Basic Econometrics - rewiev

Jerzy Mycielski

- Linear equation

$$y_i = x_{1i}\beta_1 + x_{2i}\beta_2 + \ldots + x_{Ki}\beta_K + \varepsilon_i, \text{ dla } i = 1, \ldots, N,$$

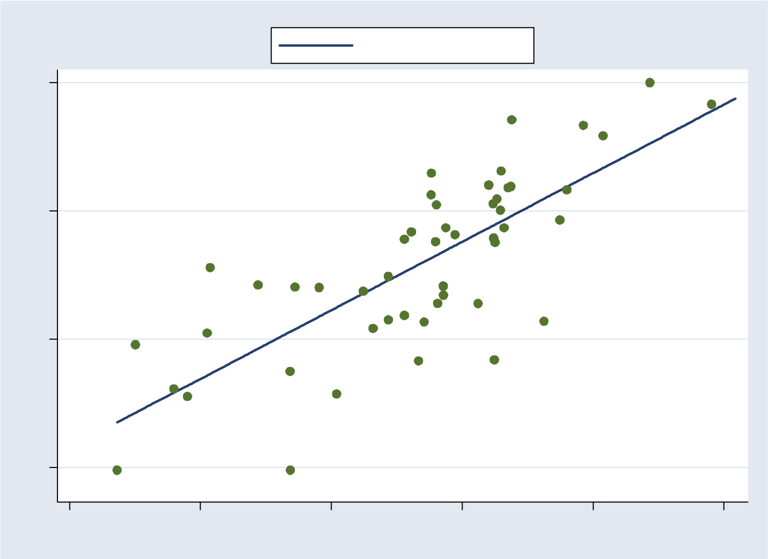
- Elements
  - dependent (endogenous) variable y_i
  - independent (exogenous) variables $x_{1i}, \ldots, x_{Ki}$
  - parameters $\beta_1, \ldots, \beta_K$
  - error term ε_i

- Similar to model

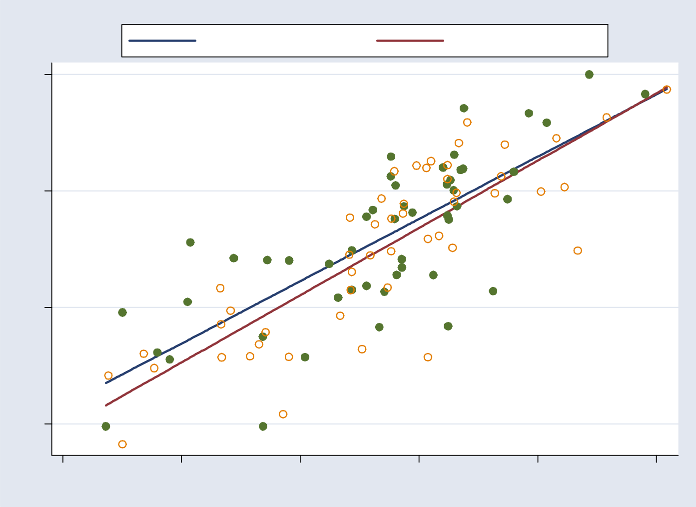
$$\widehat{y}_i = x_{1i} b_1 + x_{2i} b_2 + \ldots + x_{Ki} b_K \,.$$

- Differences:
    - fitted values $\widehat{y}_i$ instead of dependent variable y_i
    - estimates $b_1, \ldots, b_K$ instead of parameters $\beta_1, \ldots, \beta_K$
    - residuals e_i instead of error terms ε_i
- Parameters are nonrandom but estimates are random

Fitted regression line (simulated data)



Same model different sample

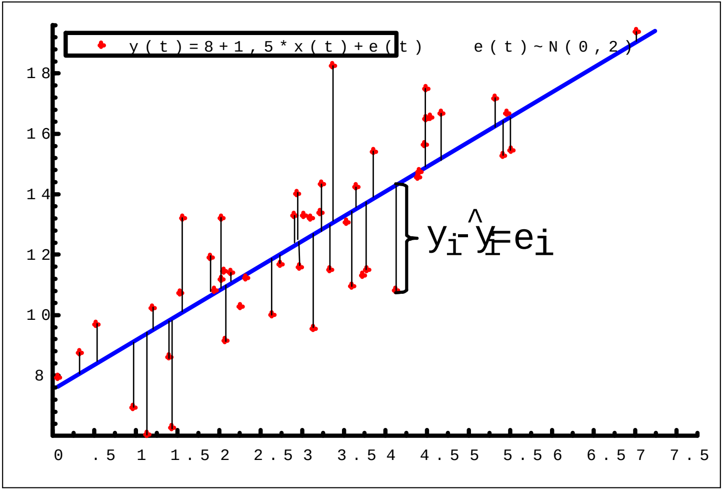


- Definition of the residual

$$e_i = y_i - x_{1i}b_1 - x_{2i}b_2 - \ldots - x_{Ki}b_K = y_i - \widehat{y}_i.$$

