Эконометрика-2 ММАЭ
Семинар 18

Лекции: А.А. Пересецкий
Семинары: Е.С. Вакуленко
Составитель: Е.Ю. Назруллаева

Problem 1. Estimating the Economic Model of Crime.

Binary response models

Becker (1968) [Becker, Gary. 1968. Crime and Punishment: An Economic Approach. Journal of Political. Economy. 78: 169-217] introduced an economic model explaining the number of crimes. The main implication of this model is that the number of crimes depends negatively on the probability to be arrested, the probability to be convicted conditional on being arrested, the probability to be imprisoned conditional on being convicted, and the average length of the imprisonment sentence. Since 1968, many empirical studies have tested the empirical implications of Becker's model, usually with cross-section data. Cornwell and Trumbull (1994) use panel data and their results suggest that the cross-section based estimates can be misleading.
The data are a random sample of 2725 Californian men, born in either 1960 or 1961, with at least one arrest prior to 1986 since age 18.

Data

The data set **"CRIME.dta"** contains data on 90 counties in North Carolina covering the years 1981 till 1987. The variables are as follows, stored in the order given below:

avgsen = average length of sentences served since age 18 (in months)
black =1 if black
born60 =1 if born in 1960
durat = recent unemployment duration (the number of quarters since the individual last had positive earnings or was released from prison)
hispan =1 if Hispanic
inc86 = reported legal earnings in 1986, tens of thousands $
inc86sq = inc86 squared
narr86 = the number of times a man was arrested
nfarr86 = # felony arrests, 1986
nparr86 = # property crime arrests, 1986
pcnv = the proportion of the prior arrests leading to conviction
pcnvsq = pcnv squared
pt86sq
ptime86 = months spent in prison in 1986
qemp86 = number of quarters in 1986 during which the man was legally employed
tottime = the number of months spent in prison prior to 1986 (since 18)

Exercise (Stata)

1. Estimate a binary model for *crime86*, where the explanatory variables are *pcnv*, *avgsen*, *tottime*, *ptime86*, *qemp86*, *inc86*, *durat*, *black*, *hispan*, *born60*

```
gen crime86=0
replace crime86=1 if narr86>0
```

"Logit" model:

```
logit crime86 pcnv avgsen tottime ptime86 qemp86 inc86 durat black hispan born60
```

"Probit" model:

```
probit crime86 pcnv avgsen tottime ptime86 qemp86 inc86 durat black hispan born60
```

AIC, BIC

```
estat ic
```

Goodness of fit – different measures for "pseudo R-squares":

```
net install sg145.pkg
fitstat, sav(r2_1)
```

Postestimation

```
predict plogit, pr
sum plogit
```

Classification table: correcty predicted y = 0 & y = 1

```
estat clas, cutoff(0.5)
```

The fraction of observations $y=1$ that are correctly predicted is termed the sensitivity, while the fraction of observations $y=0$ that are correctly predicted is known as specifcity.

Sensitivity & specificity graph: choosing the optimal cutoff level

```
lsens
estat clas, cutoff(0.27)
```

Goodness-of-fit Hosmer-Lemeshov test

```
estat gof
```

Interpretation: marginal effects

$$\frac{\partial P(y=1|x)}{\partial x_k} = g\left(\bar{x}\hat{\beta}\right)\hat{\beta}_k$$

```
mfx
mfx, at(.357787 .632294 .838752 .387156 2.30903 54.967 2.25138 1 0 1)
mfx if black==1&hisp==0&born60==1
```

$$\frac{\partial P(y=1|x)}{\partial x_k} = \frac{1}{n}\sum g\left(x_i\hat{\beta}\right)\hat{\beta}_k$$

Alternative variant: calculating of marginal effects

```
net install st0086.pkg
margeff, dummies(black hispan born60)
```

Differences between the standard command "mfx" and the package "margeff" from Stata Journal: http://www.stata-journal.com/sjpdf.html?articlenum=st0086

- MEM (marginal effect in mean values) might both underestimate and overestimate AME (average marginal effect), depending solely on the sign of the second derivative of the density function
- The difference between AME and MEM is large when the parameter estimates are large
- MEM for dummy variables: dummy variables raise a more fundamental problem if the regression model includes several dummies
- Standard errors for AME calculated using delta-method

2. Use the **ordered logit model** for *narr012* (= 0 if narr86=0; =1 if narr86=1; =2 if narr86>1), the explanatory variables are *pcnv, avgsen, tottime, ptime86, qemp86, inc86, durat, black, hispan, born60*.
   For each observation compute the estimate for the probability of no arrest in 1986.

