



MIARY OKREŚLAJĄCE STOPIEŃ ZGODNOŚCI MODELU Z DANYMI EMPIRYCZNYMI

Określane są na podstawie **reszt** modelu, czyli: $e_t = y_t - \hat{y}_t$,

gdzie: y_t - wartości rzeczywiste (zaobserwowane), $t = 1, 2, \dots, n$;

$\hat{y}_t$ - wartości teoretyczne wyznaczone na podstawie modelu, $t = 1, 2, \dots, n$.

1a) **Wariancja resztowa:**

$$S_e^2 = \frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{n - (k + 1)} = \frac{\sum_{t=1}^n e_t^2}{n - (k + 1)}$$
$$S_e^2 = \frac{\mathbf{e}^T \mathbf{e}}{n - (k + 1)} = \frac{\mathbf{Y}^T \mathbf{Y} - \mathbf{a}^T \mathbf{X}^T \mathbf{Y}}{n - (k + 1)}$$
$$S_e^2 = \frac{\mathbf{Y}^T \mathbf{Y} - \mathbf{Y}^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}}{n - (k + 1)}$$

gdzie: n - ilość obserwacji;

k - liczba zmiennych objaśniających w modelu ekonometrycznym;

e_t - reszty modelu w okresie t , czyli różnice pomiędzy wartościami rzeczywistymi

zmiennej objaśnianej y_t , a wartościami teoretycznymi wyznaczonymi z modelu $\hat{y}_t$;

$\mathbf{e}$ - wektor reszt;

$\mathbf{Y}$ - wektor zmiennych objaśnianych;

$\mathbf{X}$ - macierz zmiennych objaśniających;

$\mathbf{X}^T$ - macierz transponowana do macierzy $\mathbf{X}$;

$(\mathbf{X}^T \mathbf{X})^{-1}$ - macierz odwrotna do macierzy $\mathbf{X}^T \mathbf{X}$;

$\mathbf{a}$ - wektor ocen parametrów strukturalnych oszacowanych KMNK.

UWAGA: *Wariancji resztowej NIE interpretujemy!!!*



1b) **Odchylenie standardowe reszt** (odchylenie resztowe, odchylenie standardowe składnika losowego, standardowy błąd estymacji, średni błąd modelu, błąd standardowy reszt):

$$S_e = \sqrt{S_e^2}$$

Informuje o ile przeciętnie rzeczywiste (zaobserwowane) wartości zmiennej objaśnianej różnią się od jej wartości teoretycznych wyznaczonych na podstawie modelu ekonometrycznego.

INTERPRETACJA (ODP. przykładowa): Wartości rzeczywiste *..(zmienna Y)..* różnią się przeciętnie od ich/jej wartości teoretycznych wyznaczonych z modelu o około *..(wartość S_e)..* jednostki (zmiennej Y).

LUB:

Średni błąd modelu oznacza, że szacując model tej zmiennej objaśnianej mylimy się średnio o *..(wartość S_e)..* jednostki.

UWAGA: *Intepretuje się tylko wynik, NIE ocenia się modelu.*

2) **Współczynnik zmienności losowej** (współczynnik zmienności przypadkowej, współczynnik wyrazistości):

$$W_e = \frac{S_e}{\bar{y}} \cdot 100\% \quad (\text{oznaczenia równoważne: } W_e = V_e)$$

gdzie: S_e - odchylenie standardowe reszt modelu;

$\bar{y}$ - średnia arytmetyczna zmiennej objaśnianej Y.

Określa, jaki procent średniej wartości zmiennej objaśnianej stanowi odchylenie standardowe reszt model. Im mniejsza wartość współczynnika zmienności losowej, tym większa zgodność modelu z danymi empirycznymi.

Jeśli $W_e \leq W_e^*$, gdzie wartość krytyczna W_e^* jest równa np. 10%, to model oceniamy pozytywnie.

