Ekonometryczne modele nieliniowe

Wykład 2

Własności estymatorów i testy

1. dodatek do wykładu 1

- Słaba zbieżność (convergence in distribution)

  Ciąg zmiennych losowych $\{X_N\}_{N=1}^{\infty}$

  $F_{X_N}(x)$ - dystrybuanta X_N

  Istnieje dystrybuanta $F_X(x)$, taka że

  $\lim_{N\to\infty} F_{X_N}(x) = F_X(x)$ w każdym punkcie x,

  w którym $F_X(\cdot)$ jest ciągła.

X_N zbiega słabo do X :

$$X_N \xrightarrow{L} X$$

MNK przy warunkach pobocznych

- Restricted LS

$$J = (\mathbf{y} - \mathbf{X}\mathbf{b}_*)'(\mathbf{y} - \mathbf{X}\mathbf{b}_*) - 2\boldsymbol{\lambda}'(\mathbf{R}\mathbf{b}_* - \mathbf{r}) \rightarrow \min$$

$$\mathbf{b}_* = \mathbf{b} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{r} - \mathbf{R}\mathbf{b})$$

$$\mathbf{e}_* = \mathbf{y} - \mathbf{X}\mathbf{b}_* = \mathbf{y} - \mathbf{X}\mathbf{b} - \mathbf{X}(\mathbf{b}_* - \mathbf{b}) = \mathbf{e} - \mathbf{X}(\mathbf{b}_* - \mathbf{b})$$

$$\mathbf{e}_*'\mathbf{e}_* = \mathbf{e}'\mathbf{e} + (\mathbf{b}_* - \mathbf{b})'\mathbf{X}'\mathbf{X}(\mathbf{b}_* - \mathbf{b})$$

Test F (inny zapis)

$$\mathbf{e}_*'\mathbf{e}_* = \mathbf{e}'\mathbf{e} + (\mathbf{r} - \mathbf{R}\mathbf{b})'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{r} - \mathbf{R}\mathbf{b})$$

Wykorzystując formułę z poprzedniego wykładu:

$$F = (\mathbf{R}\mathbf{b} - \mathbf{r})'(S^2\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}(\mathbf{R}\mathbf{b} - \mathbf{r})/m \sim F(m, N-k)$$

$$F = \frac{(\mathbf{e}_*'\mathbf{e}_* - \mathbf{e}'\mathbf{e})/m}{\mathbf{e}'\mathbf{e}/(N-k)} \sim F(m, N-k)$$

Metoda największej wiarygodności

- Maximum Likelihood:

$$Funkcja\ wiarygodności = L(\boldsymbol{\theta}; \mathbf{y}) = f(\mathbf{y}; \boldsymbol{\theta})$$

- Maksymalizujemy funkcję wiarygodności względem $\boldsymbol{\theta}$ → maksymalizujemy *prawdopodobieństwo otrzymania próby takich obserwacji*, które właśnie zaobserwowaliśmy

- Alternatywna interpretacja: *funkcja parametrów warunkowa na obserwacjach*

Estymator MNW

- Ze względów obliczeniowych stosujemy:

