Методи математичної статистики в медсестринстві ЛЕКЦІЯ 6.

Кореляційний аналіз

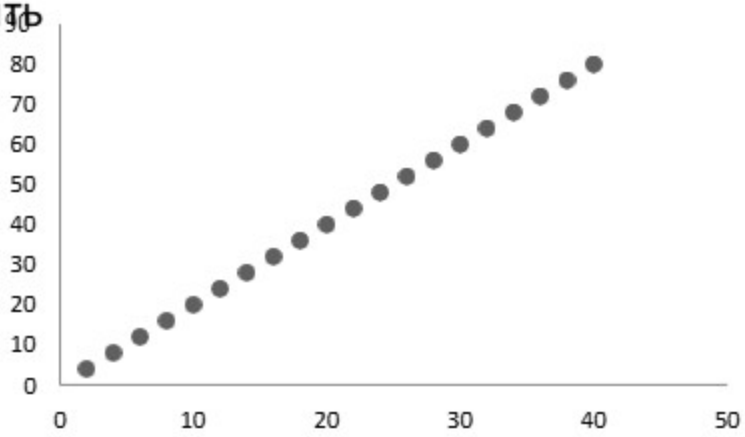
Кореляція – взаємозв'язок двох величин (змінних), при якому зміна однієї з них приводить до зміни іншої.

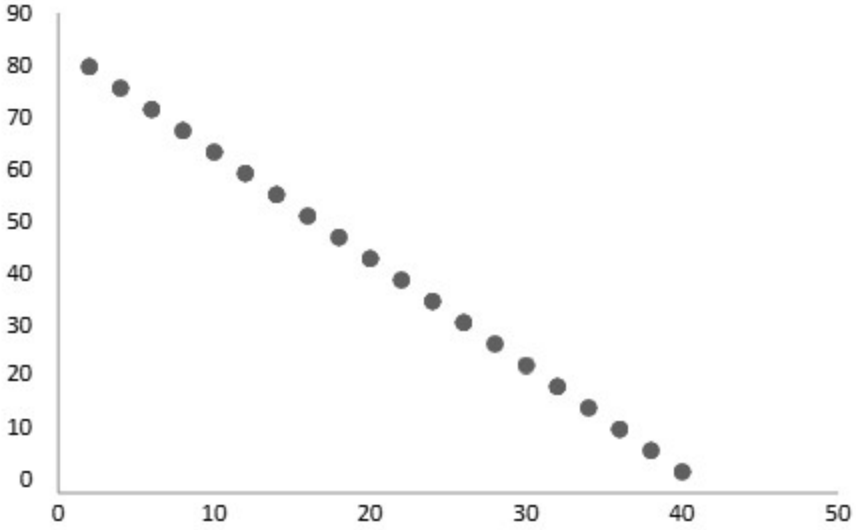
Кореляційний взаємозв'язок

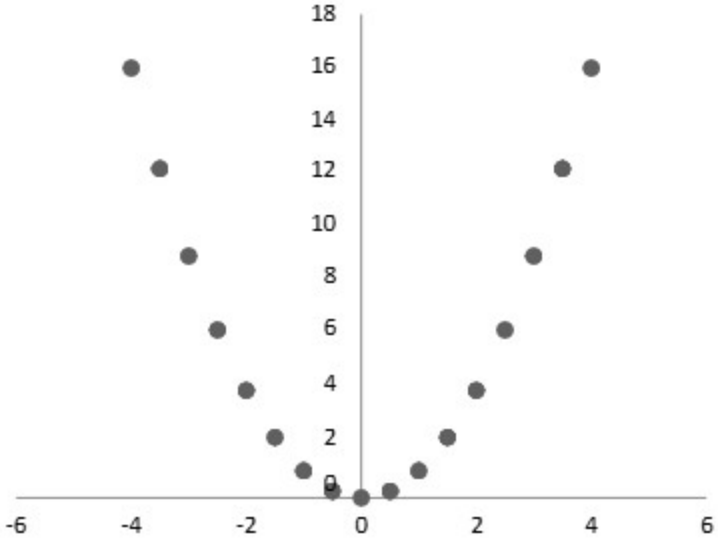
≠

Кореляційна залежність

ФУНКЦІОНАЛЬНИЙ ЗВ'ЯЗОК







КОЕФІЦІЄНТ КОРЕЛЯЦІЇ ПІРСОНА КОЕФІЦІЄНТ КОРЕЛЯЦІЇ СПІРМЕНА

1. Дані виражені у метричних шкалах
2. Розподіли змінних близькі до нормальних
3. Число ознак, що варіюються в змінних однакове
4. Кількість спостережень не менше 5.

