

Статистичні розподіли. Побудова характеристик статистичних підрозділів.

§1. Генеральна та вибіркова сукупності. Основні види вибірок.

Нехай потрібно вивчити задану (як правило, багаточисельну) сукупність об'єктів відносно деякої ознаки. Наприклад, потрібно визначити, в якій степені параметри виробів, що виготовляються, відповідають стандартним нормативам. Якщо число елементів в сукупності не дуже велике, обстеження об'єкту не пов'язане з його знищенням і не потребує великих витрат, то можна досліджувати кожний елемент окремо, фіксувати значення досліджуваної ознаки і відповідною обробкою результатів зробити той чи інший висновок про досліджувану ознаку.

Якщо ж сукупність складається із великого числа об'єктів, або дослідження пов'язане із знищенням об'єктів, або воно дороге коштує, то дослідження всіх об'єктів недоцільне. Недоцільно, наприклад, досліджувати настрок роботи всі лампочки даної партії, оскільки в результаті вся партія знищилася б. В таких випадках висновки про досліджувану ознаку роблять на основі вивчення обмеженої кількості об'єктів, належним чином вибраних із загальної сукупності.

Так як практично будь-яка ознака допускає кількісну оцінку, то, замість того, щоб говорити про сукупність досліджуваних об'єктів, можна говорити про значення досліджуваної ознаки цих об'єктів. Основними поняттями математичної статистики є генеральна та вибіркова сукупності.

Означення 1. Генеральною сукупністю називається множина числових значень деякої ознаки всіх об'єктів розглядуваної сукупності.

Означення 2. Вибірковою сукупністю, або просто вибіркою, називається множина числових значень деякої ознаки всіх об'єктів, випадковим чином вибраних із всієї сукупності розглядуваних об'єктів.

Наприклад, генеральною сукупністю є сукупність чисел, що відповідають строкам роботи всіх лампочок виготовленої партії (ці числа в дійсності залишаються невідомими), а вибірковою сукупністю буде сукупність чисел, що відповідає строкам роботи відібраних для випробувань лампочок (числа цієї сукупності визначаються із проведеного дослідження). Якщо ж вивчати не строк роботи, а, наприклад, вагу лампочки, що також має практичне значення, та сама сукупність об'єктів буде приводити до іншої генеральної сукупності і вибірки. Для простоти, якщо це не призводить до протиріч, тобто недвозначно відомо, про яку ознаку йде мова, від «генеральною вибіркою» і під «вибіркою» будемо розуміти саму сукупність досліджуваних об'єктів. Так, наприклад, будемо казати, що партія всіх електричних лампочок, що виготовлена заводом, являє собою генеральну сукупність, а множина лампочок, взятих для обстеження, складає вибірку сукупності. Підкреслимо, однак, що в математичній статистиці, як і завжди в математиці, ми абстрагуємося від конкретної природи об'єктів і вивчаємо лише їх абстрактні, а в даному випадку числові характеристики

Основну задачу математичної статистики можна сформулювати як задачу отримання обґрунтованих висновків про невідомі властивості генеральної сукупності за відомими властивостями виділеної з неї вибірки.

Число об'єктів сукупності (генеральної або вибіркової) називається об'ємом даної сукупності. Наприклад, якщо цех виготовив 2000 деталей, а для обстеження відібрано 150 деталей, то об'єм генеральної сукупності рівний 2000 ($N = 2000$), а об'єм вибірки рівний 150 ($n = 150$).

Розрізняють вибірки з поверненням і без повернення. Якщо після фіксованого значення параметра об'єкт повертається в генеральну сукупність і, таким чином, він може багато разів повторюватися у виборці, то говорять про вибірку з поверненням або з повтором. Якщо ж один раз вибраний об'єкт назад не повертається і він не може більше одного разу повторюватися у виборці, то така вибірка називається вибіркою без повернення, або без повтору. Відмітимо,

що коли об'єм вибірки набагато менший за об'єм генеральної сукупності, то відмінності між вибіркою з поверненням і без повернення практично немає.

Говорять, що вибірка репрезентативна (представницька), якщо вона достатньо повно представляє досліджувані ознаки генеральної сукупності.

Важливою умовою забезпечення репрезентативності вибірки є дотримання випадковості відбору, тобто всі об'єкти генеральної сукупності повинні мати рівні ймовірності потрапити у вибірку.

В наукових дослідженнях для формування вибірок найчастіше користуються так званими таблицями випадкових чисел, які можуть бути отримані за допомогою випадкового відбору карток або спеціальними алгоритмами на ЕОМ. Існують таблиці випадкових чисел чотиризначні, п'ятизначні і т. д. Якщо потрібно, наприклад, утворити вибірку об'єму n із генеральної сукупності об'єму $N = 100\,000$, то використовують п'ятизначну таблицю випадкових чисел і виписують будь-яку групу із n послідовних її чисел. Досліджуються ті об'єкти генеральної сукупності, номери яких співпадають ся з виписаними числами із таблиці випадкових чисел.

§2. Групування статистичних даних. Статистичні розподілі.

Дня визначення закономірностей масових випадкових явищ вивчаються статистичні дані, тобто відомості, отримані шляхом спостережень та досліджень, про значення потрібної нам ознаки. Розглянемо приклад обробки статистичних даних. Економіст, якого цікавить тарифний розряд працівників деякого підрозділу заводу, вибрав 100 документів працівників і виписав з них послідовність розрядів: 5, і, 4, 5, 4, 3, 6 і т. д. Ці статистичні дані являють собою вибірку, яка підлягає обробці.

Вивчення статистичних даних звичайно починається з їх групування в порядку зростання значення ознаки. Нехай в нашому прикладі, після впорядкування за зростанням статистичних даних, ми отримали ряд із 100 чисел 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 4, 4, 4, ..., де 1 повторюється 4 рази, 2 – 6 раз, 3 – 12 раз, 4 – 16 раз (для економії міста ми не вказали цей ряд повністю), 5 – 44 раз, 6 – 18 раз.

Означення 3. Спостережувані значення розглядуваної ознаки називаються варіантами, а послідовність варіант, записаних в порядку зростання називається вибіркоvim або варіаційним рядом.

Побудова варіаційного ряду