- Therefore

$$y_i = \widehat{y}_i + e_i = x_{1i}b_1 + x_{2i}b_2 + \ldots + x_{Ki}b_K + e_i.$$


y (t) = 8 + 1 , 5 * x (t) + e (t)
e (t) ~ N (0 , 2)
y_i - ŷ_i = e_i
1 8
1 6
1 4
1 2
1 0
8
0
. 5
1
1 . 5
2
2 . 5
3
3 . 5
4
4 . 5
5
5 . 5
6
6 . 5
7
7 . 5

- The fit is the best if the sum of squares of residuals is the smallest possible:

$$\min_{\mathbf{b}} S\left(\mathbf{b}\right) = \min_{\mathbf{b}} \sum_{i=1}^{N} \left(y_i - \widehat{y}_i\right)^2 = \min_{\mathbf{b}} \sum_{i=1}^{N} e_i^2 .$$

- Soluton of this minimization problem gives the formula for *OLS* estimator **b**
- This also explains why this estimator is called *Least Squares* estimator

- Matrix formulation of te model

$$\underbrace{\begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}}_{\mathbf{y}} = \underbrace{\begin{bmatrix} x_{11} & \cdots & x_{K1} \\ \vdots & & \\ x_{1N} & \cdots & x_{KN} \end{bmatrix}}_{\mathbf{X}} \underbrace{\begin{bmatrix} \beta_1 \\ \vdots \\ \beta_K \end{bmatrix}}_{\boldsymbol{\beta}} + \underbrace{\begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_N \end{bmatrix}}_{\boldsymbol{\varepsilon}},$$

- Therefore we can write:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

- Similarly

$$\mathbf{e} = \mathbf{y} - \mathbf{X}\mathbf{b} = \mathbf{y} - \widehat{\mathbf{y}}.$$

- We sometimes use as well the notation

$$y_i = \underbrace{\begin{bmatrix} x_{1i} & \cdots & x_{Ki} \end{bmatrix}}_{\mathbf{x}_i} \underbrace{\begin{bmatrix} \beta_1 \\ \vdots \\ \beta_K \end{bmatrix}}_{\boldsymbol{\beta}} + \varepsilon_i,$$

- so that

$$y_i = \mathbf{x}_i \boldsymbol{\beta} + \varepsilon_i, \text{ for } i = 1, \dots, N$$

- First order derivative of $S(\mathbf{b})$ w.r.t. $\mathbf{b}$:

$$\frac{\partial S(\mathbf{b})}{\partial \mathbf{b}} = -2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\mathbf{b}.$$

- First order conditions (system of normal equations)

$$\mathbf{X}'\mathbf{X}\mathbf{b} = \mathbf{X}'\mathbf{y}$$

- Solution (OLS estimator):

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

- But:
  - Matrix $\mathbf{X}$ has to be invertible (if it is not we have perfect collinearity)

- **T**otal **S**um of **S**quares

$$TSS = \sum_{i=1}^{N} (y_i - \overline{y})^2 = (\mathbf{y} - \overline{\mathbf{y}})' (\mathbf{y} - \overline{\mathbf{y}})$$

- **E**xplained **S**um of **S**quares

$$ESS = \sum_{i=1}^{N} \left(\widehat{y}_i - \overline{\widehat{y}}\right)^2 = \left(\widehat{\mathbf{y}} - \overline{\widehat{\mathbf{y}}}\right)' \left(\widehat{\mathbf{y}} - \overline{\widehat{\mathbf{y}}}\right)$$

- **R**esidual **S**um of **S**quares

$$RSS = \sum_{i=1}^{N} e_i^2 = \mathbf{e}'\mathbf{e}$$

- It can be proven that

$$TSS = ESS + RSS$$

- So we can define:

$$R^2 = \frac{ESS}{TSS} = \frac{\text{explained variation}}{\text{total variation}}$$

- and

$$0 \leq R^2 \leq 1$$

- R^2 can be interpreted as percent of total variation of dependent variable explained by the model

- Dummy variable can only take values 0 or 1
- Define a model

$$y_i = \beta_1 x_{1i} + \ldots + \beta_K x_{Ki} + \gamma D_i + \varepsilon_i.$$

- For $D_j = 0$

$$y_i = \beta_1 x_{1i} + \ldots + \beta_K x_{Ki} + \varepsilon_i.$$

- For $D_j = 1$,

$$y_j = \beta_1 x_{1j} + \ldots + \beta_K x_{Kj} + \gamma + \varepsilon_j.$$

- So the difference between expected values of y_i and y_j is equal to

$$\mathrm{E}\left(y_j\right) - \mathrm{E}\left(y_i\right) = \gamma.$$

- General case

$$y_i = \mathbf{x}_i \boldsymbol{\beta} + \gamma_0 + \sum_{s=2}^{S} D_{s,i} \gamma_s + \varepsilon_i.$$

