Linear probability, Probit, and Logit models

Linear probability model

- Often the dependent variable is a dummy variable.
- For example, if you would like to find the determinants of labor force participation, the dependent variable is 1 if the person is working(participate in the labor force), and 0 if the person is not working(does not participate in the labor force).

- Simplest way to predict the probability that the person would participate in the labor force is to estimate the following model using OLS.

$$inlf=\beta_0+\beta_1 x_1+\ldots+\beta_k x_k+u$$

Where inlf =1 if the person is in the labor force and inlf=0 if the person is not in the labor force.

This type of the model is called **the linear probability model**.

- As long as explanatory variables are not correlated with the error term, this simple method is perfectly fine.
- However, a caution is necessary. When the dependent variable is a dummy variable, $var(u|X)=X\beta[1- X\beta]$ where $X\beta$ is a short hand notation for $(\beta_0+\beta_1 x_1+\ldots+\beta_k x_k)$
- Thus, the variance depends on the values of the x-variables, which means that it is heteroskedastic. **So, you should always use the robust-standard errors.**

Exercise

- Using Mroz.dta, estimate the following linear probability model of labor force participation.

$$inlf=\beta_0+\beta_1 nwifinc+\beta_2 educ+\beta_3 exper$$
$$+\beta_4 exper+\beta_5 age+\beta_6 kidslt6+\beta_7 kidsge6$$
$$+u$$

Answer

```
. reg  inlf  nwifeinc educ exper expersq age  kidslt6 kidsge6, robust

Linear regression                                      Number of obs =     753
                                                       F(  7,   745) =   62.48
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.2642
                                                       Root MSE      =  .42713
```

inlf	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
nwifeinc	-.0034052	.0015249	-2.23	0.026	-.0063988	-.0004115
educ	.0379953	.007266	5.23	0.000	.023731	.0522596
exper	.0394924	.00581	6.80	0.000	.0280864	.0508983
expersq	-.0005963	.00019	-3.14	0.002	-.0009693	-.0002233
age	-.0160908	.002399	-6.71	0.000	-.0208004	-.0113812
kidslt6	-.2618105	.0317832	-8.24	0.000	-.3242058	-.1994152
kidsge6	.0130122	.0135329	0.96	0.337	-.013555	.0395795
_cons	.5855192	.1522599	3.85	0.000	.2866098	.8844287

- One problem of the linear probability model is that, the predicted probability can be greater than 1 or smaller than 0.
- This problem can be avoided by using Probit or Logit models which are described below.
- Nonetheless, you should note that, in many applications, the predicted probabilities fall mostly within the [0,1] interval. Thus, it is still a reasonable model, and it has been applied in many papers.

The probit and logit models

- Consider that you would like to know the determinants of the labor force participations of married women.

- For simplicity, let us consider the model with only one explanatory variable.

- In probit and logit model, we assume that there is an unobserved variable (latent variable) y* that depends on x:

$$y^*=\beta_0+\beta_1 x+u$$

- You can consider y* as the utility from participating in the labor force.
- We do not observe y*. We only observe the actual labor force participation outcome.

Y=0 if the person is not working

Y=1 if the person is working.

- Then, we assume the following.

If Y=0 (the person is not working), then y* must have been negative.

If Y=1 (the person is working), then y* must have been positive.

- This assumption can be written as:

$$\begin{cases} \text{If } Y = 0, \text{ then } y^* < 0 \\ \text{If } Y = 1, \text{ then } y^* \geq 0 \end{cases}$$

- Intuitive interpretation is the following. If the utility from participating in the labor force is negative, the person will not work. But if the utility from participating in the labor force is positive, then the person will work.

- Given the data about x and Y, we can compute the likelihood contribution for each person in the data by making the distributional assumption about the error term u.

The model:

$$y_i^* = \beta_0 + \beta_1 x_i + u_i$$

$$\begin{cases} \text{If } Y_i = 0, \text{ then } y_i^* < 0 \\ \text{If } Y_i = 1, \text{ then } y_i^* \geq 0 \end{cases}$$

The i-subscript denotes the observation id.

- When we assume that u_i follows the standard normal distribution (normal with mean 0 and variance 1), the model is called the **Probit model**.

- When we assume that u follows the logistic distribution, the model is called the **Logit model**.

