Лекція № 8

**ОСНОВИ статистичної обробки СУКУПНОСТІ**

**ВИПАДКОВИХ ВЕЛИЧИН**

*План лекції: Поняття про випадкові величини. Розподіл ймовірностей дискретної випадкової величини. Закон розподілу неперериваної випадкової величини. Нормальний розподіл безперервних випадкових величин. Розподіл Стьюдента. Інтегральна функція розподілу.*

**8.1 Поняття про випадкову величину**

Результат окремого спостереження при багаторазовому прямому вимірі будь-якої фізичної величини через наявність випадкових похибок являє собою випадкову величину.

Оскільки закономірностей у появі цих значень немає, аналіз таких величин може проводитися тільки методами теорії ймовірностей і математичної статистики. Для характеристики випадкової величини необхідно знати сукупність можливих значень цієї величини, а також ймовірності, з якими ці значення можуть з'являтися.

**8.2 Розподіл ймовірностей значень дискретної випадкової величини**

Для дискретної випадкової величини найбільш повної статистичної характеристикою є її розподіл ймовірностей: вказуються можливі значення цієї величини xi і відповідні їм ймовірності pi. Розташувавши значення x1, x2, ..., xn в порядку зростання і позначивши ймовірності p1, p2, ..., pn, отримаємо графік розподілу ймовірностей цієї дискретної величини (рис. 8.1). Сума всіх ймовірностей дорівнює 1. Найбільш вірогідне значення випадкової величини називається **модою** (x 'на рис. 8.1).



Рис. 8.1

**4.3 Закон розподілу неперервної випадкової величини**

Значення неперервної випадкової величини можуть відрізнятися один від одного як завгодно мало, тому ймовірність кожного з цих значень також безкінечно мала, і побудувати криву розподілу ймовірностей неможливо.

Щоб виявити розподіл ймовірностей в цьому випадку, розглядається деяка безліч інтервалів Δxi в діапазоні можливих значень даної випадкової величини, потім підраховують частоти ni попадання значень в кожен з цих інтервалів. Розташувавши, як і в попередньому випадку, значення xi в порядку зростання і позначивши відповідно ймовірності pi, отримаємо ступінчасту криву - **гістограму** (рис. 8.2а). Поєднавши середини верхніх відрізків гістограми ламаної кривої, одержимо **полігон частот**. Якщо взяти нескінченно малі інтервали (Δxi 0), графік втратить ступінчастий характер і перетвориться в плавну криву, звану кривою розподілу щільності імовірності f (x) для даної **неперервної випадкової величини** (рис. 8.2б). Рівняння, що описує цю криву, називається законом розподілу даної неперервної випадкової величини. Площа під всією кривою f (x) дорівнює ймовірності появи будь-якого з можливих значень xi, тобто дорівнює 1.



Важливим моментом є те, що існує оптимальне число інтервалів групування m, коли ступінчаста огинаюча гістограми найбільш близька до плавної кривої розподілу генеральної сукупності.



Рис. 8.2.а



Рис. 8.2б

Оптимальне число інтервалів можна розрахувати за формулами Старджеса

*m = 3,3lgn + 1*

Брукса і Каррузера

*m = 5lgn*

або Хайнольда і Гаеде

*m =.*

В області значень n <100 результати розрахунків за вищенаведеними фор-мулам близькі між собою.

**8.4 Нормальний розподіл безперервних випадкових величин. Розподіл Стьюдента. Інтегральна функція розподілу**

Найбільш поширеним для неперервних випадкових величин є нормальний розподіл з щільністю



де e - основа натуральних логарифмів;

       μ, σ - параметри розподілу.

Випадкові похибки багаторазових вимірювань зазвичай розподілені по нормальному закону.

Криві нормального розподілу (або функції Гаусса) (рис. 8.3) симетричні щодо ординати, що проходить через точку x = μ, і мають у цій точці єдиний максимум, рівний 1 (мода для нормального закону розподілу). При x = μ крива симетрична щодо осі ординат.

<https://uk.wikipedia.org/wiki/%D0%9D%D0%BE%D1%80%D0%BC%D0%B0%D0%BB%D1%8C%D0%BD%D0%B8%D0%B9_%D1%80%D0%BE%D0%B7%D0%BF%D0%BE%D0%B4%D1%96%D0%BB>



Рис. 8.3.



Рис. 8.4.

Графік функції Гауса означеної на двовимірній множині

<https://www.youtube.com/watch?v=Opy4FKVKzTo> 6 хв

<https://www.youtube.com/watch?v=r5-G9ZcCI6o> 13 хв

Нормальний розподіл (рис. 8.5) також називають розподілом Гауса, використання якого для обробки кінцевих сукупностей випадкових величин, якщо число n досить велике (n 30). І тут умовно вважають, що спостерігаються n значень величини X, тобто. x1, x2, …, xn є випадковою вибіркою з уявної нескінченної **генеральної сукупності.** <https://uk.wikipedia.org/wiki/%D0%9D%D0%BE%D1%80%D0%BC%D0%B0%D0%BB%D1%8C%D0%BD%D0%B8%D0%B9_%D1%80%D0%BE%D0%B7%D0%BF%D0%BE%D0%B4%D1%96%D0%BB#%D0%A1%D1%82%D0%B0%D0%BD%D0%B4%D0%B0%D1%80%D1%82%D0%BD%D0%B8%D0%B9_%D0%BD%D0%BE%D1%80%D0%BC%D0%B0%D0%BB%D1%8C%D0%BD%D0%B8%D0%B9_%D1%80%D0%BE%D0%B7%D0%BF%D0%BE%D0%B4%D1%96%D0%BB>



(рис. 8.5)

У статистиці малих вибірок (у мікростатистиці) велику роль грає інший розподіл безперервних випадкових величин - розподіл Стьюдента, щільність ймовірності якого визначається виразом:



де *Γ(α)* - гамма-функція (інтеграл Ейлера);

*tγ*-величина, що характеризує ступінь відхилення вибіркових статистичних характеристик від генеральних;

*k = n –1* - число ступенів свободи.

Значення гамма-функції для позитивного числа *b* можна обчислити за формулою:

*Γ(α) = (b - 1)!*

Графік розподілу Стьюдента (рис. 8.6) нагадує формою нормальний розподіл і зі збільшенням n наближається до нього дедалі більше (можна вважати, що з n > 30 обидва графіки практично збігаються).

<https://uk.wikipedia.org/wiki/T-%D1%80%D0%BE%D0%B7%D0%BF%D0%BE%D0%B4%D1%96%D0%BB_%D0%A1%D1%82%D1%8C%D1%8E%D0%B4%D0%B5%D0%BD%D1%82%D0%B0>



(рис. 8.6)