Econometric Methods
PC-tutorial: M-Estimation

Le Nga Tran

Winter Term 2023/24

January 9, 2024

Examples

Example 1: the impact of education on woman's fertility. The dependent variable, *children*: number of living children, is non-negative value

$$E(children|x) = exp(x\beta)$$

Example 2 [1]: y is the probability of owing home ($0 \leq y \leq 1$) and x is income

$$E(y|x) = \frac{exp(x\beta)}{1+exp(x\beta)}$$

[1] Basic Econometrics, Gujarati and Porter

M-Estimation

the nonlinear conditional expectation model (θ_0 is true parameters)

$$E(y|x) = m(x, \theta_0) \tag{1}$$

The M estimators $\hat{\theta}$ minimizes the certain sample functions

$$\min_{\theta \subset \Theta} \quad N^{-1} \sum_{i=1}^{N} q(w_i, \theta) \tag{2a}$$

$q(w_i, \theta)$ is objective function, where $w_i = (y_i, x_i)$

Non-linear least squares (NLS)

$$\min_{\theta \subset \Theta} \quad N^{-1} \sum_{i=1}^{N} q(w_i, \theta) = N^{-1} \sum_{i=1}^{N} [y_i - m(x_i, \theta)]^2 \tag{3a}$$

Stata command for NSL `nl`

Interpretation: Marginal effects at mean(MEM) and Average marginal effects (AME)

Example: we are interest in the effects of education on woman's fertility

Linear regression

$$children = \beta_0 + \beta_1 educ + z\gamma + u = x\beta + u \tag{4}$$

$\frac{\partial E(children|x)}{\partial educ} = \beta_1 = const$

Non-linear regression

$$children = exp(\beta_0 + \beta_1 educ + z\gamma) + u = exp(x\beta) + u \tag{5}$$

$\frac{\partial E(children|x)}{\partial educ} = \beta_1 exp(x\beta)$, depends on the value of educ, and others control variables. This is also called the marginal effect

Marignal effects at means (MEM): the marginal effects evaluated at the average values of the regressors.

$$MEM = \frac{\partial children}{\partial educ}|(x = \bar{x}) = \hat{\beta}_1 exp(\bar{x}\hat{\beta}) \tag{6}$$

Example: MEM of education is -0.04 means keeping other variables at their average values, one additional year of education form its average value of 5 to 6 years has negative impact on the woman's fertility .

Average marginal effects (AME): the marginal effects averaged across all observations in the data:

$$AME = \frac{1}{N}\sum_{i}^{N}(\frac{\partial children}{\partial educ}|x = x_i) = \frac{1}{N}\sum_{i}^{N}\hat{\beta}_i exp(x_i\hat{\beta}) \tag{7}$$

Example: AME of education is -0.043 means one additional year of education leads to an decrease in the number of living children by 0.043 on average while keeping other variables constant.

Exe1

Consider the Stata dataset fertil2.dta. The aim is to estimate the effects of education on women's fertility in Botswana. Estimate the model

$$children = exp(x\beta) + u \tag{8}$$

the response variable, *children*, is the number of living children. The explanatory variables x are years of schooling (*educ*), age of the woman (*age*), age squared (*agesq*), and binary indicators for ever married (*evermarr*), living in an urban area (*urban*), having electricity (*electric*), and owning a television (*tv*).

a) Estimate the model by apply the nonlinear least squares with initial values of zero. Why may this model be more appropriate than linear OLS?

$$E(y|x) = exp(x\beta)$$

M-estimator $\hat{\theta}$ minimizes

$$\min \sum_i^N (y_i - exp(x_i\beta))^2$$

Stata command for nls : `nl` Stata solves the minimization problem by running iterative method, which requires initial values for parameters.

b) Re-estimate with different initial values. Discuss.

c) Compute the average partial effects of education and age

d) Compute the partial effects of education and age at the sample average. Interpret. What may be "problematic" with using a sample average?

e) Compute the partial effects of education and age at educ = 5, evermarr = 0, urban = 0, electric = 0, tv = 0 for different ages of 15, 20, . . . , 45 years. Interpret your results.

Exe2

Consider the crime data set of Agresti and Finlay (1997) for the US states. Suppose the violent crime rate $crime_i$ (number of violent crimes per 100,000 people) can be explained by the covariates $poverty_i$ (percent of population living under poverty line), $single_i$ (percent of population that are single parents), and a constant. Assume there are no endogeneity problems like reverse causality.

a) Load the data set into Stata and perform a robust regression using the command rreg crime poverty single.
Stata then applies a robust regression technique that partially relies on the Huber estimator. Interpret the estimated parameter values. Are they statistically and economically significant?
Model:

$$crime = \beta_0 + \beta_1 poverty + \beta_2 single + u \qquad (9)$$

Robust regression (`rreg`), Stata doesn't provide robust standard errors for `rreg`

b) Compare the results to an OLS regression. Is the difference in estimated effects relevant?

c) To understand why there are differences, regress $crime_i$ on $poverty_i$ using (i) robust estimation, (ii) OLS estimation, and (iii) OLS estimation excluding the last observation (... if state!="dc"). Predict $crime_i$ in each case. Then scatter $crime_i$ against $poverty_i$ and add the three regression lines to this scatter graph. Discuss you findings.