  $l = \ln L$

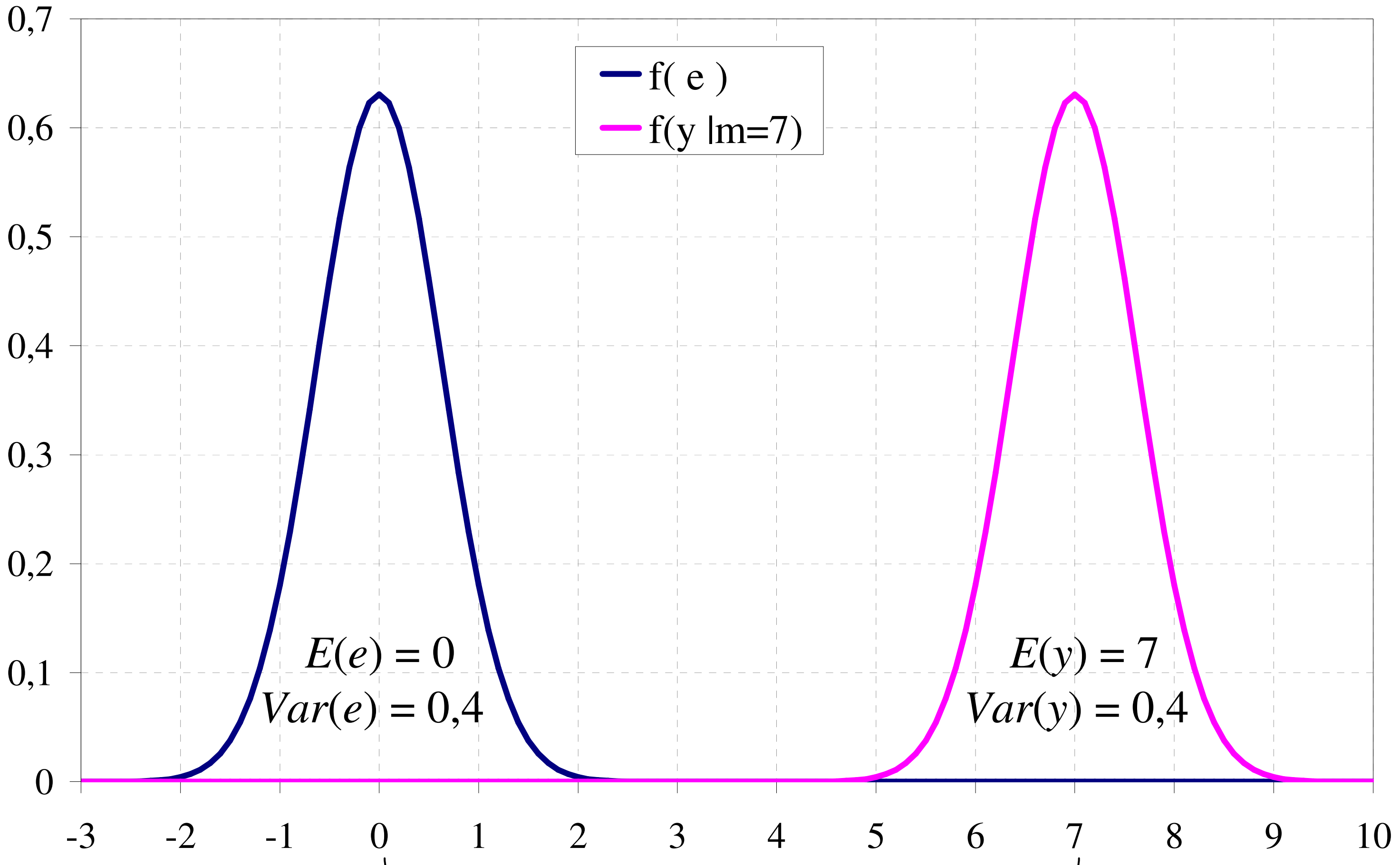
  $\hat{\boldsymbol{\theta}}$ który maksymalizuje l , także maksymalizuje L

  $$\frac{\partial l}{\partial \boldsymbol{\theta}} = \frac{1}{L}\frac{\partial L}{\partial \boldsymbol{\theta}}$$

- score $s(\boldsymbol{\theta};\mathbf{y}) = \frac{\partial l}{\partial \boldsymbol{\theta}}$
- Szukamy takiego $\hat{\boldsymbol{\theta}}$, który rozwiązuje

  $$s(\boldsymbol{\theta};\mathbf{y}) = \frac{\partial l}{\partial \boldsymbol{\theta}} = \mathbf{0}$$

Rozkład zmiennej losowej y



Przesunięcie o m=7, czyli y=m+e

Rozkład zmiennej losowej y

- Funkcja gęstości dla e:

$$f(e) = (2\pi\sigma^2)^{-1/2} \exp(-\frac{e^2}{2\sigma^2})$$

- Funkcja gęstości dla y, kiedy znamy m:
(czyli warunkowa funkcja gęstości…)

$$g(y \mid m = 7) = (2\pi\sigma^2)^{-1/2} \exp(-\frac{(y-7)^2}{2\sigma^2}) =$$

$$= (2\pi\sigma^2)^{-1/2} \exp(-\frac{e^2}{2\sigma^2}) = f(e)$$

Rozkład zmiennej losowej y

- Ogólniej, kiedy $m=\mathbf{xb}$, czyli $y=\mathbf{xb}+e$:
- Wartość oczekiwana y :