Probit model: Likelihood function example

The model

$$y_i^* = \beta_0 + \beta_1 x_i + u_i$$

$$\begin{cases} \text{If } Y_i = 0, \text{ then } y_i^* < 0 \\ \text{If } Y_i = 1, \text{ then } y_i^* \geq 0 \end{cases}$$

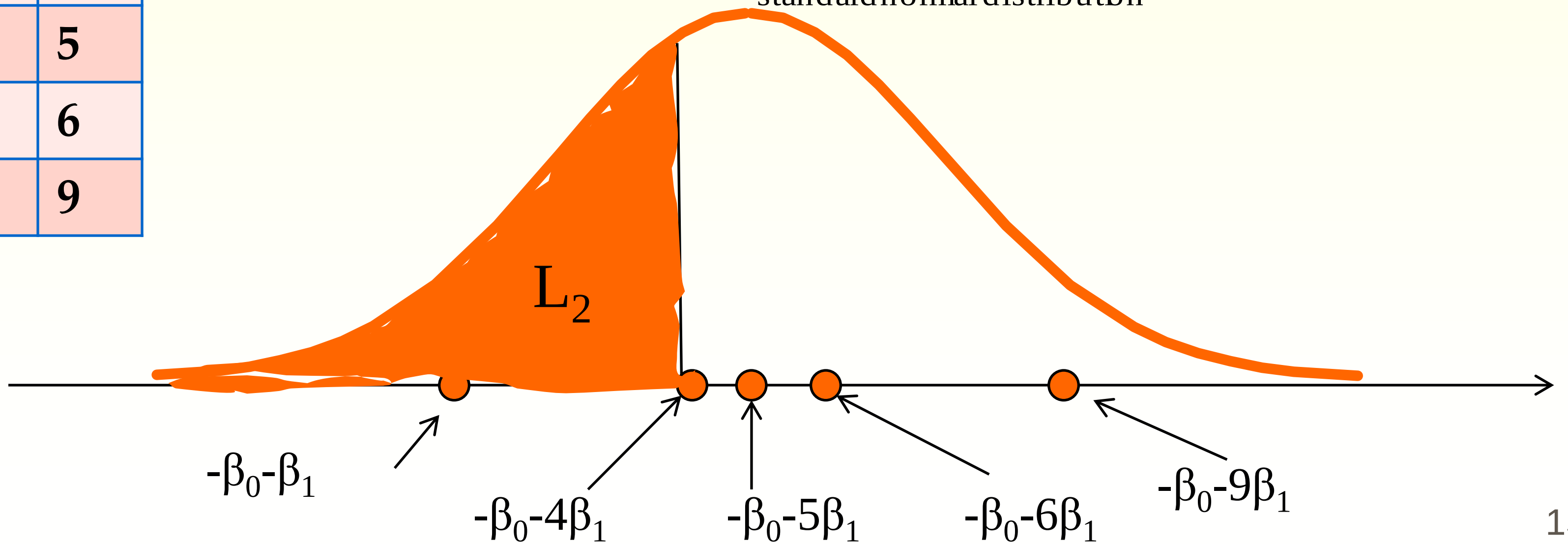
Id	Y	X
1	0	1
2	0	4
3	1	5
4	1	6
5	1	9

- Suppose that you have the following data.
- We assume that $u_i \sim N(0,1)$

- Take 2nd observation as an example. Since Y=0 for this observation, we know y*<0
- Thus, the likelihood contribution is

$$L_2 = P(y^* < 0) = P(\beta_0 + 4\beta_1 + u < 0)$$
$$= P(u < -\beta_0 - 4\beta_1)$$
$$= \int_{-\infty}^{-\beta_0 - 4\beta_1} \frac{1}{\sqrt{2\pi}} e^{-u^2} du = \underbrace{\Phi(-\beta_0 - 4\beta_1)}_{\text{Cumulative Distribution Function of standard normal distribution}}$$

Id	Y	X
1	0	1
2	0	4
3	1	5
4	1	6
5	1	9



- Now, take 3nd observation as an example. Since Y=0 for this observation, we know y*≥0
- Thus, the likelihood contribution is