- For $z_i = 1$, $z_j = s$:

$$\mathrm{E}\left(y_j\right) - \mathrm{E}\left(y_i\right) = \mathbf{x}\boldsymbol{\beta} + \gamma_0 + \gamma_s - \mathbf{x}\boldsymbol{\beta} - \gamma_0 = \gamma_s$$

1. Model is linear:

$$y_i = x_{1i}\beta_1 + \ldots + x_{Ki}\beta_K + \varepsilon_i \quad \text{for} \quad i = 1, \ldots, N.$$

or:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}.$$

2. Explanarory variables $x_{1i}, \ldots, x_{ki}$ are nonrandom for $i = 1, \ldots, N$
3. Expected value of the error therm is equal to zero:

$$\mathrm{E}\left(\varepsilon_i\right) = 0 \quad \text{dla} \quad i = 1, \ldots, N.$$

or:

$$\mathrm{E}\left(\boldsymbol{\varepsilon}\right) = 0.$$

4. Covariance (correlation) between two error terms is equal to zero:
$$\mathrm{Cov}\left(\varepsilon_i, \varepsilon_j\right) = 0 \quad \text{dla} \quad i \neq j.$$
Absence of autocorrelation
5. Variance is the same for all observations (homoscedasticity):
$$\mathrm{Var}\left(\varepsilon_i\right) = \sigma^2 \quad \text{dla} \quad i = 1, \ldots, N.$$

- Two last assumptions can be formulated as $\mathrm{Var}\left(\boldsymbol{\varepsilon}\right) = \sigma^2 \mathbf{I}$

- OLS estimator is unbiased

$$
\begin{aligned}
&\mathrm{E}\left(\mathbf{b}\right) = \mathrm{E}\left(\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{X}\beta\right) + \mathrm{E}\left(\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\varepsilon\right) \\
&= \mathrm{E}\left(\beta\right) + \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\underbrace{\mathrm{E}\left(\varepsilon\right)}_{\mathbf{0}} = \beta.
\end{aligned}
$$

- Variance of the OLS estimator is equal to

$$
\begin{aligned}
&\mathrm{Var}\left(\mathbf{b}\right) = \mathrm{Var}\left(\beta + \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\varepsilon\right) \\
&= \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\underbrace{\mathrm{Var}\left(\varepsilon\right)}_{\sigma^2\mathbf{I}}\mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1} \\
&= \sigma^2\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{I}\mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1} \\
&= \sigma^2\left(\mathbf{X}'\mathbf{X}\right)^{-1} = \circ,
\end{aligned}
$$

- It can be proven that

$$s^2 = \frac{\mathbf{e}'\mathbf{e}}{N-K} = \frac{\sum_{i=1}^{N} e_i^2}{N-K}.$$

is unbiased estimator of σ^2

- So ˚ can be estimated with

$$\hat{˚} = s^2 \left(\mathbf{X}'\mathbf{X}\right)^{-1}$$

Theorem (Gauss-Markov)

Under assumptions of CRM, OLS estimator is best linear unbiased estimator (BLUE)

- Additional assumption of normality of error term $\varepsilon \sim N\left(0, \sigma^2 \mathbf{I}\right)$ needed for derivation of statisitics distributions
- Simple hypotesis

$$\begin{cases} H_0 : \beta_k = \beta_k^* \\ H_1 : \beta_k \neq \beta_k^* \end{cases}.$$

- Test statistics

$$t = \frac{b_k - \beta_k^*}{\widehat{se}\left(b_k\right)} \sim t_{N-K}$$

- Most popular case - testing significance of the variables

$$\begin{cases} H_0 : \beta_k = 0 \\ H_1 : \beta_k \neq 0 \end{cases}.$$

- Indeed if $\beta_k = 0$ then variable is redundant in our model

$$y_i = \beta_0 + \ldots + \underbrace{\beta_k}_{0} x_{ki} + \ldots + \beta_K x_{Ki} + \varepsilon_i,$$

- Statistics $t = \frac{b_k}{\widehat{se}(b_k)}$

- General case: joint hypotesis

$$H_0 : \mathbf{H}\beta = \mathbf{h}$$

- Statistics

$$F = \frac{\left(\mathbf{e}_R'\mathbf{e}_R - \mathbf{e}'\mathbf{e}\right)/\,g}{\mathbf{e}'\mathbf{e}/\left(N-K\right)} \sim F\left(g, N-K\right),$$

- where $\mathbf{e}_R$ are residuals of the restricted model (model estimated under assumption that H_0 is true)

- Significance intervals

$$\Pr\left(b_k - \widehat{se}\left(b_k\right) t_{\frac{\alpha}{2}} < \beta_k < b_k + \widehat{se}\left(b_k\right) t_{\frac{\alpha}{2}}\right) = 1 - \alpha$$