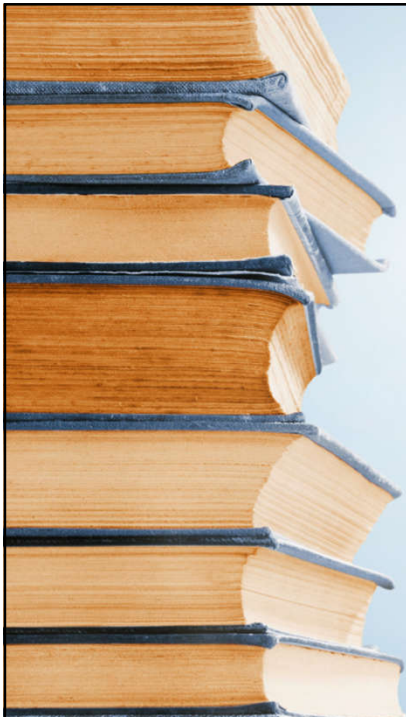




Кореляційно-регресійний аналіз

доц. Гутор Т.Г.



Всі явища суспільного життя і їх ознаки зв'язані між собою і залежать одне від одного. Взаємозв'язані ознаки діляться на факторні і результативні.

- **Факторною** називається ознака, під впливом якої змінюється залежна від неї інша ознака.
- **Результативною** називається ознака, яка змінюється під впливом факторної ознаки.



Наприклад

- Недостатня фізична активність може провокувати наявність надмірної маси тіла.
- В свою чергу надмірна маса тіла може обумовлювати появу серцево-судинних захворювань.
- Отже, в першому випадку "недостатня фізична активність" - це факторна ознака, а "надмірна маса тіла" - результативна ознака.
- В другому випадку "надмірна маса тіла" - це факторна ознака, а "серцево-судинні захворювання" - результативна ознака.



Види зв'язків у біостатистиці

- Функціональний зв'язок — залежність, при якій кожному значенню однієї ознаки відповідає єдине строго визначене значення іншої;
- Кореляційний зв'язок — це зв'язок поміж ознаками, коли значенню однієї величини однієї ознаки відповідає декілька значень іншої взаємопов'язаної з ним ознаки.



Функціональний зв'язок

- має місце при вивченні фізико-хімічних процесів явищ, в математиці, геометрії.
- Наприклад: зміна довжини сторони квадрата на певну величину приведе до зміни площі також на певну фіксовану величину.

$$a = 2 \text{ см}$$

$$S = 4 \text{ см}^2$$

$$a = 3 \text{ см}$$

$$S = 9 \text{ см}^2$$

$$a = 4 \text{ см}$$

$$S = 16 \text{ см}^2$$



Кореляційний зв'язок

- Є характерним для соціально-гігієнічних досліджень, а також досліджень в клінічній медицині.
- Наприклад, рівень температури тіла та число серцевих скорочень



Температура

36,6

37,6

38,7

ЧСС

68-72

78-82

87-93



Способи представлення кореляційної залежності



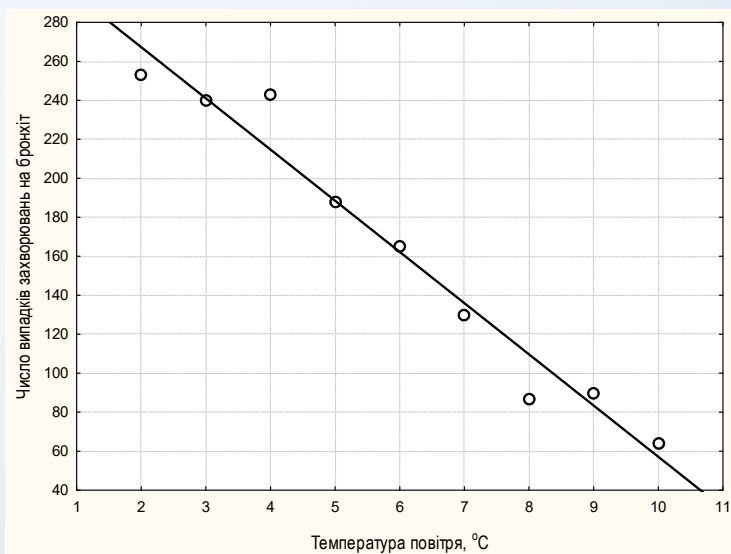
- таблиця,
- графік,
- коефіцієнт кореляції

Таблиця 1

Взаємозв'язок між температурою повітря
та числом випадків захворювання на бронхіт у місті
Львові у 2011 році



№ п/п	Середня температура повітря, °С	Число випадків захворювань на бронхіт, абс.
1	2	253
2	3	240
3	4	242
4	5	188
5	6	164
6	7	130
7	8	87
8	9	90
9	10	64



- Рис.1. Взаємозв'язок між температурою повітря та числом випадків захворювання на бронхіт у місті Львові у 2011 році



- Найпростішим способом представлення кореляційного зв'язку є розрахунок коефіцієнта кореляції – r .