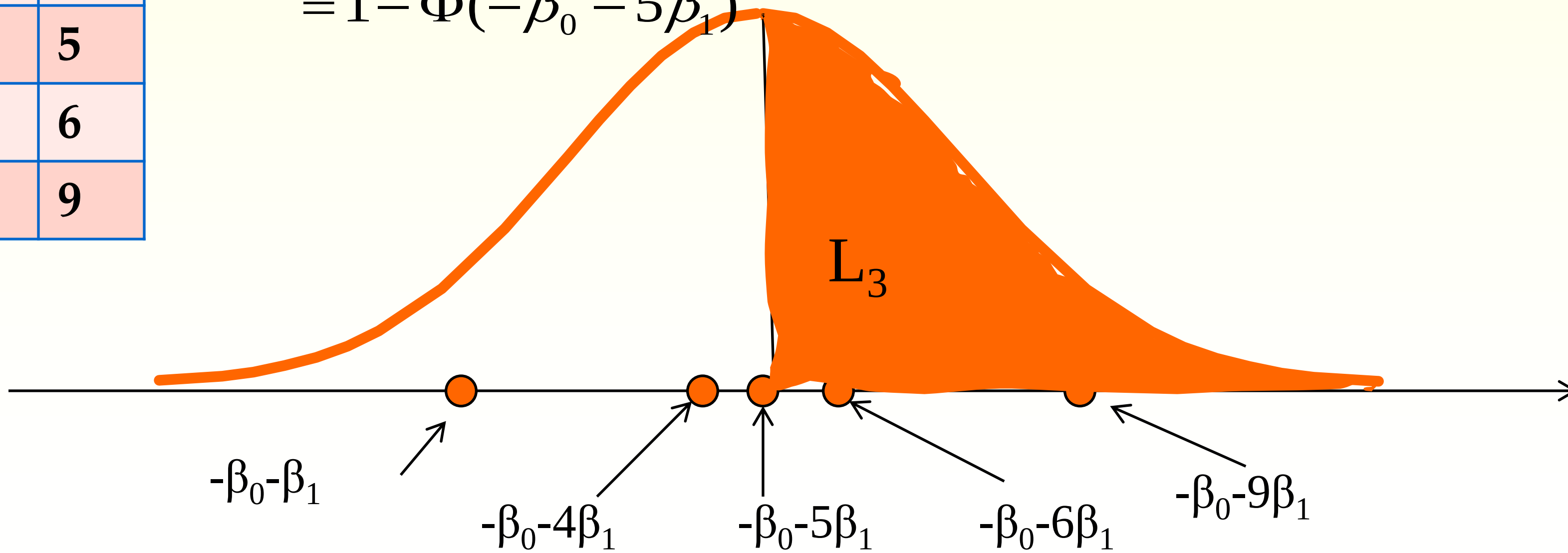
$$L_3 = P(y^* \geq 0) = P(\beta_0 + 5\beta_1 + u \geq 0)$$

$$= P(u \geq -\beta_0 - 5\beta_1)$$

$$= \int_{-\beta_0-5\beta_1}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-u^2} du = 1 - \int_{\infty}^{-\beta_0-5\beta_1} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-u^2} du$$

$$= 1 - \Phi(-\beta_0 - 5\beta_1)$$

Id	Y	X
1	0	1
2	0	4
3	1	5
4	1	6
5	1	9



Then the likelihood function is given by

Id	Y	X
1	**0**	**1**
2	**0**	**4**
3	**1**	**5**
4	**1**	**6**
5	**1**	**9**

$$L(\beta_0, \beta_1) = \prod_{i=1}^{5} L_i$$

$$= \left[\int_{\infty}^{-\beta_0 - \beta} \frac{1}{\sqrt{2\pi}} e^{-u^2} du \right] \times \left[\int_{\infty}^{-\beta_0 - 4\beta} \frac{1}{\sqrt{2\pi}} e^{-u^2} du \right]$$

$$\times \left[1 - \int_{\infty}^{-\beta_0 - 5\beta} \frac{1}{\sqrt{2\pi}} e^{-u^2} du \right] \times \left[1 - \int_{\infty}^{-\beta_0 - 6\beta} \frac{1}{\sqrt{2\pi}} e^{-u^2} du \right]$$

$$\times \left[1 - \int_{\infty}^{-\beta_0 - 9\beta} \frac{1}{\sqrt{2\pi}} e^{-u^2} du \right]$$

$$= \Phi(-\beta_0 - \beta)\Phi(-\beta_0 - 4\beta)\left[1 - \Phi(-\beta_0 - 5\beta)\right]$$

$$\times \left[1 - \Phi(-\beta_0 - 6\beta)\right]\left[1 - \Phi(-\beta_0 - 9\beta)\right]$$

Probit model
Likelihood for more general case

Id	Y	X
1	Y_1	x_1
2	Y_2	x_2
3	Y_3	x_3
:	:	:
n	Y_n	x_n

- Consider the following model.

$$y_i^* = \beta_0 + \beta_1 x_i + u_i$$

$$\begin{cases} \text{If } Y_i = 0, \text{ then } y_i^* < 0 \\ \text{If } Y_i = 1, \text{ then } y_i^* \geq 0 \end{cases}$$

$$u_i \sim N(0,1)$$

- Then, we know that

$$\begin{cases} L_i = \Phi(-\beta_0 - \beta_1 x_i) \text{ if } \mathrm{Y_i} < 0 \\ L_i = 1 - \Phi(-\beta_0 - \beta_1 x_i) \text{ if } \mathrm{Y_i} \geq 0 \end{cases}$$

- The above can be conveniently written as:

$$L_i = \left[\Phi(-\beta_0 - \beta_1 x_i)\right]^{(1-\mathrm{Y_i})} \left[1 - \Phi(-\beta_0 - \beta_1 x_i)\right]^{Y_i}$$