Коефіцієнт кореляції Пірсона («r») — в статистиці, показник кореляції (лінійної залежності) між двома змінними X та Y, який набуває значень від −1 до +1 включно. Він широко використовується в науці для вимірювання ступеня лінійної залежності між двома змінними.

$$r = \frac{\overline{xy} - \overline{x} \cdot \overline{y}}{\sqrt{(\overline{x^2} - \overline{x}^2) \cdot (\overline{y^2} - \overline{y}^2)}}$$

1. Дані виражені у неметричних шкалах (але можуть бути виражені і в метричних).
2. Характер розподілу змінних не має значення.
3. Число ознак, що варіюються в змінних однакове.
4. Кількість спостережень не менше 5.

Коефіцієнт кореляції Спірмена («r») — оцінює наскільки добре можна описати відношення між двома змінними. Якщо немає повторних значень даних, то коефіцієнт Спірмена дорівнює 1 або −1. Коефіцієнт кореляції, як і будь-яке обчислення кореляції, підходить для безперервних та дискретних змінних, у тому числі порядкових.

$$R_S = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{N(N^2 - 1)}$$

За напрямком кореляційний зв'язок може бути:
– прямий;
– зворотній.

Кореляція	Негативна	Позитивна
Відсутня	−0.09 до 0.0	0.0 до 0.09
Низька	−0.3 до −0.1	0.1 до 0.3
Середня	−0.5 до −0.3	0.3 до 0.5
Висока	−1.0 до −0.5	0.5 до 1.0

Приклад 1.

Задача: досліджується зв'язок між поглинутою дозою випромінювання (X, Гр) та кількістю аберантних клітин головного мозку (Y, %) у білих мишей n=10. Маємо таку статистику:

n	X, Гр	Y, %	X·Y	X^2	Y^2
1	3,2	59	3,2·59	$(3,2)^2$	$(59)^2$
2	2,5	44	2,5·44	$(2,5)^2$	$(44)^2$
3	4,5	85	382,5	20,25	7225
4	4,0	70	280	16	4900
5	3,0	52	156	9	2704
6	0,8	21	16,8	0,64	441
7	1,3	28	36,4	1,69	784
8	4,0	79	316	16	6241
9	3,1	41	127,1	9,61	1681
10	3,5	67	234,5	12,25	4489
Σ	29,9	546	1848,1	101,93	33882

$$r = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sqrt{(\overline{x^2} - \bar{x}^2) \cdot (\overline{y^2} - \bar{y}^2)}}$$

Приклад 2.

Задача: Після проведення на одному із підприємств гірничо-видобувної промисловості цільового медогляду, було встановлено, що при стажі роботи до 5 років рівень поширеності хронічного бронхіту склав 20,1‰; при стажі роботи 5–9 років – 38,6‰; 10–14 років – 54,0‰; 15–19 років – 52,3‰ і при стажі більше 20 років – 78,0‰. Необхідно визначити, чи існує залежність між стажем роботи на підприємствах гірничовидобувної промисловості і ймовірністю розвитку хронічного бронхіту.

Стаж роботи (роки)	Поширеність хр. бронхіту (на 1000 наявного населення)	Порядковий номер ознаки (рангів)		Різниця рангів d (x–y)	Квадрат різниці рангів
x	y	x	y		
до 5 р.	20,1	1	1	0	0
5–9 р.	38,6	2	2	0	0
10–14 р.	54,0	3	4	-1	1
15–19 р.	52,3	4	3	1	1
20 р. та більше	78,0	5	5	0	0
					$\sum d^2 = 2$

$$R_S = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{N(N^2 - 1)}$$

Завдання:

1)Короткий конспект основних понять

Задача 1. Статистичний відділ лікарні досліджує залежність між кількістю хворих та кількістю лікарів стаціонарних відділень. Для такої оцінки вони мають досвід роботи у 5 відділеннях з майже однаковими умовами. Визначити коефіцієнт кореляції Пірсона за наступною інформацією:

Кількість лікарів, y_i	25+N	30+0,1N	35	45	65
Кількість хворих, x_i	15	16	19+2N	12	28

Задача 2. Визначити коефіцієнт кореляції рангів Спірмена щодо визначення впливу кількості пацієнтів на відсоток доплат до заробітної плати сімейних лікарів. Статистичні дані для розрахунків наведено у таблиці:

Кількість хворих	94+N	82	70-0,2N	69	52	44+0,1N
Ставка %	10	16	22	25	30+N	34