$$E(y) = m = E(e + \mathbf{xb}) = \mathbf{xb}$$

$$Var(y) \equiv Var(e + m) \equiv Var(e + \mathbf{xb}) = \sigma^2$$

- Funkcja gęstości y (warunkowa na m):

$$g(y \mid m = \mathbf{xb}) = (2\pi\sigma^2)^{-1/2} \exp(-\frac{(y - \mathbf{xb})^2}{2\sigma^2}) =$$

$$= (2\pi\sigma^2)^{-1/2} \exp(-\frac{e^2}{2\sigma^2}) = f(e)$$

Funkcja wiarygodności

- Funkcja gęstości warunkowa ze względu na parametry = funkcja wiarygodności

$$L(\mathbf{y};\boldsymbol{\theta}) \equiv L(y_1,\ldots,y_N;\boldsymbol{\theta}) \equiv f(y_1,\ldots,y_N \mid \boldsymbol{\theta})$$

$$L(\mathbf{y};\boldsymbol{\theta}) \equiv f(y_1,\ldots,y_N \mid \boldsymbol{\theta}) =$$
$$= f(y_1 \mid \boldsymbol{\theta}) \cdot f(y_2 \mid y_1,\boldsymbol{\theta}) \cdot \ldots \cdot f(y_N \mid \mathbf{y}_{N-1},\boldsymbol{\theta})$$

- Gdyby $y_1,\ldots,y_N \mid \boldsymbol{\theta}$ niezależne:

$$L(\mathbf{y};\boldsymbol{\theta}) \equiv f(y_1,\ldots,y_N \mid \boldsymbol{\theta}) =$$
$$= f(y_1 \mid \boldsymbol{\theta}) \cdot f(y_2 \mid \boldsymbol{\theta}) \cdot \ldots \cdot f(y_N \mid \boldsymbol{\theta})$$

Funkcja wiarygodności

- Zazwyczaj wykorzystujemy: $\ln L$

$$\ln L(\mathbf{y};\boldsymbol{\theta}) \equiv \ln f(y_1,\ldots,y_N \mid \boldsymbol{\theta}) =$$
$$= \ln f(y_1 \mid \boldsymbol{\theta}) + \ln f(y_2 \mid y_1,\boldsymbol{\theta}) + \ldots + \ln f(y_N \mid \mathbf{y}_{N-1},\boldsymbol{\theta})$$

- Dla funkcji regresji liniowej:

$$\ln L(\mathbf{y};\boldsymbol{\theta}) =$$
$$= \ln f(e_1 + \mathbf{x}_1\mathbf{b} \mid \boldsymbol{\theta}) + \ln f(e_2 + \mathbf{x}_2\mathbf{b} \mid y_1,\boldsymbol{\theta}) + \ldots + \ln f(e_N + \mathbf{x}_N\mathbf{b} \mid \mathbf{y}_{N-1},\boldsymbol{\theta}) =$$
$$= \ln f(e_1 \mid \sigma^2) + \ln f(e_2 \mid \sigma^2) + \ldots + \ln f(e_N \mid \sigma^2)$$

Metoda Największej Wiarygodności

1. Dla ustalonych $\mathbf{x}$ i $\mathbf{b}$ wyznacz realizacje składnika losowego (reszty):

$$\hat{e}_i = y_i - \mathbf{x}_i\mathbf{b}, \quad i = 1,\dots,N$$

2. Wyznacz $\ln f(e_i)$:

$$\ln f(\hat{e}_i \,|\, \sigma^2) = \ln[(2\pi\sigma^2)^{-1/2}] - \frac{(\hat{e}_i)^2}{2\sigma^2}$$

Metoda Największej Wiarygodności

3. Wyznacz $\ln L$:

$$\ln L(\mathbf{y};\boldsymbol{\theta}) = \ln f(\hat{e}_1 \mid \sigma^2) + \ln f(\hat{e}_2 \mid \sigma^2) + ... + \ln f(\hat{e}_N \mid \sigma^2)$$

4. Optymalizuj funkcję $\ln L$ poprzez „manipulowanie” wartościami parametrów

Przykłady zastosowań

- Model regresji

$$\hat{e}_i = y_i - \mathbf{x}_i \mathbf{b}, \quad i = 1,\ldots,N$$

$$\ln f(\hat{e}_i \mid \sigma^2) = \ln[(2\pi\sigma^2)^{-1/2}] - \frac{(\hat{e}_i)^2}{2\sigma^2}$$