- Since normal distribution is symmetric, we have $\Phi(-\beta_0 - \beta_1 x_i) = 1 - \Phi(\beta_0 + \beta_1 x_i)$. Thus, we have

$$L_i = \left[1 - \Phi(\beta_0 + \beta_1 x_i)\right]^{(1-\mathrm{Y_i})} \left[\Phi(\beta_0 + \beta_1 x_i)\right]^{Y_i}$$

- Thus, the likelihood function is given by

$$L = \prod_{i=1}^{n} L_i = \prod_{i=1}^{n} \left[1 - \Phi(\beta_0 + \beta_1 x_i)\right]^{(1-\mathrm{Y_i})} \left[\Phi(\beta_0 + \beta_1 x_i)\right]^{Y_i}$$

- Usually, you maximize Log(L).
- The values of the parameters that maximize Log(L) are the estimators of the probit model.
- The MLE is done automatically by STATA.

The Logit Model
The likelihood function example

- Consider the following model

  $y_i{*}=\beta_0+\beta_1 x_i+u_i$

  $$\begin{cases} \text{If } Y_i = 0, \text{ then } y_i^* < 0 \\ \text{If } Y_i = 1, \text{ then } y_i^* \geq 0 \end{cases}$$

- In Logit model, we assume that u_i follows the logistic distribution with mean 0 and variance 1.

- The **density function** of the logistic distribution with mean 0 and variance 1 is given by the following.

$$\text{Density Function}: \mathrm{f(x)} = \frac{e^{-x}}{1-e^{-x}}$$

- The **cumulative distribution function** of the logistic distribution with mean 0 and variance 1 has the ‘closed form’.

$$\text{Cumulative Distribution Function}: \mathrm{F(x)} = \frac{1}{1+e^{-x}}$$

Id	Y	X
1	0	1
2	0	4
3	1	5
4	1	6
5	1	9

- Now, suppose that you have the following data.
- Take the 2nd observation as an example. Since Y_2=0, it must have been the case that y_2*<0.
- Thus, the likelihood contribution is:

$$
\begin{aligned}
L_2 &= P(y_2^* < 0) = P(\beta_0 + 4\beta_1 + u < 0) \\
&= P(u < -\beta_0 - 4\beta_1) \\
&= \int_{-\infty}^{-\beta_0 - 4\beta_1} \frac{e^{-u}}{1 - e^{-u}} = \frac{1}{1 + e^{-(-\beta_0 - 4\beta_1)}} = \frac{1}{1 + e^{\beta_0 + 4\beta_1}}
\end{aligned}
$$

Id	Y	X
1	0	1
2	0	4
3	1	5
4	1	6
5	1	9

- Now, take the 3rd observation as an example. Since Y_3=0, it must have been the case that y_3*<0.
- Thus, the likelihood contribution is:

$$
\begin{aligned}
L_3 &= P(y_3^* \geq 0) = P(\beta_0 + 5\beta_1 + u \geq 0) \\
&= P(u \geq -\beta_0 - 5\beta_1) \\
&= \int_{-\beta_0-5\beta_1}^{\infty} \frac{e^{-u}}{1-e^{-u}}\, du = 1 - \int_{\infty}^{-\beta_0-5\beta_1} \frac{e^{-u}}{1-e^{-u}}\, du \\
&= 1 - \frac{1}{1+e^{-(-\beta_0-5\beta_1)}} = 1 - \frac{1}{1+e^{\beta_0+5\beta_1}} \\
&= \frac{e^{\beta_0+5\beta_1}}{1+e^{\beta_0+5\beta_1}}
\end{aligned}
$$

- Thus the likelihood function for the data set is given by

Id	Y	X
1	**0**	**1**
2	**0**	**4**
3	**1**	**5**
4	**1**	**6**
5	**1**	**9**

$$L=\prod_{i=1}^{5} L_i=\frac{1}{1+e^{\beta_0+\beta_1}}\times\frac{1}{1+e^{\beta_0+4\beta_1}}$$

$$\times\frac{e^{\beta_0+5\beta_1}}{1+e^{\beta_0+5\beta_1}}\times\frac{e^{\beta_0+6\beta_1}}{1+e^{\beta_0+6\beta_1}}\times\frac{e^{\beta_0+9\beta_1}}{1+e^{\beta_0+9\beta_1}}$$

Logit model
Likelihood for more general case

Id	Y	X
1	Y_1	x_1
2	Y_2	x_2
3	Y_3	x_3
:	:	:
n	Y_n	x_n

- Consider the following model.

$$y_i^*=\beta_0+\beta_1 x_i+u_i$$

$$\begin{cases} \text{If } Y_i = 0, \text{ then } y_i^* < 0 \\ \text{If } Y_i = 1, \text{ then } y_i^* \geq 0 \end{cases}$$

We assume u_i follows the logistic distribution.

