**Прямі регресії**

Нехай над двовимірним впорядкованим випадковим вектором .проведено в однакових умовах  пар незалежних спостережень: . Випадкові змінні  та  нам невідомі, тому на основі лише цих спостережень потрібно зробити висновок про кореляцію між ними.

Обчислимо середні арифметичні вибіркових значень координат цього вектора:

, 

та їх варіанси

 , .

Варіанси  є незміщеними оцінками дисперсій випадкових змінних  та , відповідно. Тому, за аналогією з рівностями (2), (3), (6) див. Кореляційний аналіз, **вибіркова коваріація між координатами вектора**  визначається так:

, (1)

**вибіркова кореляція** 

, (2)

**вибіркова регресія другої координати на першу**

 , (3)

**вибіркова регресія першої координати на другу**

. (4)

Вибірковим аналогом дисперсії випадкової змінної



є варіанса:

,

яка при



досягає мінімуму

. (5)

З (5) випливає, що . Зокрема, якщо , то = 0. Отже, всі –  приймають однакове значення , тобто всі спостережувані точки  лежать на одній прямій .

**Зауваження 1.** Якщо , то звідси не випливає, що і в генеральній сукупності коефіцієнт кореляції , тобто, що і в ній  та  пов’язані функціональним лінійним зв’язком. Дійсно, вибіркову кореляцію і всі висновки ми отримуємо на основі конкретного статистичного розподілу представленого парами точок. Якщо здійснимо ряд спостережень навіть того ж обсягу над тими ж випадковими змінними  та , то отримаємо, взагалі кажучи, інший статистичний розподіл, для якого . ■

Якщо , то проведемо через точку  пряму, яка мінімізує суму квадратів відхилень вибіркових значень вектора  в напрямку осі *ОY*. Для цього потрібно мінімізувати функцію

. (6)

Необхідна умова екстремуму функції (6) еквівалентна системі рівнянь:

, тобто 

Звідси остаточно маємо систему:

 (7)

Друге рівняння цієї системи еквівалентне рівнянню , тому . Підставимо цей вираз у перше рівняння системи (7). Тоді послідовно отримаємо:

= 0, ,

.

Справді,



,



,

.

Аналогічно доводимо, що

,

.

Отже



- розв’язок системи рівнянь (7).

Безпосередньо переконуємося, що

, .

Отже точка з координатами  є точкою мінімуму функції (6), тобто сума відхилень точок вибірки від прямої

 (8)

в напрямку осі *OY* є найменшою. Пряма (8) називається **прямою регресії другої компонента вектора  на першу**.

Аналогічно можна показати, що пряма

 (9)

мінімізує суму квадратів відхилень від точок даної вибірки, але в напрямку осі *OX* .

Проведемо через точку  пряму , яка мінімізує суму квадратів відхилень вибіркових значень вектора  в напрямку осі *ОX*. Для цього потрібно мінімізувати функцію

. ()

Пряма (9) називається **прямою регресії першої компонента вектора  на другу**.

**Зауваження 2.** Якщо вибіркова кореляція , то обидві прямі регресії першої компонента вектора  на другу та другої координати на першу співпадають і навпаки, - якщо ці прямі співпадають, то вибіркова кореляція .

Справді, обидва рівняння прямих регресії

 та 

при  еквівалентні одному й тому ж рівнянню

.

Якщо ж обидва ці рівняння прямих регресії співпадають, то

.

Отже тоді , тобто .