

Prognoza na podstawie modelu liniowego. (jedna zmienna)

$\hat{Y} = b_0 + b_1 X$ oszacowany model ekonometryczny.

Prognoza punktowa.

Niech x_τ – przewidywana wartość cechy X w okresie prognozy.

Prognoza punktowa y_τ^* to przewidywana wartość cechy Y odpowiadająca wartości x_τ cechy X .

$$y_\tau^* = b_0 + b_1 x_\tau$$

Standardowy błąd prognozy

$$s_\tau = s_e \sqrt{1 + \frac{1}{n} + \frac{(x_\tau - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}} = s_e \sqrt{1 + \frac{\sum_{i=1}^n x_i^2 + n x_\tau^2 - 2 x_\tau \sum_{i=1}^n x_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2}}$$

Zatem należy traktować wartość prognozy jako $y_\tau^* \pm s_\tau$

Gdzie $s_e = \sqrt{s_e^2}$ **to odchylenie resztowe.**

Niech $e_i = y_i - \hat{y}_i$, gdzie $\hat{y}_i = b_0 + b_1 x_i$ wtedy

$$s_e^2 = \frac{\sum_{i=1}^n e_i^2}{n-2}$$

czyli

$$s_e^2 = \frac{\sum_{i=1}^n y_i^2 - b_0 \sum_{i=1}^n y_i - b_1 \sum_{i=1}^n x_i y_i}{n-2}$$

Jakość prognozy punktowej możemy ocenić względnym błędem prognozy punktowej

$$\delta_{punkt} = \frac{s_{\tau}}{y_{\tau}^*} \cdot 100\%$$

Prognoza przedziałowa.

Prognoza przedziałowa dla poziomu ufności $1 - \alpha$.

$$\left\langle y_{\tau}^* - u_{\alpha} s_{\tau}; y_{\tau}^* + u_{\alpha} s_{\tau} \right\rangle$$

gdzie u_{α} odczytujemy z tablicy rozkładu Studenta:

$$P(|T_{n-2}| > u_{\alpha}) = \alpha$$

Jakość prognozy przedziałowej możemy ocenić względnym błędem prognozy przedziałowej

$$\delta_{prz} = \frac{u_{\alpha} s_{\tau}}{y_{\tau}^*} \cdot 100\%$$

Prognoza na podstawie modelu liniowego. (wiele zmiennych)

Prognoza punktowa.

Niech

$$x_{\tau} = \begin{bmatrix} 1 & x_{\tau 1} & \cdots & x_{\tau k} \end{bmatrix}^T$$

będzie wektorem którego pierwszą składową jest 1, a pozostałe to przewidywane wartości kolejnych zmiennych X_i w okresie prognozy τ .

Prognoza punktowa y_{τ}^* to przewidywana wartość Y odpowiadająca wartościom $x_{\tau i}$ zmiennych X_i .

$$y_{\tau}^* = x_{\tau}^T b$$

Standardowy błąd prognozy

$$S_{\tau} = S_e \sqrt{1 + x_{\tau}^T (X^T X)^{-1} x_{\tau}}$$

Gdzie $s_e = \sqrt{s_e^2}$ **to odchylenie resztowe.**

Niech, $e = Y - \hat{Y}$, gdzie $\hat{Y} = Xb$ wtedy

$$S_e^2 = S^2 = \frac{e^T e}{n - (k+1)} = \frac{\sum_{i=1}^n e_i^2}{n - (k+1)} = \frac{1}{n - (k+1)} (Y^T Y - b^T X^T Y)$$

Uwaga: zawsze $S_{\tau} > S_e$.

Zatem należy traktować wartość prognozy jako

$$y_{\tau}^* \pm S_{\tau}$$

Jakość prognozy punktowej możemy ocenić względnym błędem prognozy punktowej

$$\delta_{punkt} = \frac{S_{\tau}}{y_{\tau}^*} \cdot 100\%$$

Prognoza przedziałowa.