- Then, we know that

$$\begin{cases} L_i = \dfrac{1}{1+e^{\beta_0+\beta_1 x_i}} \text{ if } \mathrm{Y_i} = 0 \\ L_i = \dfrac{e^{\beta_0+\beta_1 x_i}}{1+e^{\beta_0+\beta_1 x_i}} \text{ if } \mathrm{Y_i} = 1 \end{cases}$$

- The above can be conveniently written as:

$$L_i = \left[\frac{1}{1+e^{\beta_0+\beta_1 x_i}}\right]^{(1-\mathrm{Y_i})}\left[\frac{e^{\beta_0+\beta_1 x_i}}{1+e^{\beta_0+\beta_1 x_i}}\right]^{Y_i}$$

- Thus, the likelihood function is given by

$$L = \prod_{i=1}^{n} L_i = \prod_{i=1}^{n}\left[\frac{1}{1+e^{\beta_0+\beta_1 x_i}}\right]^{(1-\mathrm{Y_i})}\left[\frac{e^{\beta_0+\beta_1 x_i}}{1+e^{\beta_0+\beta_1 x_i}}\right]^{Y_i}$$

- Usually, you maximize log(L).

- The values of the parameters that maximize log(L) is the estimators of the Logit model.

Exercise

- Consider the following model.

  $y^*=\beta_0+\beta_1 nwifinc+\beta_2 educ+\beta_3 exper$
  $+\beta_4 exper^2+\beta_5 age+\beta_6 kidslt6+\beta_7 kidsge6$
  $+u$

  $inlf=0$ if $y^*<0$ (i.e, not work if $y^*<0$)

  $inlf=1$ if $y^*\geq 0$ (i.e., work if $y^*\geq 0$)

- Using Mroz.dta, estimate the parameters using both Probit and Logit models.

Probit

```
. probit inlf  nwifeinc educ exper expersq age  kidslt6 kidsge6, robust

Iteration 0:   log pseudolikelihood =  -514.8732
Iteration 1:   log pseudolikelihood = -405.78215
Iteration 2:   log pseudolikelihood = -401.32924
Iteration 3:   log pseudolikelihood = -401.30219
Iteration 4:   log pseudolikelihood = -401.30219

Probit regression                                 Number of obs   =        753
                                                  Wald chi2(7)    =     185.10
                                                  Prob > chi2     =     0.0000
Log pseudolikelihood = -401.30219                 Pseudo R2       =     0.2206
```

inlf	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
nwifeinc	-.0120237	.0053106	-2.26	0.024	-.0224323	-.0016152
educ	.1309047	.0258192	5.07	0.000	.0803	.1815095
exper	.1233476	.0188537	6.54	0.000	.086395	.1603002
expersq	-.0018871	.0006007	-3.14	0.002	-.0030645	-.0007097
age	-.0528527	.0083532	-6.33	0.000	-.0692246	-.0364807
kidslt6	-.8683285	.1162037	-7.47	0.000	-1.096084	-.6405735
kidsge6	.036005	.0452958	0.79	0.427	-.0527731	.124783
_cons	.2700768	.505175	0.53	0.593	-.7200481	1.260202

Logit

```
. logit inlf  nwifeinc educ exper expersq age  kidslt6 kidsge6, robust

Iteration 0:   log pseudolikelihood =  -514.8732
Iteration 1:   log pseudolikelihood = -406.94123
Iteration 2:   log pseudolikelihood = -401.85151
Iteration 3:   log pseudolikelihood = -401.76519
Iteration 4:   log pseudolikelihood = -401.76515

Logistic regression                               Number of obs   =        753
                                                  Wald chi2(7)    =     158.48
                                                  Prob > chi2     =     0.0000
Log pseudolikelihood = -401.76515                 Pseudo R2       =     0.2197

------------------------------------------------------------------------------
             |               Robust
        inlf |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    nwifeinc |  -.0213452   .0090781    -2.35   0.019     -.039138   -.0035523
        educ |   .2211704   .0444509     4.98   0.000     .1340483    .3082925
       exper |   .2058695   .0322914     6.38   0.000     .1425796    .2691594
     expersq |  -.0031541   .0010124    -3.12   0.002    -.0051384   -.0011698
         age |  -.0880244   .0144393    -6.10   0.000    -.1163248    -.059724
     kidslt6 |  -1.443354   .2031615    -7.10   0.000    -1.841543   -1.045165
     kidsge6 |   .0601122   .0798825     0.75   0.452    -.0964546     .216679
       _cons |   .4254524   .8597307     0.49   0.621    -1.259589    2.110494
------------------------------------------------------------------------------
```

Interpreting the parameters

- Consider the following model.

$$y_i^* = \beta_0 + \beta_1 x_i + u_i$$

$$\begin{cases} \text{If } Y_i = 0, \text{ then } y_i^* < 0 \\ \text{If } Y_i = 1, \text{ then } y_i^* \geq 0 \end{cases}$$

- In probit and Logit models, interpretation of the parameters is not straightforward.
- For the explanation purpose, consider that this model is estimating the determinants of the labor force participation.

