## Day 4: Probit, logit, Tobit and U-shape model

Learn how to conduct probit/logit regression

Learn how to conduct Tobit model

Learn how to verify U shape relationship

Learn how to use loop simplify the code

1. Probit and logit regression

Probit regression and logit regression are both methods used for modeling binary outcome variables, where the response variable can take on only two possible outcomes (usually coded as 0 and 1).

The Probit model assumes that the random disturbance term of the model follows a normal distribution; The random perturbation term corresponding to the Logit model is assumed to follow a Logistic distribution.

OLS→FE→RE

----------------------------------------------------

webuse union, clear

logit union age grade

probit union age grade

xtset idcode year

xtlogit union age grade,fe

xtprobit union age grade,fe

xtlogit union age grade,re

xtprobit union age grade,re

----------------------------------------------------

1. Tobit model

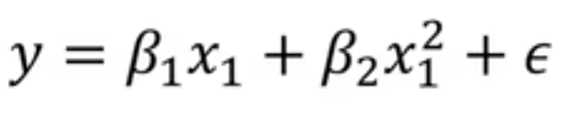
The Tobit model, also known as a censored regression model, is a statistical model used for analyzing data where the dependent variable is either left-censored, right-censored, or both.

----------------------------------------------------

sysuse nlsw88,clear

tobit wage hours ttl\_exp tenure , ll(0) ul(#) vce(cluster idcode )

1. U-shape model



----------------------------------------------------

ssc install utest, replace

clear

open the file “U-shape.dta”

twoway(scatter y x, mcolor(yellow)) (qfit y x)

applot y x,quadratic

gen x2=x^2

reg y x x2

utest x x2

sum x

----------------------------------------------------

If the "extreme point" detected is not within the data range, we cannot reject the original hypothesis.

To test whether there is a U-shape relationship between wage and ttl\_exp

----------------------------------------------------

sysuse nlsw88,clear

---------------------------------------------------

1. Loop

sysuse auto,clear

----------------------------------------------------

gen lnprice=ln(price)

generate lnmpg = ln(mpg)

generate lnrep78 = ln(rep78)

generate lnheadroom = ln(headroom)

generate lntrunk = ln(trunk)

generate lnweight = ln(weight)

generate lnlength = ln(length)

foreach x of varlist price mpg rep78 headroom trunk weight length{

gen ln`x' = ln(`x')

}

forvalue i=1(1)20{

gen x\_`i' =0

}

forvalue i=1(2)20{

drop x\_`i'

}

drop x\_\*