- Model autoregresji

$$\hat{e}_t = y_t - (a_0 + a_1 y_{t-1} + \ldots + a_1 y_{t-p}), \quad t = p+1,\ldots,N$$

$$\ln f(\hat{e}_t \mid \sigma^2) = \ln[(2\pi\sigma^2)^{-1/2}] - \frac{(\hat{e}_t)^2}{2\sigma^2}$$

Przykłady zastosowań

- Model ARMA

$$y_t = a_0 + a_1 y_{t-1} + ... + a_p y_{t-p} + e_t + b_1 e_{t-1} + ... + b_q e_{t-q}$$

  – warunkowa MNW

$$\hat{e}_t = y_t - (a_0 + a_1 y_{t-1} + ... + a_p y_{t-p} + b_1 \hat{e}_{t-1} + ... + b_q \hat{e}_{t-q}),$$

$$t = \max(p+1, q+1), ..., N$$

$$\hat{e}_t = 0, \qquad t = 1, ..., q$$

$$\ln f(\hat{e}_t \mid \sigma^2) = \ln[(2\pi\sigma^2)^{-1/2}] - \frac{(\hat{e}_t)^2}{2\sigma^2}$$

Przykłady zastosowań

- Model regresji z efektem GARCH(1,1)

$$y_t = \mathbf{x}_t \mathbf{b} + e_t, \qquad e_t \sim N(0, \sigma_t^2),$$

$$\sigma_t^2 = a_0 + a_1 e_{t-1}^2 + a_2 \sigma_{t-1}^2$$

$$a_1 > 0, \quad a_2 > 0, \quad a_1 + a_2 < 1$$

  – estymacja MNW

$$\hat{e}_t = y_t - \mathbf{x}_t \mathbf{b}$$

$$\hat{\sigma}_{\mathrm{t}}^2 = a_0 + a_1 \hat{e}_{t-1}^2 + a_2 \hat{\sigma}_{t-1}^2, \qquad t = 2, \ldots, N$$

$$\hat{e}_1 = 0, \quad \hat{\sigma}_1^2 = \hat{e}_1^2$$

$$\ln f(\hat{e}_t \mid \hat{\sigma}_t^2) = \ln[(2\pi \hat{\sigma}_t^2)^{-1/2}] - \frac{(\hat{e}_t)^2}{2\hat{\sigma}_t^2}$$

Przykłady zastosowań

- Model logitowy

$$\Pr(Y_i = 1) = \frac{\exp(\mathbf{x}_i \mathbf{b})}{1 + \exp(\mathbf{x}_i \mathbf{b})}, \qquad \Pr(Y_i = 0) = \frac{1}{1 + \exp(\mathbf{x}_i \mathbf{b})}$$

- Model probitowy

$$\Pr(Y_i = 1) = F(\mathbf{x}_i \mathbf{b}), \qquad \Pr(Y_i = 0) = 1 - F(\mathbf{x}_i \mathbf{b})$$

$F(\cdot) -$ dystrybuanta rozkladu normalnego

- Estymacja MNW

$$f(y_i \mid \boldsymbol{\theta}) = I(y_i = 1) \cdot \Pr(y_i = 1 \mid \boldsymbol{\theta}) + I(y_i = 0) \cdot \Pr(y_i = 0 \mid \boldsymbol{\theta})$$

$$\ln L(\mathbf{y}; \boldsymbol{\theta}) = \ln f(y_1 \mid \boldsymbol{\theta}) + \ln f(y_2 \mid \boldsymbol{\theta}) + \ldots + \ln f(y_N \mid \boldsymbol{\theta})$$

Własności estymatora MNW

- Zgodność $\operatorname{plim}(\hat{\boldsymbol{\theta}}) = \boldsymbol{\theta}$
- Asymptotyczna normalność $\hat{\boldsymbol{\theta}} \overset{a}{\sim} N(\boldsymbol{\theta}; I^{-1}(\boldsymbol{\theta}))$
  - Macierz informacji $I(\boldsymbol{\theta}) = E\left[\left(\frac{\partial l}{\partial \boldsymbol{\theta}}\right)'\left(\frac{\partial l}{\partial \boldsymbol{\theta}}\right)\right] = -E\left[\frac{\partial^2 l}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}\right]$
  - w praktyce łatwiej policzyć drugie wyrażenie

