Lecture 6: Binary Dependent Variable

Big Data Analytics,Spring 2019

Zhaopeng Qu

Nanjing University

4/20/2019

1. Introduction to limited dependent variable
2. The Linear Probability Model(LPM)
3. Nonlinear probability model

Introduction to limited dependent variable

Introduction

- So far the dependent variable (Y) has been continuous:

Introduction

- So far the dependent variable (Y) has been continuous:
  - testscore

Introduction

- So far the dependent variable (Y) has been continuous:
  - testscore
  - average hourly earnings

Introduction

- So far the dependent variable (Y) has been continuous:
  - testscore
  - average hourly earnings
  - GDP growth rate

Introduction

- So far the dependent variable (Y) has been continuous:
  - testscore
  - average hourly earnings
  - GDP growth rate
- What if Y is discrete?

Introduction

- So far the dependent variable (Y) has been continuous:
  - testscore
  - average hourly earnings
  - GDP growth rate
- What if Y is discrete?
  - Y= get into college, or not; X = parental income.

Introduction

- So far the dependent variable (Y) has been continuous:
  - testscore
  - average hourly earnings
  - GDP growth rate
- What if Y is discrete?
  - $Y=$ get into college, or not; $X =$ parental income.
  - $Y=$ person smokes, or not; $X =$ cigarette tax rate, income.

Introduction

- So far the dependent variable (Y) has been continuous:
  - testscore
  - average hourly earnings
  - GDP growth rate
- What if Y is discrete?
  - Y= get into college, or not; X = parental income.
  - Y= person smokes, or not; X = cigarette tax rate, income.
  - Y= mortgage application is accepted, or not; X = race, income, house characteristics, marital status ...

Introduction

- So far the dependent variable (Y) has been continuous:
  - testscore
  - average hourly earnings
  - GDP growth rate
- What if Y is discrete?
  - Y= get into college, or not; X = parental income.
  - Y= person smokes, or not; X = cigarette tax rate, income.
  - Y= mortgage application is accepted, or not; X = race, income, house characteristics, marital status ...
- Binary outcomes models:

Introduction

- So far the dependent variable (Y) has been continuous:
  - testscore
  - average hourly earnings
  - GDP growth rate
- What if Y is discrete?
  - Y= get into college, or not; $X =$ parental income.
  - Y= person smokes, or not; $X =$ cigarette tax rate, income.
  - Y= mortgage application is accepted, or not; $X =$ race, income, house characteristics, marital status ...
- Binary outcomes models:
  - Logit Probability Model(LPM)

Introduction

- So far the dependent variable (Y) has been continuous:
  - testscore
  - average hourly earnings
  - GDP growth rate
- What if Y is discrete?
  - Y= get into college, or not; X = parental income.
  - Y= person smokes, or not; X = cigarette tax rate, income.
  - Y= mortgage application is accepted, or not; X = race, income, house characteristics, marital status ...
- Binary outcomes models:
  - Logit Probability Model(LPM)
  - Logit

Introduction

- So far the dependent variable (Y) has been continuous:
  - testscore
  - average hourly earnings
  - GDP growth rate
- What if Y is discrete?
  - Y= get into college, or not; X = parental income.
  - Y= person smokes, or not; X = cigarette tax rate, income.
  - Y= mortgage application is accepted, or not; X = race, income, house characteristics, marital status ...
- Binary outcomes models:
  - Logit Probability Model(LPM)
  - Logit
  - Probit

The Linear Probability Model(LPM)

The linear probability model

- If a outcome variable is **binary**,then the expecation of it is

$$E[Y] = 1 \times Pr(Y = 1) + 0 \times Pr(Y = 0) = Pr(Y = 1)$$

$$E[Y|X_{1i}, ..., X_{ki}] = Pr(Y = 1|X_{1i}, ..., X_{ki})$$

The linear probability model

- If a outcome variable is **binary**,then the expecation of it is

$$E[Y] = 1 \times Pr(Y = 1) + 0 \times Pr(Y = 0) = Pr(Y = 1)$$

- Then we have the probability of Y conditional on X

$$E[Y|X_{1i}, ..., X_{ki}] = Pr(Y = 1|X_{1i}, ..., X_{ki})$$

The linear probability model

- The conditional expectation equals the probability that $Y_i = 1$ conditional on $X_{1i}, ..., X_{ki}$:

$$E[Y|X_{1i}, ..., X_{ki}] = Pr(Y = 1|X_{1i}, ..., X_{ki}) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + ... + \beta_k X_{ki}$$

The linear probability model

- The conditional expectation equals the probability that $Y_i = 1$ conditional on $X_{1i}, ..., X_{ki}$:

$$E[Y|X_{1i}, ..., X_{ki}] = Pr(Y = 1|X_{1i}, ..., X_{ki}) = \beta_0+\beta_1 X_{1i}+\beta_2 X_{2i}+...+\beta_k X_{ki}$$

- The population coefficient β_j equals the change in the probability that $Y_i = 1$ associated with a unit change in X_j.

$$\frac{\partial Pr(Y_i = 1|X_{1i}, ..., X_{ki})}{\partial X_j} = \beta_j$$

The linear probability model

- The conditional expectation equals the probability that $Y_i = 1$ conditional on $X_{1i}, ..., X_{ki}$:

$$E[Y|X_{1i}, ..., X_{ki}] = Pr(Y = 1|X_{1i}, ..., X_{ki}) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + ... + \beta_k X_{ki}$$

- The population coefficient β_j equals the change in the probability that $Y_i = 1$ associated with a unit change in X_j.

$$\frac{\partial Pr(Y_i = 1|X_{1i}, ..., X_{ki})}{\partial X_j} = \beta_j$$

- β_j is *the change in the probability that* $Y = 1$ *associated with a unit change in* X_j

The linear probability model

- Almost all of the tools of Multiple OLS regression can carry over to the LPM model.