   $y^* = x\beta + u : \ u \sim N(0,1)$ ordered probit model, $u \sim logistic$ ordered logit model

$$y = \begin{cases} 0, & y^* \le c_1 \\ 1, & c_1 < y^* \le c_2 \\ 2, & c_2 < y^* \le c_3 \\ \vdots & \\ J, & y^* > c_J \end{cases}$$

gen narr012=0
replace narr012=1 if narr86==1
replace narr012=2 if narr86>1
ologit narr012 pcnv avgsen tottime ptime86 qemp86 inc86 durat black hispan born60

Interpretation: a "+" sign of the coefficient tells whether the choice probabilities shift to higher categories when the independent variable increases.

predict pologit, p outcome(0)

$$P(y=0|x) = P(y^* \le c_1 | x) = G(x'\beta + \varepsilon \le c_1) = G(c_1 - x'\beta)$$
$$P(y=1|x) = P(c_1 < y^* \le c_2 | x) = G(c_2 - x'\beta) - G(c_1 - x'\beta)$$
$$P(y=2|x) = P(c_2 < y^* \le c_3 | x) = G(c_3 - x'\beta) - G(c_2 - x'\beta)$$
$$\vdots$$
$$P(y=J|x) = P(y^* > c_J | x) = 1 - G(c_J - x'\beta)$$

mfx, predict(outcome(0))
…
mfx, predict(outcome(2))

$$\frac{\partial P(y=0|x)}{\partial x_k}=-g\left(c_1-x\hat{\beta}\right)\hat{\beta}_k$$
$$\frac{\partial P(y=1|x)}{\partial x_k}=-\left[g\left(c_2-x\hat{\beta}\right)-g\left(c_1-x\hat{\beta}\right)\right]\beta_k$$
$$\frac{\partial P(y=2|x)}{\partial x_k}=\left[g\left(c_2-x\hat{\beta}\right)-g\left(c_3-x\hat{\beta}\right)\right]\beta_k$$
$$\vdots$$
$$\frac{\partial P(y=J|x)}{\partial x_k}=g\left(c_J-x\hat{\beta}\right)\beta_k$$

Wald test for the equivalence of the estimates for cut levels:

test _b[/cut1]=_b[/cut2]

3. Compare the prediction value of these two models. For example? Compare the number of right predictions for a threshold 0.5 (i.e. no arrests predicted if the probability of this event is more than 0.5)

count if plogit<=0.5&crime86==0
count if plogit>0.5&crime86==1

count if pologit>0.5&narr012==0
count if pologit<=0.5&narr012!=0

Multinomial models

Multinomial logit

Пусть есть $j=1,...,m$ неупорядоченных альтернатив

"случайная полезность"
$P(y=j\,|\,x)=P\left(u_j+\varepsilon_j>u_k+\varepsilon_k\right)$

Цель: нужно знать плотность совместного распределения случайных ошибок ε, имеет простое аналитическое представление для распределения экстремальных значений:

$$F(z)=\exp\left(-e^{-z}\right)$$

Тогда

$$P(y=j\,|\,x)=\frac{\exp\left(u_j\right)}{\exp\left(u_1\right)+...+\exp\left(u_m\right)}$$

Предполагаем

$$u_j=x'\beta_j$$

Нормировка (для идентифицируемости модели)

$$u_1=x'\beta_1=0$$

Следовательно

$$P(y=1|x)=\frac{1}{1+\exp(u_2)+...+\exp(u_m)}$$

$$P(y=j|x)=\frac{\exp(u_j)}{1+\exp(u_2)+...+\exp(u_m)}$$

Problem 2. *(Wooldridge – CrossSection&Panel).*
The data in PENSION.dta are a subset of data used by Papke (1998) in assessing the impact of allowing individuals to choose their own allocations on asset allocation in pension plans. Initially, Papke codes the responses ''mostly bonds,'' ''mixed,'' and ''mostly stocks'' as 0, 50, and 100, and uses a linear regression model estimated by OLS. The binary explanatory variable choice is unity if the person has choice in how his or her pension fund is invested.

