**Елементи математичної статистики**

**План:**

1. Предмет і методи математичної статистики.
2. Генеральна і вибіркові сукупності.
3. Статистичний розподіл вибірки.
4. Графічне зображення статистичних розподілів.
5. Середні вибіркові характеристики.
6. Предмет і методи математичної статистики

Математична статистика займається встановленням закономірностей масових явищ і процесів на основі результатів спостережень або експериментів над ними. Для виявлення таких закономірностей використовуються математичні методи:

* + збирання і систематизації статистичних даних;
  + наукового аналізу цих даних.

Математична статистика опирається на теорію ймовірностей. Застосування математичної статистики в економіці дозволяє будувати економічні моделі і оцінювати їхні параметри, перевіряти твердження (гіпотези) про властивості економічних показників і форми їх взаємозв’язку. Одержані результати є основою для економічного аналізу і прогнозування, дають можливість приймати обґрунтовані економічні рішення у бізнесі в умовах невизначеності.

1. Генеральна і вибіркові сукупності

Нехай потрібно дослідити сукупність однорідних об’єктів відносно деякої якісної чи кількісної ознаки. Якісні ознаки характеризують деякі властивості або стан об’єкта, наприклад, стать, професія, якість продукції тощо. Кількісні ознаки одержують в результаті вимірювання або спостереження і виражаються числами, наприклад, маса, об’єм, прибуток тощо. Кількісні ознаки позначаються через *Х*, *Y, Z*, ...; вони можуть бути дискретними або неперервними, тобто набувати будь-яких числових значень в деяких межах.

Можливі два способи дослідження ознаки сукупності: суцільне спостереження і вибіркове спостереження. При суцільному спостереженні вивчається кожний об’єкт сукупності, а при вибірковому – деякої частини, відібраної із всієї сукупності. При цьому вся сукупність об’єктів, яка підлягає дослідженню, називається *генеральною*, а та її частина, яка попала на перевірку або дослідження, *вибірковою сукупністю* або *вибіркою*. Число

об’єктів у генеральній сукупності і у вибірці називається їх *обсягами*. За результатами вивчення ознаки у вибірці робиться висновок про властивості генеральної сукупності.

Дослідження вибіркової сукупності дає такі переваги:

1. практичність (із-за великого обсягу неможливо охопити всю генеральну сукупність);
2. виграш у часі;
3. зменшення затрат праці;
4. достатньо високу достовірність;
5. економія продукції.

Інформація про генеральну сукупність, одержана на основі вибірки, завжди буде мати деяку похибку, оскільки ґрунтується на вивченні тільки частини об’єктів. Тому дуже важливо так організувати вибіркове спостереження, щоб згадана інформація була найбільш повною і давала правильне уявлення про генеральну сукупність. Неправильно сформована вибірка дасть викривлене уявлення про сукупність.

Розрізняють два способи відбору елементів генеральної сукупності у вибірку: *випадковий* і *невипадковий*. Випадковий відбір означає, що кожен об’єкт сукупності має рівний шанс потрапити до вибірки. Якщо сукупність містить групу об’єктів, то випадковий відбір повинен забезпечити представництво у вибірці об’єктів кожної із цих груп. Такий відбір називають *типовим*. Наприклад, продукція виробляється різними цехами, на різних конвеєрах. Тому при вивченні якості виробів підприємства доцільно робити вибірку із продукції кожного цеху і кожного конвеєра. Так, при наявності трьох цехів і двох ліній у кожному, продукцію доцільно розбити на шість груп.

Випадковий відбір, як правило, проводять за допомогою спеціальних таблиць випадкових чисел. Для цього елементи сукупності нумеруються і з таблиці випадкових чисел, відкритої навмання на довільній сторінці, виписуються підряд номери елементів, які повинні ввійти у вибірку. Останнім часом використовують комп’ютер, який генерує випадкові числа.

Серед невипадкових вибірок виділяють два види вибірок: *серійний* і *механічний*. Серійним називають відбір, при якому об’єкти вибираються із генеральної сукупності не по одному, а серіями, на які попередньо поділена сукупність. При механічному відборі елементи генеральної сукупності відбираються за наперед встановленим правилом. Наприклад, до 5%-вої вибірки входить кожен двадцятий член генеральної сукупності.