INTERPRETACJA (ODP.): Odchylenie standardowe reszt modelu stanowi około *..(wartość W_e procentowo)..* średniej wartości *..(zmienna Y)..*

Zakładając wartość krytyczną $W_e^* = \dots$ (np. $W_e^* = 10\%$) otrzymujemy, iż $W_e \leq W_e^*$, zatem model uznaje się za dobrze dopasowany, zgodny z danymi empirycznymi. Model oceniamy pozytywnie.

LUB

Zakładając wartość krytyczną $W_e^* = \dots$ (np. $W_e^* = 10\%$) otrzymujemy, iż $W_e > W_e^*$, zatem model uznaje się za źle dopasowany, nie jest zgodny z danymi empirycznymi. Model oceniamy negatywnie.



Analiza wariancji zmiennej objaśnianej:

Całkowitą zmienność zmiennej objaśnianej określoną jako: $\sum_{t=1}^n (y_t - \bar{y})^2$

$$\sum_{t=1}^n (y_t - \bar{y})^2 = \sum_{t=1}^n (y_t - \bar{y})^2 + \sum_{t=1}^n (y_t - y_t)^2$$

można rozbić na 2 części:

$$\sum_{t=1}^n (y_t - \bar{y})^2 = \sum_{t=1}^n (y_t - \bar{y})^2 + \sum_{t=1}^n e_t^2$$

gdzie: $\sum_{t=1}^n (y_t - \bar{y})^2 = SST$ - całkowita zmienność zmiennej objaśnianej; suma kwadratów

odchyłeń wartości rzeczywistych od średniej zmiennej objaśnianej;

$$\sum_{t=1}^n (y_t - \bar{y})^2 = SSR$$
 - zmienność zmiennej objaśnianej wyjaśniona przez model;

suma kwadratów odchyłeń wartości teoretycznych od średniej zmiennej objaśnianej;

$$\sum_{t=1}^n (y_t - y_t)^2 = \sum_{t=1}^n e_t^2 = SSE$$
 - zmienność zmiennej objaśnianej nie wyjaśniona przez

model; suma kwadratów odchyłeń wartości rzeczywistych od wartości teoretycznych zmiennej objaśnianej, czyli suma kwadratów reszt.

Zachodzi więc: $SST = SSR + SSE$ lub $1 = \frac{SSR}{SST} + \frac{SSE}{SST}$

3) Współczynnik determinacji:

$$R^2 = \frac{SSR}{SST} = 1 - \frac{\sum_{t=1}^n (y_t - y_t)^2}{\sum_{t=1}^n (y_t - \bar{y})^2}$$

$$R^2 = 1 - \frac{\mathbf{e}^T \mathbf{e}}{\mathbf{Y}^T \mathbf{Y} - n \cdot \bar{y}^2} = 1 - \frac{\mathbf{Y}^T \mathbf{Y} - \mathbf{a}^T \mathbf{X}^T \mathbf{Y}}{\mathbf{Y}^T \mathbf{Y} - n \cdot \bar{y}^2}$$

$$R^2 = \mathbf{R}_0^T \mathbf{R}^{-1} \mathbf{R}_0$$

gdzie: $\mathbf{R}_0$ - macierz współczynników korelacji r_j pomiędzy zmienną objaśnianą $\mathbf{Y}$ a zmiennymi objaśniającymi $\mathbf{X}_j$;



$\mathbf{R}^{-1}$ - macierz odwrotna do macierzy $\mathbf{R}$, czyli macierzy współczynników korelacji r_{ij} pomiędzy zmiennymi objaśniającymi X_i oraz X_j .

Wartość współczynnika R^2 należy do przedziału $[0,1]$.

Aby wyrazić wartość współczynnika procentowo, należy otrzymany wynik (ułamek) $\cdot 100\%$.

Informuje, w jakim stopniu zmienność zmiennej objaśnianej została wyjaśniona przez model.

Jeśli $R^2 \geq 0,7$ (70%) to model uznaje się za dobrze dopasowany do danych. Model oceniamy pozytywnie.

