

Zadanie dodatkowe nr 4

Tomasz Terlecki

Jakub Kuciński

23 maja 2020

1 Równanie dziewiąte

Rozważmy równanie regresji względem wielu zmiennych.

$$\begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \approx \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad (1)$$

Najlepsze rozwiązanie w sensie aproksymacji średniokwadratowej jest osiągnięte dla

$$\beta = (X^T X)^{-1} X^T Y. \quad (2)$$

2 Zastosowanie

Chcemy wyznaczyć prostą regresji zachorowań i zgonów względem piramidy wieku. W tym celu użyjemy danych dotyczących ludności 16 państw oraz liczby zachorowań i zgonów na dzień 12. maja. W naszym przypadku zmienne X_k odpowiadają liczbie osób w danych dziesięcioletnich przedziałach wiekowych. Wartości x_{ij} odpowiadają liczbie osób w przedziale wiekowym $[10(j-1), 10j-1]$ w kraju o numerze i . Inaczej mówiąc, kolumny odpowiadają grupom wiekowym, a wiersze ludności państw. Stosując wzór (2) dla wybranych przez nas krajów otrzymaliśmy:

Wiek	Współczynnik zachorowań	Wiek	Współczynnik zgonów
Const	-4260,602771	Const	-2397,93372
0-9	0,017547692	0-9	-0,00189355
10-19	0,059865071	10-19	0,006303402
20-29	-0,139278188	20-29	-0,00884239
30-39	-0,004668705	30-39	0,003931422
40-49	0,017537948	40-49	0,003929802
50-59	0,089679259	50-59	-0,00299812
60-69	-0,021341225	60-69	-0,01135733
70-79	0,10915987	70-79	0,021529316
80-89	-0,109840658	80-89	-0,00265009
90-99	-0,854519127	90-99	-0,04890375
100+	15,69150843	100+	1,843252609

3 Wnioski

Z obliczeń można wnioskować, że liczba osób w wieku przynajmniej 100 lat znacząco wpływa na liczbę zachorowań i zgonów. Wzrost liczby osób w tej grupie wiekowej istotnie zwiększa liczbę chorych i śmierci. Widzimy, że wzrost liczby osób w grupach wiekowych 0-9, 10-19, 40-49, 50-59, 70-79, 100+ zwiększa liczbę zachorowań, natomiast wzrost liczby osób w grupach 20-29, 30-39, 60-69, 80-89, 90-99 zmniejsza liczbę zachorowań.

Podobnie dla liczby zgonów widzimy, że wzrost liczby osób w grupach 10-19, 30-39, 40-49, 70-79, 100+ zwiększa liczbę zachorowań, natomiast wzrost liczby osób w grupach 0-9, 20-29, 50-59, 60-69, 80-89, 90-99 zmniejsza liczbę zachorowań.

Współczynniki liczby zachorowań dla grup wiekowych 0-9, 30-39, 40-49 są bliższe zeru niż dla pozostałych grup, więc liczba osób w tych grupach znacznie mniej wpływa na liczbę osób chorych w populacji. Najbardziej znacząca jest liczba osób w grupach 20-29, 70-79, 80-89, 90-99, 100+.

Analogicznie liczba osób w grupach wiekowych 0-9, 50-59 i 80-89 najmniej wpływa na liczbę zgonów, a liczba osób w grupach 60-69, 70-79, 90-99 i 100+ wpływa najbardziej.

Widzimy też, że tendencja wzrostowa lub malejąca co do liczby zachorowań i zgonów jest utrzymywana dla grup wiekowych 60+ (tzn. znak współczynnika wzrostu zachorowań i zgonów jest taki sam).

4 Obliczenia

Poniżej znajduje się link udostępniający skoroszyt w excelu zawierający obliczenia.

https://uniwroc-my.sharepoint.com/:x:/g/personal/309694_uwr_edu_pl/EVaoCq8vCSF0oiqgpL9a7jYBhQn0gSwwPe=Io3o2U