Własności estymatora MNW

$$\frac{\partial[\partial l/\partial \boldsymbol{\theta}]}{\partial \theta_1} = \begin{bmatrix} \frac{\partial^2 l}{\partial \theta_1^{\ 2}} & \frac{\partial^2 l}{\partial \theta_1 \partial \theta_2} & \cdots & \frac{\partial^2 l}{\partial \theta_1 \partial \theta_k} \end{bmatrix}$$

$$\frac{\partial^2 l}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} = \begin{bmatrix} \frac{\partial^2 l}{\partial \theta_1^{\ 2}} & \frac{\partial^2 l}{\partial \theta_1 \partial \theta_2} & \cdots & \frac{\partial^2 l}{\partial \theta_1 \partial \theta_k} \\ \frac{\partial^2 l}{\partial \theta_2 \partial \theta_1} & \frac{\partial^2 l}{\partial \theta_2^{\ 2}} & \cdots & \frac{\partial^2 l}{\partial \theta_2 \partial \theta_k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 l}{\partial \theta_k \partial \theta_1} & \frac{\partial^2 l}{\partial \theta_k \partial \theta_2} & \cdots & \frac{\partial^2 l}{\partial \theta_k^{\ 2}} \end{bmatrix}$$

- To nie to samo co $\left(\frac{\partial l}{\partial \boldsymbol{\theta}}\right)' \left(\frac{\partial l}{\partial \boldsymbol{\theta}}\right)$

Własności estymatora MNW

- Asymptotycznie efektywny estymator:
  - dla jednego parametru $\sqrt{n}(\hat{\theta}-\theta)\xrightarrow{L}N(0,\sigma^2)$.
    Jeśli jakiś inny estymator jest zgodny i ma asymptotyczny rozkład normalny, to wariancja $\sqrt{n}\tilde{\theta}$ jest większa lub równa σ^2.
  - dla wielu parametrów $\sqrt{n}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})\xrightarrow{L}N(\mathbf{0},\mathbf{V})$.
    Jeśli jakiś inny estymator jest zgodny i ma asymptotyczny rozkład normalny, to $\tilde{\mathbf{V}}-\mathbf{V}$ jest macierzą dodatnio półokreśloną.

Własności estymatora MNW

- Niezmienniczość (invariance):
  - jeśli $\hat{\boldsymbol{\theta}}$ estymator MNW dla $\boldsymbol{\theta}$ i $g(\boldsymbol{\theta})$ ciągła funkcja $\boldsymbol{\theta}$, to $g(\hat{\boldsymbol{\theta}})$ jest estymatorem MNW dla $g(\boldsymbol{\theta})$.

- „Score" ma wartość oczekiwaną zero i wariancję $I(\boldsymbol{\theta})$

Estymacja modelu liniowego

$$y_t = \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + \ldots + \beta_k x_{kt} + u_t$$

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{u} \qquad \mathbf{u} \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$$

$$f(\mathbf{u}) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left(-\frac{\mathbf{u'u}}{2\sigma^2}\right)$$

$$f(\mathbf{y} \mid \mathbf{X}) = f(\mathbf{u}) \left| \frac{\partial \mathbf{u}}{\partial \mathbf{y}} \right| = f(\mathbf{u})$$

$$l = \ln f(\mathbf{y} \mid \mathbf{X}) = -\frac{N}{2}\ln(2\pi) - \frac{N}{2}\ln(\sigma^2) - \frac{\mathbf{u'u}}{2\sigma^2}$$

$$= -\frac{N}{2}\ln(2\pi) - \frac{N}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}(\mathbf{y} - \mathbf{X\beta})'(\mathbf{y} - \mathbf{X\beta})$$

Estymacja modelu liniowego

- Wektor nieznanych parametrów:

$$\boldsymbol{\theta} = [\boldsymbol{\beta}', \sigma^2]'$$

- Po maksymalizacji logarytmu funkcji wiarygodności mamy:

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

$$\hat{\sigma}^2 = \frac{\hat{\mathbf{u}}'\hat{\mathbf{u}}}{N}$$ obciążony estymator, ale zgodny