Probit case

- Note that

$$P(y_i^* \geq 0) = 1 - \Phi(-\beta_0 - \beta_1 x_i) = \Phi(\beta_0 + \beta_1 x_i)$$

=the probability that the person participates in the labor force.

- Therefore, if β_1 is positive, an increase in x would increase the probability that the person participates in the labor force.
- If β_1 is negative, then an increase in x will decrease the probability that the person participates in the labor force.

Example

```
. probit inlf  nwifeinc educ exper expersq age  kidslt6 kidsge6, robust

Iteration 0:   log pseudolikelihood =  -514.8732
Iteration 1:   log pseudolikelihood = -405.78215
Iteration 2:   log pseudolikelihood = -401.32924
Iteration 3:   log pseudolikelihood = -401.30219
Iteration 4:   log pseudolikelihood = -401.30219

Probit regression                                 Number of obs   =        753
                                                  Wald chi2(7)    =     185.10
                                                  Prob > chi2     =     0.0000
Log pseudolikelihood = -401.30219                 Pseudo R2       =     0.2206
```

inlf	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
nwifeinc	-.0120237	.0053106	-2.26	0.024	-.0224323	-.0016152
educ	.1309047	.0258192	5.07	0.000	.0803	.1815095
exper	.1233476	.0188537	6.54	0.000	.086395	.1603002
expersq	-.0018871	.0006007	-3.14	0.002	-.0030645	-.0007097
age	-.0528527	.0083532	-6.33	0.000	-.0692246	-.0364807
kidslt6	-.8683285	.1162037	-7.47	0.000	-1.096084	-.6405735
kidsge6	.036005	.0452958	0.79	0.427	-.0527731	.124783
_cons	.2700768	.505175	0.53	0.593	-.7200481	1.260202

Increase in non wife income will decrease the probability that the woman works.

Increase in education will increase the probability that the woman works.

Increase in the number of kids who are younger than 6 will decrease the probability that the woman works.

The Partial Effects of Probit model (Continuous variable case)

- The sign of the parameters can tell you if the probability of the labor force participation will increase or decrease.
- We also want to know “by how much” the probability increases or decreases.
- This can be done by computing the **partial effects** (sometimes called the **marginal effects**).

- The increase in the probability due to an increase in x-variable by one unit can be computed by taking the derivative of $\Phi(\beta_0+\beta_1 x_i)$ with respect to x_i.

Partial effect(Marginal effect)=The increase in the probability that the person works due to one unit increase in xi

$$=\frac{\partial\Phi(\beta_0+\beta_1 x_i)}{\partial x_i}=\underbrace{\phi(\beta_0+\beta_1 x_i)}\beta_1$$

Partial effect

Note, this is the **density function,** not the cumulative distribution function.

- As you can see, the partial effect depends on the value of the x-variable. Therefore, it is different for different person in the data.

- However, we want to know the overall effect of x on the probability.
- There are two ways to do so.

1. Partial effect at average
2. The average partial effect

- Partial effect at average

$$PEA = \phi(\hat{\beta}_0 + \hat{\beta}_1 \bar{x})\hat{\beta}_1$$

Stata computes this automatically. This is called the marginal effect in Stata,

This is the partial effect evaluated at the average value of x.

- Average partial effect (APE)

This is not done automatically, but easy to do manually.

$$APE = \frac{1}{n}\sum_{i=1}^{n}\phi(\hat{\beta}_0 + \hat{\beta}_1 x_i)\hat{\beta}_1$$

You compute the partial effect for each person, then take the average.

The partial effect of the probit model (Discrete variable case)

- Consider the following labor force participation model

$$y^* = \beta_0 + \beta_1 x_i + \beta_2 D_i + u_i$$

$$\begin{cases} \text{If } Y_i = 0, \text{ then } y_i^* < 0 \\ \text{If } Y_i = 1, \text{ then } y_i^* \geq 0 \end{cases}$$

where D_i is a dummy variable that is 1 if the person lives with parents, and 0 otherwise.