W przeciwnym wypadku, gdy $R^2 < 0,7$ (70%), model uznaje się za źle dopasowany do danych. Model weryfikujemy negatywnie.

INTERPRETACJA (ODP.): Zmienność zmiennej objaśnianej została wyjaśniona przez model w ..(wartość R^2 procentowo)... Model uznaje się za dobrze /lub/ źle dopasowany do danych. Model weryfikujemy pozytywnie /lub/ negatywnie.

4) Współczynnik zbieżności:

$$\varphi^2 = 1 - R^2 \quad \text{gdyż: } \varphi^2 + R^2 = 1$$

$$\varphi^2 = \frac{SSE}{SST} = \frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{\sum_{t=1}^n (y_t - \bar{y})^2}$$

$$\varphi^2 = \frac{\mathbf{e}^T \mathbf{e}}{\mathbf{Y}^T \mathbf{Y} - n \cdot \bar{y}^2} = \frac{\mathbf{Y}^T \mathbf{Y} - \mathbf{a}^T \mathbf{X}^T \mathbf{Y}}{\mathbf{Y}^T \mathbf{Y} - n \cdot \bar{y}^2}$$

Wartość współczynnika φ^2 należy do przedziału $[0,1]$.

Informuje, w jakim stopniu zmienność zmiennej objaśnianej NIE została wyjaśniona przez model.

Jeśli $\varphi^2 \leq 0,3$ (30%) to model uznaje się za dobrze dopasowany do danych. Model oceniamy pozytywnie.

W przeciwnym wypadku, gdy $\varphi^2 > 0,3$ (30%), model uznaje się za źle dopasowany do danych. Model weryfikujemy negatywnie.

INTERPRETACJA (ODP.): Zmienność zmiennej objaśnianej nie została wyjaśniona przez model w ..(wartość φ^2 procentowo)... Model uznaje się za dobrze /lub/ źle dopasowany do danych. Model weryfikujemy pozytywnie /lub/ negatywnie.

5) Współczynnik korelacji wielorakiej:

$$R = \sqrt{R^2}$$

Wartość współczynnika R należy do przedziału $[0,1]$.

Współczynnik korelacji wielorakiej można potraktować jako współczynnik korelacji między wartościami rzeczywistymi zmiennej objaśnianej a jej wartościami teoretycznymi wyznaczonymi z modelu ekonometrycznego.

Informuje, w jakim stopniu są skorelowane ze sobą empiryczne (Y) i teoretyczne ($\hat{Y}$) wartości zmiennej objaśnianej.

Weryfikacja hipotezy o istotności współczynnika korelacji wielorakiej R :

I. Formujemy hipotezy:

$H_0 : R = 0$ (współczynnik korelacji wielorakiej nie jest istotny)

$H_1 : R \neq 0$ (współczynnik korelacji wielorakiej jest istotny)

II. Obliczamy statystykę: $F = \frac{R^2}{1-R^2} \cdot \frac{n-(k+1)}{k}$

III. Odczytujemy z tablic rozkładu Fishera-Snedecora wartość krytyczną $F^* = F_{\alpha, r_1, r_2}$ dla zadanego poziomu istotności α oraz $r_1 = k$ i $r_2 = n - (k + 1)$

IV. Porównujemy statystykę F i F^* i piszemy wniosek:

- Jeśli $F > F^*$ to należy odrzucić hipotezę zerową H_0 na korzyść hipotezy alternatywnej H_1 . Oznacza to, że współczynnik korelacji wielorakiej jest istotny, a dopasowanie modelu ekonometrycznego do danych jest dostatecznie wysokie.

- Jeśli $F \leq F^*$ to nie ma podstaw by odrzucić hipotezę zerową H_0 . Oznacza to, że współczynnik korelacji wielorakiej jest nieistotnie różny od zera, a stopień dopasowania modelu ekonometrycznego do danych jest zbyt słaby.