Estymacja modelu liniowego

- Macierz informacji

$$I(\boldsymbol{\theta}) = I\begin{pmatrix} \boldsymbol{\beta} \\ \sigma^2 \end{pmatrix} = \begin{bmatrix} \frac{1}{\sigma^2}(\mathbf{X'X}) & \mathbf{0} \\ \mathbf{0} & \frac{N}{2\sigma^4} \end{bmatrix}$$

- … i jej odwrotność

$$I^{-1}\begin{pmatrix} \boldsymbol{\beta} \\ \sigma^2 \end{pmatrix} = \begin{bmatrix} \sigma^2(\mathbf{X'X})^{-1} & \mathbf{0} \\ \mathbf{0} & \frac{2\sigma^4}{N} \end{bmatrix}$$

Estymacja modelu liniowego

- Wartość funkcji wiarygodności dla oszacowanych parametrów:

$$L(\mathbf{b},\hat{\sigma}^2) = (2\pi e)^{-N/2}(\hat{\sigma}^2)^{-N/2} = \left(\frac{2\pi e}{N}\right)^{-N/2}(\hat{\mathbf{u}}'\hat{\mathbf{u}})^{-N/2}$$

$$= c \cdot (\hat{\mathbf{u}}'\hat{\mathbf{u}})^{-N/2}$$

Test ilorazu wiarygodności

- Likelihood ratio (LR) test:

$$H_0 : \mathbf{R}\boldsymbol{\beta} = \mathbf{r}$$

- Iloraz wiarygodności:

$$\lambda = \frac{L(b_*, \sigma_*^2)}{L(b, \sigma^2)}$$

- Statystyka testowa:

$$LR = -2 \ln \lambda = 2[\ln L(b, \sigma^2) - \ln L(b_*, \sigma_*^2)] \overset{a}{\sim} \chi^2(m)$$

Test ilorazu wiarygodności

- F. wiarygodności modelu z restrykcjami:

$$l_* = l - \boldsymbol{\mu}'(\mathbf{R}\boldsymbol{\beta} - \mathbf{r})$$

- Estymator $\boldsymbol{\beta}$ identyczny jak dla MNK przy warunkach pobocznych

$$\hat{\mathbf{u}}_* = \mathbf{y} - \mathbf{X}\mathbf{b}_* \qquad \tilde{\sigma}^2 = \frac{\hat{\mathbf{u}}_*'\hat{\mathbf{u}}_*}{N}$$

$$L(\mathbf{b}_*, \tilde{\sigma}^2) = c \cdot (\hat{\mathbf{u}}_*'\hat{\mathbf{u}}_*)^{-N/2}$$

Test ilorazu wiarygodności

- Formuła testu LR dla modelu liniowego

$$LR = n[\ln(\hat{\mathbf{u}}_*'\hat{\mathbf{u}}_*) - \ln(\hat{\mathbf{u}}'\hat{\mathbf{u}})] =$$

$$= n\ln\left(1 + \frac{\hat{\mathbf{u}}_*'\hat{\mathbf{u}}_* - \hat{\mathbf{u}}'\hat{\mathbf{u}}}{\hat{\mathbf{u}}'\hat{\mathbf{u}}}\right) =$$

$$= n\ln\left(\frac{1}{1 - (\hat{\mathbf{u}}_*'\hat{\mathbf{u}}_* - \hat{\mathbf{u}}'\hat{\mathbf{u}})/\hat{\mathbf{u}}'\hat{\mathbf{u}}}\right)$$

Test Walda

- Analogicznie do MNK można wyprowadzić statystykę testu Walda dla MNW:

$$(\mathbf{Rb}-\mathbf{r})'(\mathbf{R}I^{-1}(\boldsymbol{\beta})\mathbf{R}')^{-1}(\mathbf{Rb}-\mathbf{r}) \overset{a}{\sim} \chi^2(m)$$

$$W = (\mathbf{Rb}-\mathbf{r})'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}(\mathbf{Rb}-\mathbf{r})/\hat{\sigma}^2 \overset{a}{\sim} \chi^2(m)$$

$$\hat{\mathbf{u}}_*'\hat{\mathbf{u}}_* = \hat{\mathbf{u}}'\hat{\mathbf{u}} + (\mathbf{r}-\mathbf{Rb})'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{r}-\mathbf{Rb})$$

$$W = \frac{n(\hat{\mathbf{u}}_*'\hat{\mathbf{u}}_* - \hat{\mathbf{u}}'\hat{\mathbf{u}})}{\hat{\mathbf{u}}'\hat{\mathbf{u}}} \sim \chi^2(m)$$

Test mnożnika Lagrange'a

$$s(\boldsymbol{\theta};\mathbf{y}) = \frac{\partial l}{\partial \boldsymbol{\theta}} = \mathbf{0}$$

- Langrange Multiplier (LM) test – score test:

$$LM = s'(\mathbf{b}_*)I^{-1}(\mathbf{b}_*)s(\mathbf{b}_*) \overset{a}{\sim} \chi^2(m)$$

- Do testowania wystarczy oszacować model z restrykcjami!