- You want to know the effect of living with parents on the probability that the woman participates in the labor force.
- This is a dummy variable. In such a case, the partial effect formula described before is not a good approximation.
- For a dummy variable case, a better way to compute the partial effect is given in the next slides.

- The partial effect at average for discrete case:

$$\text{Partial effect at average} = \underbrace{\Phi(\hat{\beta}_0 + \hat{\beta}_1\bar{x} + \hat{\beta}_2 \times 1)}_{\substack{\text{Participation probability} \\ \text{when living with parents (at average)}}} - \underbrace{\Phi(\hat{\beta}_0 + \hat{\beta}_1\bar{x} + \hat{\beta}_2 \times 0)}_{\substack{\text{Participation probability} \\ \text{when not living with parents (at average)}}}$$

- This is computed automatically by STATA.

The average partial effect is computed as:

$$(\text{Partial effect})_{\text{i}} = \underbrace{\Phi(\hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\beta}_2 \times 1)}_{\substack{\text{Participation probability} \\ \text{when living with parents}}} - \underbrace{\Phi(\hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\beta}_2 \times 0)}_{\substack{\text{Participation probability} \\ \text{when not living with parents}}}$$

Then, the average partial effect is computed as the sample average of the partial effect.

$$\text{Average partial effect} = \frac{1}{n}\sum_{\text{i}=1}^{\text{n}} (\text{Partial effect})_{\text{i}}$$

Exercise

- Using Mroz.dta, estimate the following model.

$$y^*=\beta_0+\beta_2 educ+\beta_3 exper+u$$

inlf=0 if $y^*<0$ (i.e, not work if $y^*<0$)

inlf=1 if $y^*\geq 0$ (i.e., work if $y^*\geq 0$)

$u \sim N(0,1)$

Compute the effect of education on the labor force participation. Compute both the partial effect at average and the average partial effect.

```
. probit inlf educ exper

Iteration 0:   log likelihood =  -514.8732
Iteration 1:   log likelihood = -456.47099
Iteration 2:   log likelihood = -455.55276
Iteration 3:   log likelihood = -455.55204

Probit regression                                 Number of obs   =        753
                                                  LR chi2(2)      =     118.64
                                                  Prob > chi2     =     0.0000
Log likelihood = -455.55204                       Pseudo R2       =     0.1152

------------------------------------------------------------------------------
        inlf |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |    .104287   .0219681     4.75   0.000     .0612303    .1473437
       exper |   .0597903   .0065816     9.08   0.000     .0468907      .07269
       _cons |  -1.713151    .280658    -6.10   0.000     -2.26323   -1.163071
------------------------------------------------------------------------------

. egen aveduc=mean(educ)

. egen avexper=mean(exper)

. gen avxbeta=_b[_cons ]+_b[educ ]*aveduc+_b[exper]*avexper

. gen partial_av=normalden(avxbeta)*_b[educ]

. su partial_av

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
  partial_av |       753    .0407492           0   .0407492   .0407492
```

Computing partial effect at average, manually.

Partial effect

```
. probit inlf educ exper

Iteration 0:   log likelihood =  -514.8732
Iteration 1:   log likelihood = -456.47099
Iteration 2:   log likelihood = -455.55276
Iteration 3:   log likelihood = -455.55204

Probit regression                                 Number of obs   =        753
                                                  LR chi2(2)      =     118.64
                                                  Prob > chi2     =     0.0000
Log likelihood = -455.55204                       Pseudo R2       =     0.1152

------------------------------------------------------------------------------
        inlf |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |    .104287   .0219681     4.75   0.000     .0612303    .1473437
       exper |   .0597903   .0065816     9.08   0.000     .0468907      .07269
       _cons |  -1.713151    .280658    -6.10   0.000    -2.26323   -1.163071
------------------------------------------------------------------------------

. mfx, varlist(educ)

Marginal effects after probit
      y  = Pr(inlf) (predict)
         =  .58075593
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
    educ |   .0407492      .00858    4.75   0.000   .023932  .057567   12.2869
------------------------------------------------------------------------------
```

Computing partial effect at average automatically.

```
. probit inlf educ exper, robust

Iteration 0:   log pseudolikelihood =  -514.8732
Iteration 1:   log pseudolikelihood = -456.47099
Iteration 2:   log pseudolikelihood = -455.55276
Iteration 3:   log pseudolikelihood = -455.55204

Probit regression                                 Number of obs   =        753
                                                  Wald chi2(2)    =      80.67
                                                  Prob > chi2     =     0.0000
Log pseudolikelihood = -455.55204                 Pseudo R2       =     0.1152

------------------------------------------------------------------------------
             |               Robust
        inlf |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |    .104287   .0217204     4.80   0.000     .0617157    .1468583
       exper |   .0597903   .0078708     7.60   0.000     .0443638    .0752169
       _cons |  -1.713151   .2802903    -6.11   0.000     -2.26251   -1.163792
------------------------------------------------------------------------------

. predict xbeta, xb

. gen partial =normalden(xbeta)*_b[educ]

. su partial

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
     partial |       753     .036153    .0072086   .0043549   .0416045
```

Computing average partial effect, "manually".

