ЗАХОДИ, ЩО ВИЗНАЧУЮТЬ СТУПЕНЬ ВІДПОВІДНОСТІ МОДЕЛІ З ЕМПІРИЧНИМИ ДАНИМИ

Вони визначаються на основі модельних **залишків**, тобто.: $e_t = y_t - y_t$,

де: y_t - фактичні значення (дотримані), $t = 1, 2, \ldots, n$;

y_t - теоретичні значення, визначені на основі моделі, $t = 1, 2, \ldots, n$.

1a) **Залишкова дисперсія:**

$$S_e^2 = \frac{\sum_{t=1}^{n}\left(y_t - y_t\right)^2}{n-(k+1)} = \frac{\sum_{t=1}^{n} e_t^2}{n-(k+1)}$$

$$S_e^2 = \frac{\mathbf{e}^T\mathbf{e}}{n-(k+1)} = \frac{\mathbf{Y}^T\mathbf{Y} - \mathbf{a}^T\mathbf{X}^T\mathbf{Y}}{n-(k+1)}$$

$$S_e^2 = \frac{\mathbf{Y}^T\mathbf{Y} - \mathbf{Y}^T\mathbf{X}\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{Y}}{n-(k+1)}$$

де: n - кількість спостережень;

k - кількість незалежних змінних в економетричній моделі;

e_t - решта моделі в період t, тобто відмінності між реальними значеннями залежної змінної y_t і теоретичними значеннями, визначеними з моделі y_t;

$\mathbf{e}$ - вектор залишків;

$\mathbf{Y}$ - вектор пояснюваних змінних;

$\mathbf{X}$ - матриця пояснювальних змінних;

$\mathbf{X}^T$ - матриця, транспонована в матрицю X ;

$\left(\mathbf{X}^T\mathbf{X}\right)^{-1}$ - обернена матриця $X^T X$;

$\mathbf{a}$ - вектор оцінок розрахункових параметрів конструкції KMNK.

<u>ПРИМІТКА.</u> *Ми НЕ інтерпретуємо залишкову дисперсію!!!*

Joanna Grochowska eTrapez
www.etrapez.pl
tel. 534 864 486

1b) **Стандартне відхилення залишків** (залишкове відхилення, стандартне відхилення випадкового члена, стандартна помилка оцінки, середня помилка моделі, стандартна помилка залишків):

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

Він повідомляє, наскільки в середньому реальні (спостережувані) значення залежної змінної відрізняються від її теоретичних значень, визначених на основі економетричної моделі.

ІНТЕРПРЕТАЦІЯ (одп. Зразковий): реальні значення ..(змінна *Y*).. відрізняються в середньому від своїх теоретичних значень, визначених з моделі, приблизно на ..(значення S_e).. одиниць (змінна *Y*).

АБО:

Середня помилка моделі означає, що при оцінці моделі цієї поясненої змінної ми помиляємося в середньому на ..(значення S_e).. одиниць.

ПРИМІТКА*: Інтерпретується лише результат, модель НЕ оцінюється.*

2) **Коефіцієнт випадкової варіації** (коефіцієнт випадкової варіації, коефіцієнт наочності)

$$W_e = \frac{S_e}{\overline{y}} \cdot 100\%$$ (*еквівалентні маркування*: $W_e = V_e$)

де: S_e - стандартне відхилення залишків моделі;

$\overline{y}$ - середнє арифметичне залежної змінної *Y*.

Визначає, який відсоток середнього значення поясненої змінної є стандартним відхиленням залишків моделі. Чим менше значення коефіцієнта випадкової варіації, тим більша відповідність моделі емпіричним даним.

Якщо $W_e \le W_e^*$, де критичне значення W_e^* дорівнює, наприклад 10%, ми оцінюємо модель позитивно.

ІНТЕРПРЕТАЦІЯ (РЕП.): Стандартне відхилення залишків моделі становить приблизно ..(відсоткове значення W_e).. середнього значення ..(змінна *Y*)...

Прийнявши критичне значення *We* = ...* (наприклад, *We* = 10%*), отримуємо, що $W_e \le W_e^*$, отже, модель вважається добре підібраною моделлю, що узгоджується з емпіричними даними. Ми оцінюємо модель позитивно.

АБО

Joanna Grochowska eTrapez
www.etrapez.pl
tel. 534 864 486

Прийнявши критичне значення *We* =...* (наприклад, *We* = 10%*) отримуємо, що $W_e > W_e^{*}$, отже, модель вважається поганою, вона не відповідає емпіричним даним. Ми оцінюємо модель негативно.

Аналіз дисперсії залежної змінної:

Загальна мінливість залежної змінної визначається як: $\sum_{t=1}^{n}\left(y_t - \overline{y}\right)^2$

можна розбити на 2 частини:

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

$$\sum_{t=1}^{n}\left(y_t - \overline{y}\right)^2 = \sum_{t=1}^{n}\left(y_t - \overline{y}\right)^2 + \sum_{t=1}^{n} e_t^{\,2}$$

де: $\sum_{t=1}^{n}\left(y_t - \overline{y}\right)^2 = SST$ - сумарна мінливість залежної змінної; сума квадратів відхилень реальних значень від середнього значення залежної змінної;

$\sum_{t=1}^{n}\left(y_t - \overline{y}\right)^2 = SSR$ - мінливість залежної змінної, що пояснюється моделлю; сума квадратів відхилень теоретичних значень від середнього значення залежної змінної;

$\sum_{t=1}^{n}\left(y_t - y_t\right)^2 = \sum_{t=1}^{n} e_t^{\,2} = SSE$ - мінливість поясненої змінної, не поясненої моделлю; сума квадратів відхилень реальних значень від теоретичних значень залежної змінної, тобто сума квадратів залишків.