Test mnożnika Lagrange'a

- Dla modelu liniowego

$$s(\boldsymbol{\theta}) = \begin{bmatrix} \dfrac{\partial l}{\partial \boldsymbol{\beta}} \\ \dfrac{\partial l}{\partial \sigma^2} \end{bmatrix} = \begin{bmatrix} \dfrac{1}{\sigma^2}\mathbf{X}'\mathbf{u} \\ -\dfrac{N}{\sigma^2} + \dfrac{\mathbf{u}'\mathbf{u}}{2\sigma^4} \end{bmatrix}$$

- Dla $\mathbf{Rb}_* = \mathbf{r}$

$$s(\tilde{\boldsymbol{\theta}}) = \begin{bmatrix} \dfrac{1}{\tilde{\sigma}^2}\mathbf{X}'\hat{\mathbf{u}}_* \\ 0 \end{bmatrix}$$

Test mnożnika Lagrange'a

- Po wyprowadzeniu:

$$LM = \begin{bmatrix} \dfrac{1}{\tilde{\sigma}^2}\mathbf{X}'\hat{\mathbf{u}}_* \\ 0 \end{bmatrix}' \begin{bmatrix} \tilde{\sigma}^2(\mathbf{X}'\mathbf{X})^{-1} & \mathbf{0} \\ \mathbf{0} & \dfrac{2\tilde{\sigma}^4}{N} \end{bmatrix} \begin{bmatrix} \dfrac{1}{\tilde{\sigma}^2}\mathbf{X}'\hat{\mathbf{u}}_* \\ 0 \end{bmatrix} = \ldots$$

$$= \frac{n\hat{\mathbf{u}}_*'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\hat{\mathbf{u}}_*}{\hat{\mathbf{u}}_*'\hat{\mathbf{u}}_*} = \ldots$$

$$= nR^2_{(\hat{\mathbf{u}}_* \leftarrow \mathbf{X})} = \frac{n(\hat{\mathbf{u}}_*'\hat{\mathbf{u}}_* - \hat{\mathbf{u}}'\hat{\mathbf{u}})}{\hat{\mathbf{u}}_*'\hat{\mathbf{u}}_*}$$

Porównanie testów

- Która statystyka największa?

$$LR = n\ln\left(1+\frac{\hat{\mathbf{u}}_*'\hat{\mathbf{u}}_* - \hat{\mathbf{u}}'\hat{\mathbf{u}}}{\hat{\mathbf{u}}'\hat{\mathbf{u}}}\right) \approx n\frac{\hat{\mathbf{u}}_*'\hat{\mathbf{u}}_* - \hat{\mathbf{u}}'\hat{\mathbf{u}}}{\hat{\mathbf{u}}'\hat{\mathbf{u}}} - \frac{1}{2}\left(\frac{\hat{\mathbf{u}}_*'\hat{\mathbf{u}}_* - \hat{\mathbf{u}}'\hat{\mathbf{u}}}{\hat{\mathbf{u}}'\hat{\mathbf{u}}}\right)^2$$

$$W = \frac{n(\hat{\mathbf{u}}_*'\hat{\mathbf{u}}_* - \hat{\mathbf{u}}'\hat{\mathbf{u}})}{\hat{\mathbf{u}}'\hat{\mathbf{u}}}$$

$$LM = \frac{n(\hat{\mathbf{u}}_*'\hat{\mathbf{u}}_* - \hat{\mathbf{u}}'\hat{\mathbf{u}})}{\hat{\mathbf{u}}_*'\hat{\mathbf{u}}_*}$$

- W ≥ LR ≥ LM .

Pytania dodatkowe

- Jaką formę modelu („z restrykcjami” czy „bez restrykcji”) należy oszacować przy stosowaniu testu F, Walda, LM i LR?