавництво "Аспект-Поліграф", 2008. — 348 с.
2. Гранков В.П. Выборочное наблюдение. — Изд.2-е, перераб. и доп. М.: Госстатиздат, 1963.
3. Девятко И.Ф. Методы социологического исследования. — Екатеринбург: Изд-во Урал. ун-та, 1998. — 208 с.
4. Дружинин Н.К. Выборочное наблюдение и эксперимент. — М.: Статистика, 1977. — 176 с.
5. Йейтс Ф. Выборочный метод в переписах и исследованиях. — М.: Статистика, 1965. — 433 с.
6. Матюха И.Я. Статистика бюджетов населения. — М.: Статистика, 1967. — 247 с.
7. Чайнов А. В. Бюджетные исследования. История и методы. — М.: Новый агроном, 1929.

*Стаття надійшла до редакції 30.03.2010 р.*