SampleSelection

直接使用sampleSelection包自带的示例

参照（Greene，2003）的示例example 22.8, page 786

require(sampleSelection)

## Loading required package: sampleSelection  
## Loading required package: maxLik  
## Loading required package: miscTools  
##   
## Please cite the 'maxLik' package as:  
## Henningsen, Arne and Toomet, Ott (2011). maxLik: A package for maximum likelihood estimation in R. Computational Statistics 26(3), 443-458. DOI 10.1007/s00180-010-0217-1.  
##   
## If you have questions, suggestions, or comments regarding the 'maxLik' package, please use a forum or 'tracker' at maxLik's R-Forge site:  
## https://r-forge.r-project.org/projects/maxlik/

data(Mroz87)  
Mroz87$kids <- ( Mroz87$kids5 + Mroz87$kids618 > 0 )  
  
  
# Heckman两阶段估计  
heck<- heckit( lfp ~ age + I( age^2 ) + faminc + kids + educ,  
 wage ~ exper + I( exper^2 ) + educ + city, data=Mroz87 )  
  
# 极大似然估计  
ml<-selection( lfp ~ age + I( age^2 ) + faminc + kids + educ,  
 wage ~ exper + I( exper^2 ) + educ + city, data=Mroz87 )   
  
summary(heck)

## --------------------------------------------  
## Tobit 2 model (sample selection model)  
## 2-step Heckman / heckit estimation  
## 753 observations (325 censored and 428 observed)  
## 14 free parameters (df = 740)  
## Probit selection equation:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) -4.16e+00 1.40e+00 -2.96 0.00313  
## age 1.85e-01 6.60e-02 2.81 0.00508  
## I(age^2) -2.43e-03 7.74e-04 -3.14 0.00178  
## faminc 4.58e-06 4.21e-06 1.09 0.27654  
## kidsTRUE -4.49e-01 1.31e-01 -3.43 0.00064  
## educ 9.82e-02 2.30e-02 4.27 2.2e-05  
## Outcome equation:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) -0.971200 2.059351 -0.47 0.64  
## exper 0.021061 0.062465 0.34 0.74  
## I(exper^2) 0.000137 0.001878 0.07 0.94  
## educ 0.417017 0.100250 4.16 3.6e-05  
## city 0.443838 0.315898 1.41 0.16  
## Multiple R-Squared:0.126, Adjusted R-Squared:0.116  
## Error terms:  
## Estimate Std. Error t value Pr(>|t|)  
## invMillsRatio -1.098 1.266 -0.87 0.39  
## sigma 3.200 NA NA NA  
## rho -0.343 NA NA NA  
## --------------------------------------------

summary(ml)

## --------------------------------------------  
## Tobit 2 model (sample selection model)  
## Maximum Likelihood estimation  
## Newton-Raphson maximisation, 6 iterations  
## Return code 1: gradient close to zero  
## Log-Likelihood: -1581   
## 753 observations (325 censored and 428 observed)  
## 13 free parameters (df = 740)  
## Probit selection equation:  
## Estimate Std. error t value Pr(> t)  
## (Intercept) -4.12e+00 1.40e+00 -2.94 0.00327  
## age 1.84e-01 6.59e-02 2.79 0.00521  
## I(age^2) -2.41e-03 7.72e-04 -3.12 0.00182  
## faminc 5.68e-06 4.42e-06 1.29 0.19838  
## kidsTRUE -4.51e-01 1.30e-01 -3.46 0.00054  
## educ 9.53e-02 2.32e-02 4.12 3.9e-05  
## Outcome equation:  
## Estimate Std. error t value Pr(> t)  
## (Intercept) -1.963024 1.198221 -1.64 0.10  
## exper 0.027868 0.061551 0.45 0.65  
## I(exper^2) -0.000104 0.001839 -0.06 0.95  
## educ 0.457005 0.073230 6.24 4.4e-10  
## city 0.446529 0.315921 1.41 0.16  
## Error terms:  
## Estimate Std. error t value Pr(> t)  
## sigma 3.108 0.114 27.3 <2e-16  
## rho -0.132 0.165 -0.8 0.42  
## --------------------------------------------

## Example using binary outcome for selection model.  
## We estimate the probability of womens' education on their  
## chances to get high wage (> $5/hr in 1975 USD), using PSID data  
## We use education as explanatory variable  
## and add age, kids, and non-work income as exclusion restrictions.  
  
m <- selection(lfp ~ educ + age + kids5 + kids618 + nwifeinc,  
 wage >= 5 ~ educ, data = Mroz87 )  
summary(m)

## --------------------------------------------  
## Tobit 2 model (sample selection model)  
## Maximum Likelihood estimation  
## BHHH maximisation, 10 iterations  
## Return code 2: successive function values within tolerance limit  
## Log-Likelihood: -653.2   
## 753 observations (325 censored and 428 observed)  
## 9 free parameters (df = 744)  
## Probit selection equation:  
## Estimate Std. error t value Pr(> t)  
## (Intercept) 0.43031 0.47597 0.90 0.37  
## educ 0.15622 0.02381 6.56 5.4e-11  
## age -0.03471 0.00765 -4.54 5.7e-06  
## kids5 -0.89054 0.11266 -7.90 2.7e-15  
## kids618 -0.03816 0.03932 -0.97 0.33  
## nwifeinc -0.02095 0.00432 -4.85 1.2e-06  
## Outcome equation:  
## Estimate Std. error t value Pr(> t)  
## (Intercept) -4.5213 0.5611 -8.06 7.7e-16  
## educ 0.2879 0.0369 7.80 6.2e-15  
## Error terms:  
## Estimate Std. error t value Pr(> t)  
## rho 0.116 0.271 0.43 0.67  
## --------------------------------------------