Так воно і є: $SST = SSR + SSE$ або $1 = \frac{SSR}{SST} + \frac{SSE}{SST}$

3) **Коефіцієнт детермінації:**

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

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

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

де: $\mathbf{R}_0$ - матриця коефіцієнтів кореляції r_j між залежною змінною Y та пояснювальною змінною X_j;

$\mathbf{R}^{-1}$ - матриця, обернена до матриці $\mathbf{R}$, тобто матриця коефіцієнтів кореляції r_{ij} між пояснювальними змінними X_i та X_j.

Значення коефіцієнта R^2 знаходиться в межах [0,1].
Щоб виразити значення фактора у відсотках, слід отримати отриманий результат (дріб) • 100%.
Він повідомляє, якою мірою мінливість залежної змінної була пояснена моделлю.
Якщо $R^2 \geq 0,7 \ (70\%)$ модель вважається хорошою відповідною для даних. Ми оцінюємо модель позитивно.
Інакше, коли модель $R^2 < 0,7 \ (70\%)$ вважається поганою відповідністю даним. Ми верифікуємо модель негативно.

ІНТЕРПРЕТАЦІЯ (одп.): мінливість залежної змінної була пояснена моделлю w ..(відсоток R_2).. . Модель вважається *добре/або/погано* пристосованою до даних. Ми перевіряємо модель *позитивно/або/негативно*.

Joanna Grochowska eTrapez
www.etrapez.pl
tel. 534 864 486

4) **Коефіцієнт збіжності:**

$$\varphi^2 = 1 - R^2 \qquad gdy\dot{z}: \quad \varphi^2 + R^2 = 1$$

$$\varphi^2 = \frac{SSE}{SST} = \frac{\sum_{t=1}^{n}\left(y_t - y_t\right)^2}{\sum_{t=1}^{n}\left(y_t - \overline{y}\right)^2}$$

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

Значення коефіцієнта φ^2 знаходиться в межах [0,1].
Він повідомляє, якою мірою мінливість залежної змінної НЕ була пояснена моделлю.
Якщо $\varphi^2 \leq 0,3 \ (30\%)$ модель вважається хорошою відповідною для даних. Ми оцінюємо модель позитивно.
Інакше, коли модель вважається поганою відповідністю даним $\varphi^2 > 0,3 \ (30\%)$. Ми верифікуємо модель негативно.

ІНТЕРПРЕТАЦІЯ (одп.): Змінність залежної змінної не була пояснена моделлю в ..(відсоткове значення φ^2).. . Модель вважається такою, що *добре/або/погано* відповідає даним. Ми перевіряємо модель *позитивно/або/негативно*.

5) **Коефіцієнт множинної кореляції:**

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

Значення коефіцієнта R знаходиться в межах [0,1].
Коефіцієнт множинної кореляції можна розглядати як коефіцієнт кореляції між фактичними значеннями поясненої змінної та її теоретичними значеннями, визначеними з економетричної моделі.
Він повідомляє, якою мірою емпіричні (Y) та теоретичні $(\hat{Y})$ значення пояснюваної змінної корелюють один з одним.

Перевірка гіпотези про значущість коефіцієнта множинної кореляції R:

I. Формулюємо гіпотези:

   $H_0 : R = 0$ (коефіцієнт множинної кореляції не є значущим)

   $H_1 : R \neq 0$ (значний коефіцієнт множинної кореляції)

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

III. Ми зчитуємо критичне значення з таблиць розподілу Фішера-Снедекора

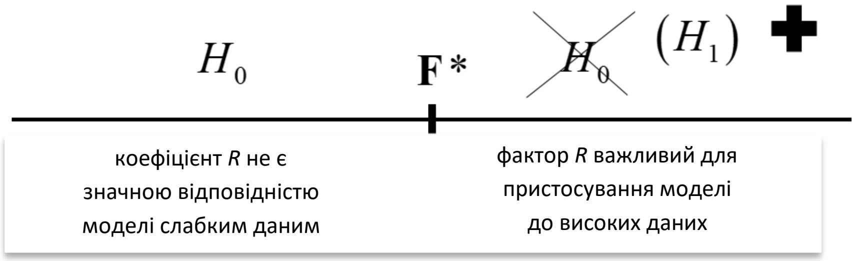
$$F^* = F_{\alpha, r_1, r_2}$$

для даного рівня значущості α та $r_1 = k$ і $r_2 = n-(k+1)$

IV. Порівнюємо статистику F і F^* та робимо висновок:

- Якщо необхідно відхилити нульову гіпотезу H_0 на користь альтернативної гіпотези H_1. Це означає, що коефіцієнт множинної кореляції є значним, і відповідність економетричної моделі даним є достатньо високою.

- Якщо $F \leq F^*$ немає підстав відхиляти нульову гіпотезу H_0. Це означає, що коефіцієнт множинної кореляції суттєво не відрізняється від нуля, а ступінь відповідності економетричної моделі даним занадто слабка.



Joanna Grochowska eTrapez
www.etrapez.pl
tel. 534 864 486