

Сутність та використання методів статистичного аналізу для характеристики динаміки фінансових інструментів

У статистичному аналізі загалом, та, фондового ринку, зокрема, важливим є дотримання певної послідовності етапів. Це значною мірою сприяє якості результатів аналізу та раціональному використанню ресурсів для його проведення.

Найпершим і одним з найважливіших етапів будь-якого статистичного аналізу є визначення структури даних, їх класифікації, розподілу і перетворення даних. Для коректного застосування того чи іншого методу особливу увагу в проведенні статистичного аналізу приділяють виду досліджуваної випадкової величини та законам їх розподілу. Основою статистичного аналізу є сукупність інформації щодо досліджуваного об'єкта (дані). Їх класифікація потрібна для того, щоб обрати найбільш доречний для їх аналізу метод. Набори даних класифікують за такими критеріями:

- 1) кількість ознак, якими описують досліджуваний об'єкт (одномірні і багатомірні дані);
- 2) можливість упорядкування даних в часі (просторові дані, динамічний ряд, просторово-динамічні дані);
- 3) шкала виміру ознаки (якісні (номінальні та порядкові), кількісні (дискретні та неперервні));
- 4) способи отримання інформації (первинні та вторинні дані).

Одномірні дані (одна ознака) містять інформацію лише про одну ознаку щодо кожної одиниці сукупності. Для характеристики цієї ознаки використовують такі методи: розрахунок середніх показників та показників варіації; групування даних і побудова варіаційних рядів; графічне представлення; класифікація та виявлення аномальних значень.

Багатомірні дані (дві і більше ознаки) містять значення декількох ознак щодо кожної одиниці сукупності (наприклад, на сайті АТ «ФОНДОВА БІРЖА ПФТС» є інформація про останні офіційні котировки на ФБ ПФТС за результатами торгової сесії: код, назва, остання офіційна котировка, тип, площадка). Такі дані аналізують за допомогою наступних методів: визначення основних характеристик за кожною ознакою; аналіз наявності та ступеня залежностей між цими ознаками; вивчення видів залежності однієї змінної від інших; класифікації спостережень для отримання однорідних груп та виявлення аномальних значень ознаки; аналіз динамічних рядів та прогнозування.

Якісні (атрибутивні) дані характеризують властивості одиниці статистичної сукупності, які не визначити кількісно. Такі ознаки вимірюють за номінальною та порядковою шкалами (табл. 2.1).

Кількісні (числові) дані – це показники, виражені числами, їх отримують через вимірювання або підрахунки. Вони вказують на обсяги явища. Такі ознаки вимірюють за метричною шкалою і їх поділяють на дискретні та неперервні (табл. 2.1).

Відповідно до іншої класифікації ознаки метричної шкали поділяють на ознаки адитивної та відносної шкали. Відмінність між ними – це поняття істинного нульового значення. Його суть тому, що значення нуль за відотною шкалою – це відсутність явища (наприклад, заробітна плата), на відміну від ознаки за адитивною шкалою (температура 0° не означає її відсутність). Детальніше про шкали у [43, с. 11-12].

Дуже важливим аспектом користування числовою інформацією є розуміння принципів її збору та методології обчислення показників. Таке розуміння забезпечують метадані – опис статистичних концепцій, змінних та алгоритмів перетворення первинних даних в інформацію (табл. 2.1).

Таблиця 2.1 – Характеристика шкал вимірювання ознак

Назва	Опис	Показники описової статистика	Приклад	Представлення
<i>Якісні дані</i>				
Номінальна шкала	Назви та категорії, які не можна впорядкувати	Частота, мода	Вид економічної діяльності, організаційна форма бізнесу	Стовпчикова, секторна діаграми;
Порядкова шкала	Назви та категорії, які можна впорядкувати	Частота, мода, медіана	Рівень ризику від провадження бізнесу: високий, середній, незначний	атрибутивний ряд розподілу; таблиця спряженості
<i>Кількісні дані (метрична шкала):</i>				
Дискретні ознаки	Цілі числа	Частота, мода, медіана, показники варіації, асиметрії та ексцесу	Кількість філій або місцевих одиниць у підприємства, кількість різних видів економічної діяльності	Стовпчикова діаграма, гістограма, боксплот ряді розподілу: дискретний та неперервний
Неперервні ознаки	Набуває будь-яких значень у визначених межах		Чистий прибуток/збиток, середня вартість активів, власний капітал	Лінійна діаграма; діаграма розсіювання, ряди розподілу

Джерело: Складено за: [43, 44, 45, 46, 47]

Далі у цьому розділі детальніше розглянемо прийоми попереднього етапу статистичного аналізу кількісних даних: створення рядів розподілу, побудова гістограм та боксплотів, обчислення показників описової статистики, цензурування і перетворення статистичних сукупностей.

