Лекція 09.

Оцінка якості моделі

Оцінка якості показує, наскільки теоретичні обчислення по побудованій моделі відхиляються від експериментальних даних. Наявність зв'язку двох змінних називається кореляцією.

Якщо оцінка якості застосовується до дослідження, то вона вирішує задачу: є чи зв'язок між входом X і виходом Y й оцінює силу цього зв'язку.

1. Лінійний коефіцієнт кореляції

Лінійний коефіцієнт кореляції вказує, є чи між двома рядами X й Y лінійна залежність й яка сила. Обчислюється по наступній формулі:

$$\mathrm{KR} = \frac{m_{xy} - m_x m_y}{\sigma_x \sigma_y}$$

m_x , m_y , m_{xy} — математичне очікування x, y, xy:

$$m_x = \frac{1}{n}\sum_{i=1}^{n} x_i \quad m_y = \frac{1}{n}\sum_{i=1}^{n} y_i \quad m_{xy} = \frac{1}{n}\sum_{i=1}^{n} x_i y_i$$

Дисперсія σ_x^2 й σ_y^2 показує, наскільки розкидані крапки від середньої величини:

$$\sigma_x^2 = \sum_{i=1}^{n} \frac{(x_i - m_x)^2}{n} \quad \sigma_y^2 = \sum_{i=1}^{n} \frac{(y_i - m_y)^2}{n}$$

Лінійний коефіцієнт кореляції може мати знак плюс або мінус. Позитивна його величина свідчить про прямий зв'язок між X й Y. Чим ближче KR до +1, тих зв'язок більше тісна. Негативна величина його свідчить про зворотний зв'язок; у цьому випадку границею є –1. Близькість KR до нуля свідчить про слабкий зв'язок між X й Y (див. мал. 9.1).


Y
X
Сильная положительная корреляция (KR стремится к 1)
Y
X
Корреляция отсутствует (KR близко к 0)
Y
X
Сильная отрицательная корреляция (KR стремится к –1)

Рис. 9.1.

2. Нелінійний коефіцієнт кореляції

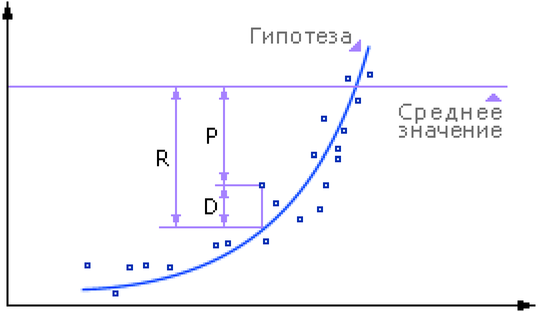

Рис. 9.2.

Нелінійний коефіцієнт кореляції обчислюється по наступній формулі:

$$\mathrm{KNR} = \sqrt{P - \frac{D}{P}}$$

P — розкид між реальними точками й середньою величиною:

$$P = \sum_{i=1}^{n} \frac{(y_i^{\mathrm{P.}} - y_i^{\mathrm{C.}})^2}{n}$$

D — розкид між гіпотетичною кривою й реальними точками:

?? $$R = \sum_{i=1}^{n} \frac{(y_i^{\text{Г.}} - y_i^{\text{С.}})^2}{n}$$

R — розкид між гіпотезою й середньою величиною:

?? $$R = \sum_{i=1}^{n} \frac{(y_i^{\text{Г.}} - y_i^{\text{С.}})^2}{n}$$

3. Коефіцієнт кореляції двох динамічних рядів

X й Y представляються у вигляді рядів z_i й u_i для того, щоб виключити постійну складову: $z_i = x_i – m_x$
$u_i = y_i – m_y$

$$r = \frac{\sum_{i=1}^{n} z_i u_i}{\sqrt{\sum_{i=1}^{n} z_i^2 \cdot \sum_{i=1}^{n} u_i^2}}$$

При r –> 1 має місце тісна кореляція. При r –> 0 процеси взаємно ортогональны, кореляції ні, процеси не зв'язані один з одним.

4. Кореляція усередині динамічного ряду

Досліджується сила зв'язку між минулим і сьогоденням одного процесу. Для цього сигнал порівнюють із самим собою, зрушеним у часі, і обчислюють коефіцієнт кореляції двох динамічних рядів (див. п. 3).

5. Пошук періодичності ряду

Є чи періодичність у динамічному ряді, можна з'ясувати, проробивши пряме перетворення Фур'є й розглянувши спектр досліджуваного сигналу. Про це розповідається в лекції 07.

6. Залежність динаміки ряду Z від двох динамічних факторів X й Y

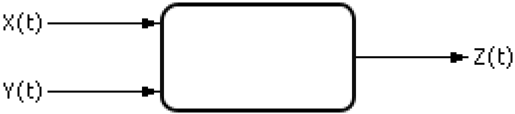

Рис. 9.5.

Коефіцієнт множинної кореляції R:

$$R = \sqrt{\frac{r_{xz}^2 + r_{xy}^2 - 2\, r_{xy} \cdot r_{xz} \cdot r_{yz}}{1 - r_{xy}^2}}$$

$$r_{xz} = \frac{m_{xz} - m_x m_z}{\sigma_x\, \sigma_z} \quad r_{xy} = \frac{m_{xy} - m_x m_y}{\sigma_x\, \sigma_y}$$

$$r_{yz} = \frac{m_{yz} - m_y m_z}{\sigma_y\, \sigma_z}$$

7. Зв'язок двох ознак

Формула

$$K = \frac{ad - bc}{ad + bc}$$

де K — це коефіцієнт асоціацій, дозволяє з'ясувати, чи є який-небудь зв'язок між двома ознаками. Якщо даний коефіцієнт близький до одиниці, то в цьому випадку можна говорити про існування такого зв'язку.

Приклад. Спробуємо за допомогою даної формули з'ясувати, є чи зв'язок між ростом і вагою людини? Нехай у нашому розпорядженні є дані про вагу й ріст 500 чоловік:

Таблиця 9.1.

	Вага < 67 кг.	**Вага > 67 кг.**
Ріст < 167 див.	a = 304 чіл.	b = 17 чіл.
Ріст > 167 див.	c = 112 чіл.	d = 67 чіл.

По формулі: K = (304 · 67 – 17 · 112)/(304 · 67 + 17 · 112) = 0.83. Так как величина 0.83 близька до 1, то можна говорити про існування певного зв’язку між вагою та ростом.