Зауважимо, що при дослідженні таких вибірок не завжди можна застосувати ймовірнісні методи.

1. Статистичний розподіл вибірки

Нехай *Х* – випадкова величина, яка уособлює певну кількісну ознаку об’єктів генеральної сукупності. Із цієї сукупності одержано вибірку ознаки

*X* : *X*1 , *X* 2 ,...*X n* , де Хі , *i* 1, *n*– незалежні однаково розподілені випадкові

величини, розподіл кожної із яких співпадає із розподілом *Х*. Різні значення,

*i*

які приймає кількісна ознака, наприклад,

*X* х *i* 1, *k* 

називаються

*варіантами* або *реалізаціями***.** Послідовність варіант, розміщених у зростаючому порядку

*x*1 *x*2 ... *xk*

називають *варіаційним рядом*. Якщо варіанта

*xi* зустрічається *ni*

раз у

вибірці, то число *ni* називають *частотою* варіанти *х і.* Відношення = називають *відносною частотою* варіанти, причому = 1. В залежності від того, яких значень можуть набувати варіанти, варіаційні ряди поділяються на *дискретні* та *неперервні (інтервальні).*

***Означення*.** Перелік варіант варіаційного ряду і відповідних їм частот або відносних частот називають *статистичним розподілом* вибірки:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *xi* | *x*1 | *x*2 | ... | *xk* |
| *ni* | *n*1 | *n*2 | ... | *nk* |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *xi* | *x*1 | *x*2 | ... | *xk* |
| *wi* | *w*1 | *w*2 | ... | *wk* |

**Приклад.** Проведено обстеження величини річного продукту (млн. грн.) на десяти підприємствах харчової промисловості: 2, 5, 1, 2, 2, 5, 1, 2, 1,

2.

Побудувати статистичний розподіл вибірки.

Розв’язання. Досліджується ознака *Х* – величина річного прибутку (млн. грн.) на підприємстві харчової промисловості. Маємо варіаційний ряд:

*x*1 1, *x*2

2, *x*3

5.

Будуємо статистичні розподіли частот

|  |  |  |  |
| --- | --- | --- | --- |
| *xi* | 1 | 2 | 5 |
| *ni* | 3 | 5 | 2 |

3

*n* *ni*

*i* 1

3 5 2 10;

|  |  |  |  |
| --- | --- | --- | --- |
| *xi* | 1 | 2 | 5 |
| *wi* | 0,3 | 0,5 | 0,2 |

і відносних частот



1. Графічне зображення статистичних розподілів

Графічна ілюстрація статистичних даних надає їм наочності і часто при- водить до спрощення аналізу цих даних. В залежності від вигляду варіа- ційного ряду і поставленої задачі будують різні графіки статистичних роз- поділів, зокрема *полігон, гістограму, кумулятивну криву.*

Розглянемо дискретний статистичний розподіл вибірки. Полігоном частот (або відносних частот) називають ламану лінію, відрізки якої з’єднують

точки з координатами *xi* , *ni* 

(або *xi* , *wi* ),

*i* 1, *k* .

У випадку неперервного статистичного розподілу гістограмою частот (або відносних частот) називають східчасту фігуру у вигляді послідовності

прямокутників з основами

*hi* *i* *i* 1

і висотами *fi* (або *gi*). Площа *Si і*-*го*

прямокутника

*Si* *hi* *fi*

*k*

*ni* .

Звідси, площа гістограми частот дорівнює

обсягу вибірки: *Si* *n*.

**Приклад.** На кондитерській фабриці автомат розливає рідкий шоколад у форми для одержання шоколадних плиток. Для контролю навмання відібрано 20 плиток із партії готової продукції. Результати обстеження наведені в таблиці:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Маса, г | 96-98 | 98-100 | 100-102 | 102-104 |
| К-ть плиток | 4 | 8 | 6 | 2 |

Побудувати: 1) полігон і гістограму відносних частот; 2) кумулятивну криву.