Перше уявлення про досліджувану сукупність дає її візуальне представлення. Тому першим кроком для аналізу числової інформації є створення ряду розподілу та побудова гістограми. Ряд розподілу – це сукупність, впорядкована у групи за певною ознакою. Це таблиця, яка складається з двох стовпців – варіант (x) і частот (f).

Для аналізу ряду розподілу використовують такі аналітичні показники:

- 1) середина інтервалу – півсума верхньої та нижньої меж інтервалу;
 - 2) частка (w_i) – відносна частота; частота, розділена на обсяг сукупності;
 - 3) кумулятивна частка ($\text{Cum } i$) – сума часток даного інтервалу та всіх попередніх;
 - 4) щільність інтервалу (g_i) – відношення частоти інтервалу до його ширини;
- її використовують для рядів розподілу з нерівними інтервалами.

Залежно від виду ряду розподілу варіанти є категоріями (атрибутивний ряд розподілу), цілими числами (дискретний) або інтервалами (інтервальний ряд розподілу).

Створення атрибутивного та дискретного рядів розподілу передбачає підрахунок кількості одиниць сукупності із заданим значенням варіанта (детальніше про це, наприклад у [43, с. 83-86]). Для побудови інтервального ряду розподілу необхідно вирішити низку питань. Перша за все, визначити, скільки інтервалів доцільно утворити та якою має бути їх ширина. Використання цих способів детально описані, наприклад, у [43, с. 72-76, 86-98].

У програмі Excel побудувати ряд розподілу можна декількома способами (табл. 2.2).

Таблиця 2.2 – Способи побудови ряду розподілу у програмі Excel

Спосіб побудови	Примітки
Формула масиву FREQUENCY	<ul style="list-style-type: none"> • треба задати межі інтервалів; • перевагою цього способу є те, що при зміні значень сукупності програма сама коригує розрахунки
Вбудований пакет аналізу Data Analysis / Нystogram	<ul style="list-style-type: none"> • інтервали Excel називає скринями (bins) • межі інтервалів краще також задати, хоча можливий варіант, коли Excel сам їх визначає як відношення розмаху варіації до кількості інтервалів; • на нашу думку, варто так задати інтервали самостійно, бо за замовчуванням Excel виставляє ширину першого інтервалу так, щоб його частота дорівнювала 1, тож ширина першого інтервалу буде відрізнятися від інших і це вже не буде ряд розподілу з однаковими інтервалами; • Excel будує гістограму, якщо вибрати опцію Chart output

Продовження таблиці 2.2

Спосіб побудови	Примітки
Функції DCOUNT, DAVERAGE та COUNTIFS, AVERAGEIFS	<ul style="list-style-type: none"> • DCOUNT, COUNTIFS – дають кількість одиниць сукупності, які відповідають заданим умовам; • DAVERAGE, AVERAGEIFS – дають середнє значення одиниць сукупності, які відповідають заданим умовам; • корисні для створення аналітичного ряду розподілу

Джерело: Складено за: [43-49]

Для представлення/візуалізації ряду розподілу використовують такі графіки: гістограму, полігон частот, кумуляту, огіву, боксплот (див. табл. 2.3).

Таблиця 2.3 – Способи візуалізації ряду розподілу у програмі Excel

Діаграма	Примітка
Гістограма	<ul style="list-style-type: none"> • показує розподіл даних в розрізі інтервалів значень (або часових періодів); • висота кожного прямокутника вказує на частоту, частку (щільність) інтервалу, над яким він знаходиться; • показує форму розподілу, пропуски у значеннях, нетипові одиниці сукупності; • якщо середини верхніх сторін прямокутників з'єднати відрізками прямої, утвориться полігон частот
Кумулята	<ul style="list-style-type: none"> • полігон кумулятивних частот; • використовують зокрема для виявлення квантилів розподілу; • якщо поміняти місцями ординати кумуляти, отримаємо огіву
Боксплот (коробковий графік)	<ul style="list-style-type: none"> • представляє: максимальне та мінімальне значення, перший та третій квантилі й медіану; • показує асиметричність розподілу та її напрям; • дозволяє виявити нетипові значення та розмах варіації; • можна представити декілька розподілів на одному графіку; • його можна розташувати горизонтально та вертикально; • варто розглядати разом із гістограмою розподілу

Джерело: Складено за: [43-49]

Показники описової статистики визначають на попередньому етапі статистичного дослідження. Разом з гістограмами вони забезпечують кількісне уявлення про тип та характер розподілу. Розрізняють такі групи показників: 1) центру розподілу; 2) варіації; 3) форми розподілу.