Види кореляційних зв'язків



- Коефіцієнт парної кореляції відображає характер зв'язку лише двох однак. Він дає характеристику узагальненого "неочищеного" зв'язку між параметрами. При цьому можливий вплив інших факторів, які не враховуються.
- Множинний коефіцієнт кореляції – визначає взаємозв'язок між трьома та більше ознаками і показує сукупність впливу кожної з них. При цьому для відображення чистого взаємозв'язку між конкретним фактором та рівнем здоров'я, включаючи вплив інших застосовується *парціальний коефіцієнт кореляції*.

Кореляційна залежність характеризується



- направленістю,
- силою зв'язку,
- вірогідністю,
- формою



Направленість зв'язку

- **визначається за алгебраїчним знаком коефіцієнта кореляції**
- **Прямий зв'язок** (позитивний "+") динаміка параметрів є однонаправленою — збільшення одного параметра обумовлює збільшення іншого (зростання екологічного навантаження обумовлює зростання захворюваності населення)
- **Зворотний зв'язок** (негативний, від'ємний "-") — динаміка параметрів є різнонаправленою — збільшення одного параметра обумовлює зменшення іншого (при збільшенні віку дітей спостерігається зниження рівня захворюваності).



Сила зв'язку

слабка	середня	сильна
$r = 0,01 - 0,29$	$r = 0,30 - 0,69$	$r = 0,70 - 0,99$

За значенням коефіцієнта кореляції можна зробити такі висновки:

- якщо r набуває значення близьке до -1 , то між факторами існує щільний зворотній зв'язок.
- якщо $r = 0$, то зв'язок відсутній
- якщо r близьке до $+1$, то між факторами існує щільний прямий зв'язок
- якщо $r = 1$ або $= -1$, то між факторами існує функціональний зв'язок.



Достовірність зв'язку

- **достовірний зв'язок** поміж явищами чи ознаками при $p < 0,05$. Тобто встановлена залежність є і в досліджуваній групі, і в генеральній сукупності.
- **недостовірний зв'язок** поміж явищами чи ознаками при $p > 0,05$. Тобто встановлена залежність є в досліджуваній групі, а в генеральній сукупності її немає.



Форма зв'язку

- **Прямолінійна** — рівномірна зміна одного параметра відповідає рівномірним змінам іншого параметра (при незначних коливаннях)
- **Криволінійна** — рівномірна зміна одного параметра відповідає нерівномірним змінам іншого параметра (нерівномірність має певну закономірність)

Методи обрахунку коефіцієнта кореляції



- ранговий коефіцієнт (Спірмена)
- метод прямолінійної кореляції (Пірсона)

Ранговий коефіцієнт (Спірмена) використовується



- при невеликому числі спостережень ($n \leq 30$)
- коли немає необхідності в точних розрахунках сили зв'язку,
- коли ознаки мають не тільки кількісні, але і напівкількісні чи якісні значення,
- коли ряди розподілу мають відкриті варіанти (наприклад <20 або >40)

Метод прямолінійної кореляції (Пірсона) використовується



- коли ознаки мають кількісні значення,
- коли є потреба в точних розрахунках сили зв'язку.

Регресія



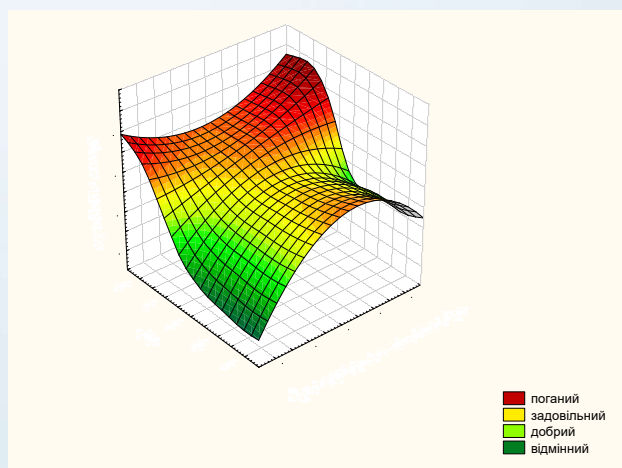
- функція, яка дозволяє за величиною одного явища (X) визначати очікувану величину іншого явища (Y).
- Коефіцієнт регресії показує, на яку величину в середньому зміниться другий параметр при зміні першого на певну одиницю виміру

Множинний коефіцієнт регресії



- відображає зв'язок одночасно комплексу факторів з досліджуваною результативною ознакою

Взаємозв'язок стану здоров'я з віком та матеріальним становищем пацієнтів сімейних лікарів м. Львова



Дякую за увагу

ПРОШУ ВАШИХ ЗАПИТАНЬ
ЗА ТЕМАТИКОЮ ПРОЧИТАНОЇ
ВАМ ЛЕКЦІЇ