Розв’язання. Для кількісної ознаки *Х* – маса шоколаду у формах, маємо

*n*

неперервний статистичний розподіл. Обчислимо відносні частоти появи ознаки *Х* на відповідних інтервалах; маємо ряд розподілу:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Інтервал | 96-98 | 98-100 | 100-102 | 102-104 |
| *wi* | 0,2 | 0,4 | 0,3 | 0,1 |

*wi*  *i*

#### n

Довжина кожного інтервалу

*hi* *i*

*i*1 2,

*i* 1,4.

Визначаємо

щільності *gi*:

*g*1 

0,2

2

0,1;

*g*2 

0,4

2

0,2;

*g*3 

0,3

2

0,15;

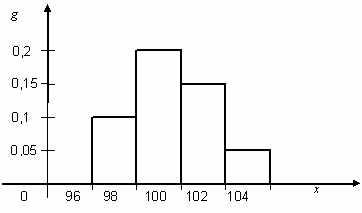
*g*4 

0,1

2

0,05.

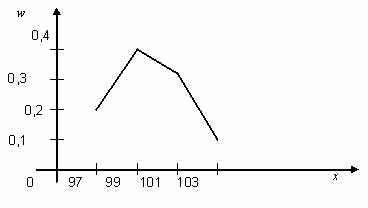
Гістограма відносних частот має вигляд .



Перейдемо від інтервального до дискретного статистичного розподілу від- носних частот

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Х* | 97 | 99 | 101 | 103 |
| *wi* | 0,2 | 0,4 | 0,3 | 0,1 |

Побудуємо полігон відносних частот (рис. 6.3).



1. Середні вибіркові характеристики

Середні величини характеризують значення ознаки, навколо якого концентруються варіанти статистичного розподілу (центральна тенденція розподілу). Існують різні форми середніх: арифметична, гармонічна, квадратична, геометрична тощо. Найбільш поширеною із них є середня арифметична.

Розглянемо дискретний статистичний розподіл вибірки обсягу *п*:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *xi* | *x*1 | *x*2 | ***...*** | *xk* |
| *ni* | *n*1 | *n*2 | ***...*** | *nk* |

Величина, яка обчислюється за формулою

називається *вибірковою середньою* такого розподілу (середнє арифметичне

значення).

За вибіркове середнє неперервного статистичного розподілу приймається середнє арифметичне відповідного дискретного розподілу.

Наведемо основні властивості вибіркової середньої.

1. . Якщо всі варіанти збільшити (зменшити) в одне і те ж число раз **, то вибіркова середня збільшиться (зменшиться) у тільки ж разів

2. Якщо всі варіанти збільшити (зменшити) на одне і те ж число *С*, то вибіркова середня збільшиться (зменшиться) на таке ж число

3. Сума всіх відхилень значень ознаки від вибіркової середньої дорівнює нулю:

4. Вибіркова середня алгебраїчної суми кількох ознак дорівнює сумі вибіркових середніх цих ознак

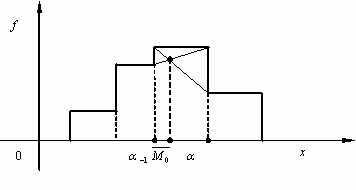
Крім вибіркової середньої, для аналізу статистичного розподілу застосовують *структурні* або *порядкові середні.* Із них найбільш поширені мода і медіана.

***Означення.*** *Модою*

*M* 0 статистичного розподілу називається варіанта, яка

має найбільшу частоту появи.

Але простіше моду можна знайти графічно за гістограмою частот: визначається абсциса точки перетину діагоналей, які з’єднують вершини прямокутника, що належить модальному інтервалу, із найближчими вершинами двох прилеглих прямокутників (рис. 6.5).



***Означення.*** *Медіаною* статистичного розподілу називають варіанту, яка припадає на середину розподілу варіаційногоряду.Для дискретного варіаційного ряду із непарною кількістю елементів медіана дорівнює серединній варіанті, а для ряду із парною кількістю елементів – півсумі двох серединних варіант.

Питання для самоконтролю

1. Дайте поняття генеральної та вибіркової сукупностей.
2. Які переваги дає дослідження вибіркової сукупності?
3. Що називають варіантами, варіаційним рядом?
4. Що таке частота, відносна частота варіантів?
5. Що таке полігон частот, відносних частот?
6. Що називають гістограмою частот, відносних частот?
7. Що називають модою, медіаною?