Показники центру розподілу для незгрупованих даних

До показників першої групи віднесемо найбільше та найменше значення, середні величини та квантилі розподілу (табл. 2.4).

Таблиця 2.4 – Показники центру розподілу

Показник	Функція Excel	Суть показника
Найбільше значення (max)	MAX, MAXA, LARGE	Найбільше значення в сукупності
Найменше значення (min)	MIN, MINA, SMALL	Найменше значення в сукупності
Мода (Mo)	MODE	Значення, яке в сукупності зустрічається найчастіше. Ряди розподілу поділяють на одно-, дво-, моно- та багатомодальні
Медіана	MEDIAN	Ділить сукупність на дві рівні частини
Середні: Арифметична	AVERAGE, AVERAGEA,	Середнє значення; початковий момент першого порядку
Гармонійна	HARMEAN	Обернена до середньої арифметичної з обернених величин
Геометрична	GEOMEAN	Корінь n-го ступеня з n множників
Квантилі: квартилі ($\xi_{0,25}$, $\xi_{0,5}$, $\xi_{0,75}$) децилі ($\xi_{0,1}$, $\xi_{0,2}$, ..., $\xi_{0,9}$) процентилі ($\xi_{0,01}$, $\xi_{0,02}$, ..., $\xi_{0,99}$)	QUARTILE PERCENTILE	Ділять сукупність відповідно на 4, 10 та 100 рівних частин
к-й початковий момент	–	Узагальнююча формула для розрахунку показників ряду розподілу

Джерело: Складено за: [43-49]

Показники варіації для незгрупованих даних

Показниками другої групи є розмах варіації, міжквартильна ширина, середнє лінійне відхилення та середнє квадратичне відхилення (табл. 2.5).

При виборі статистичних показників для аналізу слід враховувати поняття робастності. Це нечутливість до нетипових значень. Зокрема, медіана є центром розподілу і на неї ніяк не впливають аномально великі чи малі значення ознаки.

Таблиця 2.5 – Показники варіації для незгрупованих даних

Показник	Функція Excel	Суть показника
Розмах варіації	MAX - MIN	Різниця між найбільшим та найменшим по сукупності
Міжквартильна ширина	QUARTILE _{0,75} – QUARTILE _{0,25}	Різниця між найбільшим та найменшим 50 % сукупності
Середнє абсолютне відхилення	AVEDEV	Середнє абсолютних значень відхилень варіантів від середньої арифметичної
Дисперсія для генеральної сукупності	VAR	Середнє квадратів відхилень варіантів від середньої арифметичної
Дисперсія для вибірки	VARA	
Середнє квадратичне відхилення	STDEV	

Джерело: Складено за: [43-49]

Для розрахунку ж середньої арифметичної використовують всі значення ознаки, тому її величина дуже чутлива до нетипових значень. Тому при аналізі розподілів із сильною асиметрією як центр розподілу використовувати медіану.

Саме зазначені базові речі та підходи дозволяють виявити особливості статистичного аналізу даних на сучасному фондовому ринку з його основними етапами та підходами до систематизації показників.

Список використаних джерел

- 43.Матковський С. О., Марець О. Р. Теорія статистики : навч. посіб. Львів. : нац. ун-т ім. І. Франка, 2010. 534 с.

44. Сигел Э. Практическая бизнес-статистика : уч. пособие. Москва : Издательский дом "Вильямс", 2008. 1056 с.
45. Матковський С. О., Вдовин М. Л., Панчишин Т. В. Статистика : навч. посіб. Львів : нац. ун-т ім.І.Франка, 2010. 340 с.
46. Минько А. А. Статистика в бизнесе. Руководство менеджера и финансиста. Москва : Эксмо, 2008. 504 с.
47. Мхитаряна В. С. Статистика: учебник для бакалавров. Москва : Издательство Юрайт, 2013. 590 с.
48. Матковський С. О., Гринькевич О.С., Вдовин М. Л., Вільчинська О.М. Бізнес-статистика: навч. посібник. Київ : Алерта. 284 с.
49. Лапач С.Н. Статистика в науке и бизнесе : уч. пособие. Киев : МОРИОН, 2002. 640 с.