Exercise

- Use JPSC1.dta to estimate the following model.

  $y^*=\beta_0+\beta_2 exper+\beta_3(livetogether)+u$

  Work=0 if $y^*<0$ (i.e, not work if $y^*<0$)

  Work=1 if $y^*\geq 0$ (i.e., work if $y^*\geq 0$)

  $u \sim N(0,1)$

  Livetogehter is a dummy variable for those who are living with parents.

Q1. Estimate the effect of living with parents on the labor force participation.

```
. probit work exp livetogether

Iteration 0:   log likelihood = -5052.9067
Iteration 1:   log likelihood = -3612.6274
Iteration 2:   log likelihood = -3517.5372
Iteration 3:   log likelihood = -3515.8581
Iteration 4:   log likelihood = -3515.8574

Probit regression                                 Number of obs   =       7537
                                                  LR chi2(2)      =    3074.10
                                                  Prob > chi2     =     0.0000
Log likelihood = -3515.8574                       Pseudo R2       =     0.3042

------------------------------------------------------------------------------
        work |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         exp |   .1967633   .0043554    45.18   0.000     .1882269    .2052996
livetogether |   .2209021   .0363567     6.08   0.000     .1496442      .29216
       _cons |  -2.225386   .0454465   -48.97   0.000    -2.314459   -2.136313
------------------------------------------------------------------------------

. egen avexp=mean(exp)

. gen xbeta1=_b[_cons ]+_b[exp ]*avexp+_b[livetogether]*1

. gen xbeta0=_b[_cons ]+_b[exp ]*avexp+_b[livetogether]*0

. gen p1=normal(xbeta1)

. gen p0=normal(xbeta0)

. gen partial_av=p1-p0
. su  partial_av

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
  partial_av |      7537    .0844965           0   .0844965   .0844965
```

Computing partial effect at average, “manually”.

Partial effect at average

```
. probit work exp livetogether

Iteration 0:   log likelihood = -5052.9067
Iteration 1:   log likelihood = -3612.6274
Iteration 2:   log likelihood = -3517.5372
Iteration 3:   log likelihood = -3515.8581
Iteration 4:   log likelihood = -3515.8574

Probit regression                                 Number of obs   =       7537
                                                  LR chi2(2)      =    3074.10
                                                  Prob > chi2     =     0.0000
Log likelihood = -3515.8574                       Pseudo R2       =     0.3042

------------------------------------------------------------------------------
        work |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         exp |   .1967633   .0043554    45.18   0.000     .1882269    .2052996
livetogether |   .2209021   .0363567     6.08   0.000     .1496442      .29216
       _cons |  -2.225386   .0454465   -48.97   0.000    -2.314459   -2.136313
------------------------------------------------------------------------------

. mfx, varlist(livetogether)

Marginal effects after probit
      y  = Pr(work) (predict)
         =  .37468279
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
liveto~r*|   .0844965      .01401    6.03   0.000   .057046  .111947   .337402
------------------------------------------------------------------------------
(*) dy/dx is for discrete change of dummy variable from 0 to 1
```

Computing the partial effect at average automatically.

```
. probit work exp livetogether

Iteration 0:   log likelihood = -5052.9067
Iteration 1:   log likelihood = -3612.6274
Iteration 2:   log likelihood = -3517.5372
Iteration 3:   log likelihood = -3515.8581
Iteration 4:   log likelihood = -3515.8574

Probit regression                                 Number of obs   =       7537
                                                  LR chi2(2)      =    3074.10
                                                  Prob > chi2     =     0.0000
Log likelihood = -3515.8574                       Pseudo R2       =     0.3042

------------------------------------------------------------------------------
        work |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         exp |   .1967633   .0043554    45.18   0.000     .1882269    .2052996
livetogether |   .2209021   .0363567     6.08   0.000     .1496442      .29216
       _cons |  -2.225386   .0454465   -48.97   0.000    -2.314459   -2.136313
------------------------------------------------------------------------------

. gen xb1=_b[_cons]+_b[exp]*exp+_b[livetogether ]*1

. gen xb0=_b[_cons]+_b[exp]*exp+_b[livetogether ]*0

. gen pp1=normal(xb1)

. gen pp0=normal(xb0)

. gen partial2=pp1-pp0

. browse

. su partial2

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
    partial2 |      7537    .0594117    .0239495   .0000651   .0879483
```

Computing the average partial effect "manually".