Ekonometria

Robert Pietrzykowski

Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

WNE Zarządzanie Studia stacjonarne 2014 - 2015

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + \epsilon_i$$

$$y = X\beta + \epsilon$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{k1} \\ 1 & x_{12} & x_{22} & \cdots & x_{k2} \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{kn} \end{bmatrix} x \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

- Zakładamy, że każda obserwacja zmiennej objaśnianej y_i powstaje jako suma iloczynów nieznanych parametrów $\beta_i, \quad i = 0, 1, 2, \ldots, k$ i zmiennych objaśniających x_{ki} oraz uzupełniana jest przez efekt losowy ϵ_i.
- Zakładamy, że obserwacje dotyczące zmiennych objaśnianych są stałe (inne określenie zmiennej x zmienna niezależna) to znaczy niezależne od efektów losowych ϵ ($\beta_k \cdot x_{ki}$ - wyjaśnia część zmiennej y, ϵ - dotyczy części niewyjaśnionej: patrz test F i równanie wariacyjne)
- **A zatem zmienna objaśniana jest losowa, gdyż jest funkcją efektów losowych** ϵ
- **Poza tym zmienna objaśniana jest funkcją liniową zmiennych objaśnianych i względem parametrów równania regresji**
- Poza tym zakładamy, że liniowy model regresji został **dobrze wyspecyfikowany** tzn. zawiera wszystkie ważne zmienne do wyjaśnienia zmienności zmiennej objaśnianej oraz ma poprawną postać matematyczną

- Elementy macierzy $\mathbf{X}$ są nielosowe.

  Rozważamy zależność deterministyczną tzn. związek przyczynowo–skutkowy. W przypadku zależności stochastycznej (losowość zmiennych objaśnianych) są one niezależne od efektów losowych ϵ. W ekonomii pierwsze założenie zwykle nie jest spełnione. Dlatego dla modelu KMRL przyjmujemy drugie założenie jako wystarczające, a zatem:

$$E(y|\mathbf{X}) = E(y) \quad oraz \quad var(y|\mathbf{X}) = var(y)$$

- Rząd macierzy $\mathbf{X}$ jest równy liczbie szacowanych parametrów k, a k jest mniejsze od liczby obserwacji n co można zapisać:

$$rz(\mathbf{X}) = k < n$$

Z powyższego założenia wynika, że obserwacje na każdej zmiennej objaśniającej nie mogą być jednakowe, czyli kolumna macierzy $\mathbf{X}$ nie może być kombinacją liniową innych kolumn tej macierzy (każda zmienna wnosi do równania regresji inne informacje), a zatem:

$$rz(\mathbf{X}'\mathbf{X}) = k$$

Wektor błędów losowych ϵ ma warunkową wartość oczekiwaną przy konkretnej macierzy $\mathbf{X}$ róną wektorowi zerowemu, a zatem:

$$E(\epsilon|\mathbf{X}) = E(\epsilon) = 0$$

Z powyższego wynika, że efekty losowe nie wpływają na zmienne objaśniające. Poza tym efekty losowe dodatnie i ujemne wzajemnie się znoszą, a oczekiwany ich efekt jest równy zero czyli czynniki pominięte w modelu nie wpływają na średnią wartość zmiennej objaśnianej.

$$E(\epsilon) = E\begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix} = \begin{bmatrix} E(\epsilon_1) \\ E(\epsilon_2) \\ \vdots \\ E(\epsilon_n) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Wariancje efektów losowych są takie same dla wszystkich obserwacji i równe nieznanej wariancji σ^2. Po drugie kowariancje efektów losowych są nieskorelowane czyli równe zero, co można zapisać w postaci:

$$var(\epsilon|\mathbf{X}) = E(\epsilon\epsilon'|\mathbf{X}) = var(\epsilon) = E(\epsilon\epsilon') = \sigma^2\mathbf{I}$$

lub w sposób rozwinięty:

$$E\left(\begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix} x \begin{bmatrix} \epsilon_1 & \epsilon_2 & \cdots & \epsilon_n \end{bmatrix}\right) = \begin{bmatrix} E(\epsilon_1^2) & E(\epsilon_1\epsilon_2) & \cdots & E(\epsilon_1\epsilon_n) \\ E(\epsilon_2\epsilon_1) & E(\epsilon_2^2) & \cdots & E(\epsilon_2\epsilon_n) \\ \vdots & \vdots & \dots & \vdots \\ E(\epsilon_n\epsilon_1) & E(\epsilon_n\epsilon_2) & \cdots & E(\epsilon_n^2) \end{bmatrix}$$

wiedząc, że zgodnie z definicją:

$$D^2(\epsilon) = E(\epsilon_i - E(\epsilon_i))^2$$

oraz kowariancja:

$$E((\epsilon_i - E(\epsilon_i))(\epsilon_j - E(\epsilon_j)))$$

przy założeniu $E(\epsilon_i) = 0$
otrzymujemy:

$$\begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix} = \sigma^2 \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

- Jednakowe wariancje efektów losowych nazywamy **homoskedastycznością**
- Przypadek zerowych kowariancji dla różnych zaburzeń losowych ϵ_i i ϵ_j nazywamy **brakiem autokorelacji zaburzeń**. Oznacza to, że zaburzenia losowe dla różnych obserwacji są niezależne, a przez to nieskorelowane, a wiec nie maja tendencji do gromadzenia się np. wokół dodatnich lub ujemnych (lub naprzemiennie dodatnich i ujemnych wartości).
- Warunkowe rozkłady efektów losowych mają identyczne niezależne rozkłady czyli $\epsilon_i \backsim IID(0, \sigma^2)$ (Independent and Identically Distributed – niezależne i identycznie rozłożone)

- Zaburzenia losowe mają n-wymiarowy rozkład normalny:

$$\epsilon \backsim N(0, \sigma^2 \mathbf{I})$$

- Z wcześniejszych założeń wynika, że każdy efekt losowy ma rozkład normalny o średniej 0 i wariancji σ^2 i rozkład ten jest niezależny od rozkładu innego efektu losowego.