Prognoza przedziałowa dla poziomu ufności $1 - \alpha$.

$$\left\langle y_{\tau}^* - u_{\alpha} S_{\tau}; y_{\tau}^* + u_{\alpha} S_{\tau} \right\rangle$$

gdzie u_{α} odczytujemy z tablicy rozkładu Studenta:

$$P\left(\left|T_{n-(k+1)}\right| > u_{\alpha}\right) = \alpha.$$

Jakość prognozy przedziałowej możemy ocenić względnym błędem prognozy przedziałowej

$$\delta_{prz} = \frac{u_{\alpha} S_{\tau}}{y_{\tau}^*} \cdot 100\%$$

Model tendencji rozwojowej

Gdy X jest zmienną czasową $x_i = t$ ($t = 1, 2, \dots, n$) tzn. model regresji ma postać

$$\hat{Y} = b_0 + b_1 t$$

wówczas taki model nazywamy modelem tendencji rozwojowej lub modelem trendu liniowego

Wtedy korzystając z własności:

$$(*) \quad \sum_{t=1}^n t = \frac{n(n+1)}{2}, \quad \sum_{t=1}^n t^2 = \frac{n(n+1)(2n+1)}{6},$$
$$\bar{t} = \frac{n+1}{2} \quad \sum (t - \bar{t})^2 = \sum t^2 - n(\bar{t})^2$$

mamy

$$b_1 = \frac{n \sum t y_t - \sum t \sum y_t}{n \sum t^2 - (\sum t)^2} = \frac{12 \sum (t - \bar{t}) y_t}{n(n^2 - 1)}$$

$$b_0 = \bar{y} - b_1 \bar{t} = \bar{y} - b_1 \frac{n+1}{2}$$

Wariancja resztowa

Niech $e_i = y_i - \hat{y}_i$, (gdzie $\hat{y}_i = b_0 + b_1 x_i$) to reszty modelu, wtedy

$$s_e^2 = \frac{\sum_{i=1}^n e_i^2}{n-2} \quad \text{czyli} \quad s_e^2 = \frac{\sum_{t=1}^n y_t^2 - b_0 \sum_{t=1}^n y_t - b_1 \sum_{t=1}^n t y_t}{n-2}$$

$s_e = \sqrt{s_e^2}$ oznacza średnie (standardowe) odchylenie od trendu liniowego.

Dopasowanie modelu do danych empirycznych oceniamy też współczynnikiem determinacji

$$R^2 = \frac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2} = 1 - \frac{\sum e_i^2}{\sum(y_i - \bar{y})^2} =$$

$$= \frac{b_0 \sum y_i + b_1 \sum ty_t - n\bar{y}^2}{\sum y_t^2 - n(\bar{y})^2} = \frac{b_1(\sum ty_t - n\bar{t} \bar{y})}{\sum y_t^2 - n(\bar{y})^2} = r^2$$

Prognoza

Prognoza punktowa.

Niech t_τ – okres prognozy.

Prognoza punktowa y_τ to przewidywana wartość cechy Y w okresie t_τ .

$$y_\tau = b_0 + b_1 t_\tau$$

Standardowy błąd prognozy punktowej

$$s_\tau = s_e \sqrt{1 + \frac{1}{n} + \frac{(t_\tau - \bar{t})^2}{\sum_{t=1}^n (t - \bar{t})^2}} = s_e \sqrt{1 + \frac{\sum_{t=1}^n t^2 + nt_\tau^2 - 2t_\tau \sum_{t=1}^n t}{n \sum_{t=1}^n t^2 - \left(\sum_{t=1}^n t\right)^2}}$$

Wzór ten można uprościć korzystając z własności (*).

$$s_{\tau} = s_e \sqrt{1 + \frac{\frac{n(n+1)(2n+1)}{6} + nt_{\tau}^2 - 2t_{\tau} \frac{n(n+1)}{2}}{\frac{n^2(n+1)(2n+1)}{6} - \frac{n^2(n+1)^2}{4}}} =$$

$$= s_e \sqrt{1 + \frac{2(2n+1) + \frac{12t_{\tau}^2}{n+1} - 12t_{\tau}}{n^2 - n}}$$

Zatem należy traktować wartość prognozy jako $y_{\tau} \pm s_{\tau}$

Jakość prognozy punktowej możemy ocenić względnym błędem prognozy punktowej

$$\delta_{punkt} = \frac{s_{\tau}}{y_{\tau}} \cdot 100\%$$

Prognoza przedziałowa dla poziomu ufności $1 - \alpha$.

$$\left\langle y_{\tau} - u_{\alpha} S_{\tau}; y_{\tau} + u_{\alpha} S_{\tau} \right\rangle$$

gdzie u_{α} odczytujemy z tablicy rozkładu Studenta:

$$P(|T_{n-2}| > u_{\alpha}) = \alpha$$

Jakość prognozy przedziałowej możemy ocenić względnym błędem prognozy przedziałowej

$$\delta_{prz} = \frac{u_{\alpha} s_{\tau}}{y_{\tau}} \cdot 100\%$$