id – family identifier
pyears – years in pension plan
prftshr – =1 if profit sharing plan (employee receives a percentage of those profits based on the company's earnings)
choice – =1 if can choose method invest
female – =1 if female
married – =1 if married
age – age in years
educ – highest grade completed
finc25 – $15,000 < faminc92 <= $25,000
finc35 – $25,000 < faminc92 <= $35,000
finc50 – $35,000 < faminc92 <= $50,000
finc75 – $50,000 < faminc92 <= $75,000
finc100 – $75,000 < faminc92 <= $100,000
finc101 – $100,000 < faminc92
wealth89 – net worth, 1989, $1000
black – =1 if black
stckin89 – =1 if owned stock in 1989
irain89 – =1 if had IRA in 1989 (individual retirement savings program)
pctstck – 0=mstbnds, 50=mixed, 100=mststcks

OLS
reg pctstck choice age educ wealth89 prftshr

Controlling for age, education, gender, race, marital status, income (via a set of dummy variables), wealth, and whether the plan is profit sharing, gives the OLS estimate $\hat{\beta}_{choice}=12.05$. This result means that, other things equal, a person having choice has about 12 percentage points more assets in stocks.

mlogit pctstck choice age educ wealth89 prftshr, baseoutcome(0)

Mlogit: the independence of irrelevant alternatives (IIA) property from rational choice theory!
The ratios of the probabilities (odds), ex. $\frac{P(y=j)}{P(y=0)}$, are independent of one another.
Otherwise: use **nested models** or **multinomial probit**.

Interpretation: $\hat{\beta}_j$ – parameters of a binary choice logit model between alternative j and the base alternative 0. So a positive coefficient from *mlogit* means that as the regressor increases, we are more likely to choose alternative j than the base alternative 0.

mlogit pctstck choice age educ wealth89 prftshr, rr baseoutcome(0)

Relative-risk ratio (relative odds) for mlogit models:

$$\frac{P(y_i = j)}{P(y_i = 1)} = exp(x_i \beta_j)$$

or

Interpretation: $exp(\beta_j)$ – gives proportionate change in the relative risk of choosing alternative j rather than alternative 0 when x_i changes by 1 unit.

Problem 3. *Задача 12.14 из учебника Катышева, Магнуса и Пересецкого.*

Пусть $y_t^* = x_t'\beta + \varepsilon_t$, где ошибки ε_t имеют плотность распределения $f(x)$ и

$$y_t = \begin{cases} \alpha_1, & \textit{если } y_t^* \le \alpha_1, \\ y_t^*, & \textit{если } \alpha_1 < y_t^* < \alpha_2, \\ \alpha_2 & \textit{если } y_t^* \ge \alpha_2. \end{cases}$$

a) Найдите распределение y_t.
b) Найдите логарифмическую функцию правдоподобия для оценивания вектора β.
c) Найдите $(\partial Ey)/(\partial x)$.

Problem 4. *(Задача составлена на основе задачи 12.5 из Сборника задач к начальному курсу эконометрики Катышева, Магнуса, Пересецкого и Головань).*

Рассмотрим модель бинарного выбора $\mathbb{P}\{Y_t = 1\} = F(\alpha + \beta d_t)$, где d – фиктивная переменная (принимающая значения 0 или 1). Ниже представлены результаты 100 наблюдений:

	$y=0$	$y=1$
$d=0$	10	30
$d=1$	40	20

A. Используя логит-модель, оцените параметры α и β.

B. При помощи теста отношения правдоподобия протестируйте гипотезу $H_0 : \beta = 0$ на уровне значимости 5%.

C. *Протестируйте гипотезу $H_0 : \beta = 0$ на уровне значимости 5% с помощью теста Вальда и теста множителей Лагранжа.