



Regresja

Wzory

Funkcja regresji I rodzaju

Funkcję regresji I rodzaju $g(x)$ wyznaczamy przyporządkowując wartościom cechy niezależnej (oznaczonej najczęściej jako X) średnie warunkowe cechy zależnej (najczęściej Y).

$$g(x_i) = E(Y | X = x_i)$$

Empiryczną linię regresji rysujemy na wykresie łącząc odpowiednie punkty odcinkami.

Funkcja regresji II rodzaju

Jeżeli Y zależy od X , oszacowanie **funkcji regresji II rodzaju** ma postać:

$$\hat{y}_i = a_0 + a_1 x_i,$$

gdzie:

$$a_1 = \frac{\sum (x_i - \bar{X})(y_i - \bar{Y})}{\sum (x_i - \bar{X})^2}$$

$$a_0 = \bar{Y} - a_1 \bar{X}$$

Jeżeli X zależy od Y , oszacowanie **funkcji regresji II rodzaju** ma postać:

$$\hat{x}_i = b_0 + b_1 y_i,$$

gdzie:

$$b_1 = \frac{\sum (x_i - \bar{X})(y_i - \bar{Y})}{\sum (y_i - \bar{Y})^2}$$

$$b_0 = \bar{X} - b_1 \bar{Y}$$

Uwagi

Trzeba pamiętać, że sumy we wzorach oznaczają sumowanie wszystkich jednostek w próbie i w niektórych zadaniach może być konieczne dodatkowe mnożenie przez liczebności n_i , o ile występują.

Współczynnik korelacji liniowej Pearsona: $r_{xy} = \pm \sqrt{a_1 \cdot b_1}$, znak przed pierwiastkiem jest zawsze taki sam, jak współczynników a_1 i b_1 .

Dodatkowo współczynniki a_1 i b_1 można wyznaczyć ze wzorów:

$$a_1 = r_{xy} \frac{S(Y)}{S(X)} \quad b_1 = r_{xy} \frac{S(X)}{S(Y)}$$

Badania dokładności dopasowania oszacowanej funkcji regresji

Wariancja resztowa

$$S^2(u_i) = \frac{\sum (y_i - \hat{y}_i)^2}{n - k}, \text{ jeśli } Y \text{ zależy od } X$$

$$S^2(z_i) = \frac{\sum (x_i - \hat{x}_i)^2}{n - k}, \text{ jeśli } X \text{ zależy od } Y$$

, gdzie k oznacza liczbę parametrów w funkcji regresji (w przypadku prostej $k = 2$)

Współczynnik zbieżności ϕ^2

$$\phi_{yx}^2 = \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{Y})^2}, \text{ jeśli } Y \text{ zależy od } X$$

$$\phi_{xy}^2 = \frac{\sum (x_i - \hat{x}_i)^2}{\sum (x_i - \bar{X})^2}, \text{ jeśli } X \text{ zależy od } Y$$

Współczynnik ten wskazuje, w ilu procentach zmiana wartości jednej zmiennej nie została wyjaśniona regresją.

Zachodzi:

$$\phi_{xy}^2 + r_{xy}^2 = 1$$

Korelacja wielu zmiennych

Korelacja cząstkowa

Badamy wzajemną korelację dwóch cech, z pominięciem wpływu pozostałych.

Współczynnik korelacji cząstkowej:

$$r_{ij.abzc...} = \frac{-P_{ij}}{\sqrt{P_{ii} \cdot P_{jj}}}, \text{ gdzie:}$$

$r_{ij.abzc...}$ - współczynnik korelacji cząstkowej pomiędzy cechami o indeksach i i j , z pominięciem cech o indeksach $a, b, z, c, \dots$

P_{ij}, P_{ii}, P_{jj} - odpowiednie dopełnienia algebraiczne macierzy współczynników korelacji

liniowej Pearsona:
$$P = \begin{bmatrix} 1 & r_{12} & r_{13} & \dots & r_{1n} \\ r_{21} & 1 & r_{23} & \dots & r_{2n} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ r_{n1} & r_{n2} & r_{n3} & \dots & 1 \end{bmatrix}.$$

Korelacja wieloraka

Badamy korelację wybranej cechy z wszystkimi pozostałymi łącznie

Współczynnik korelacji wielorakiej:

$$R_w = R_{i.abc...} = \sqrt{1 - \frac{|P|}{|R|}}, \text{ gdzie:}$$

$R_{i.abc...}$ - współczynnik korelacji wielorakiej cechy o indeksie i i wszystkich pozostałych cech

$|P|$ - wyznacznik macierzy współczynników korelacji liniowej Pearsona
$$P = \begin{bmatrix} 1 & r_{12} & r_{13} & \dots & r_{1n} \\ r_{21} & 1 & r_{23} & \dots & r_{2n} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ r_{n1} & r_{n2} & r_{n3} & \dots & 1 \end{bmatrix}$$

$|R|$ - wyznacznik macierzy współczynników korelacji liniowej Pearsona, z której wykreślono wiersz z numerem i oraz kolumnę z numerem i .

Regresja dla trzech zmiennych

Jeżeli Y zależy od dwóch zmiennych X_1 i X_2 , oszacowanie **funkcji regresji II rodzaju** ma postać:

$$\hat{y}_i = a_0 + a_1 x_{1i} + a_2 x_{2i}.$$

gdzie:

$$a_1 = \frac{S_1}{S_2} \cdot \frac{r_{12} - r_{13} \cdot r_{23}}{1 - r_{23}^2}$$

$$a_2 = \frac{S_1}{S_3} \cdot \frac{r_{13} - r_{12} \cdot r_{23}}{1 - r_{23}^2}$$

$$a_0 = \bar{Y} - a_1 \bar{X}_1 - a_2 \bar{X}_2$$

Indeks 1 mają współczynniki dotyczące zmiennej Y , indeks 2 współczynniki dotyczące zmiennej X_1 , a indeks 3 współczynniki dotyczące zmiennej X_2 .

Współczynnik korelacji wielorakiej:

$$R_w = R_{1.23} = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2r_{12}r_{13}r_{23}}{1 - r_{23}^2}}$$

Wariancja resztowa:

$$S^2(u) = S_1^2(1 - R_w^2)$$