The linear probability model

- Almost all of the tools of Multiple OLS regression can carry over to the LPM model.
  - Assumptions are the same as for general multiple regression model.

The linear probability model

- Almost all of the tools of Multiple OLS regression can carry over to the LPM model.
  - Assumptions are the same as for general multiple regression model.
  - The coefficients can be estimated by OLS.

The linear probability model

- Almost all of the tools of Multiple OLS regression can carry over to the LPM model.
  - Assumptions are the same as for general multiple regression model.
  - The coefficients can be estimated by OLS.
  - t-statistic and F-statistic can be constructed as before.

The linear probability model

- Almost all of the tools of Multiple OLS regression can carry over to the LPM model.
  - Assumptions are the same as for general multiple regression model.
  - The coefficients can be estimated by OLS.
  - t-statistic and F-statistic can be constructed as before.
  - the errors of the LPM are **always heteroskedastic**, so it is essential that heteroskedasticity-robust s.e. be used for inference.

The linear probability model

- Almost all of the tools of Multiple OLS regression can carry over to the LPM model.
  - Assumptions are the same as for general multiple regression model.
  - The coefficients can be estimated by OLS.
  - t-statistic and F-statistic can be constructed as before.
  - the errors of the LPM are **always heteroskedastic**, so it is essential that heteroskedasticity-robust s.e. be used for inference.
  - R^2 is not a useful statistic now.

The linear probability model

- Advantages of the linear probability model:

The linear probability model

- Advantages of the linear probability model:
  - easy to estimate

The linear probability model

- Advantages of the linear probability model:
  - easy to estimate
  - Coeffcient estimates are easy to interpret

The linear probability model

- Advantages of the linear probability model:
  - easy to estimate
  - Coeffcient estimates are easy to interpret
- Disadvantages of the linear probability model

The linear probability model

- Advantages of the linear probability model:
  - easy to estimate
  - Coeffcient estimates are easy to interpret
- Disadvantages of the linear probability model
  - Predicted probability can be above 1 or below 0!(it doesn't make sense)

The linear probability model

- Advantages of the linear probability model:
    - easy to estimate
    - Coeffcient estimates are easy to interpret
- Disadvantages of the linear probability model
    - Predicted probability can be above 1 or below 0!(it doesn't make sense)
    - Error terms are heteroskedastic

An Example: Mortgage applications

- Most individuals who want to buy a house apply for a mortgage at a bank.

An Example: Mortgage applications

- Most individuals who want to buy a house apply for a mortgage at a bank.
- Not all mortgage applications are approved.

An Example: Mortgage applications

- Most individuals who want to buy a house apply for a mortgage at a bank.
- Not all mortgage applications are approved.
- What determines whether or not a mortgage application is approved or denied?

An Example: Mortgage applications

- eg.Boston HMDA data: a data set on mortgage(住房贷款) applications collected by the Federal Reserve Bank in Boston.

Variable	Description	Mean	SD
deny	$= 1$ if application is denied	0.120	0.325
pi_ratio	monthly loan payments / monthly income	0.331	0.107
black	$= 1$ if applicant is black	0.142	0.350

An Example: Mortgage applications

- Does the payment to income ratio affect whether or not a mortgage application is denied?

$$\widehat{deny} = -0.080 + 0.604\ P/I\ ratio$$
$$(0.032)(0.098)$$

An Example: Mortgage applications

- Does the payment to income ratio affect whether or not a mortgage application is denied?

$$\widehat{deny} = -0.080 + 0.604\ P/I\ ratio$$
$$(0.032)(0.098)$$

- The estimated OLS coefficient on the payment to income ratio $\hat{\beta}_1 = 0.60$

An Example: Mortgage applications

- Does the payment to income ratio affect whether or not a mortgage application is denied?

$$\widehat{deny} = -0.080 + 0.604\ P/I\ ratio$$
$$(0.032)(0.098)$$

- The estimated OLS coefficient on the payment to income ratio $\hat{\beta}_1 = 0.60$
- The estimated coefficient is significantly different from 0 at a 1% significance level.

An Example: Mortgage applications

- Does the payment to income ratio affect whether or not a mortgage application is denied?

$$\widehat{deny} = -0.080 + 0.604\ P/I\ ratio$$
$$(0.032)(0.098)$$

- The estimated OLS coefficient on the payment to income ratio $\hat{\beta}_1 = 0.60$
- The estimated coefficient is significantly different from 0 at a 1% significance level.
- How should we interpret $\hat{\beta}_1$?

An Example: Mortgage applications

- Does the payment to income ratio affect whether or not a mortgage application is denied?

$$\widehat{deny} = -0.080 + 0.604\ P/I\ ratio$$
$$(0.032)(0.098)$$

- The estimated OLS coefficient on the payment to income ratio $\hat{\beta}_1 = 0.60$
- The estimated coefficient is significantly different from 0 at a 1% significance level.
- How should we interpret $\hat{\beta}_1$?
  - An original one: "payments/monthly income ratio increase *1*,then **probability being denied** will also increase *0.6*."

An Example: Mortgage applications

- Does the payment to income ratio affect whether or not a mortgage application is denied?

$$\widehat{deny} = -0.080 + 0.604\ P/I\ ratio$$
$$(0.032)(0.098)$$

- The estimated OLS coefficient on the payment to income ratio $\hat{\beta}_1 = 0.60$
- The estimated coefficient is significantly different from 0 at a 1% significance level.
- How should we interpret $\hat{\beta}_1$?
  - An original one: "payments/monthly income ratio increase *1*,then **probability being denied** will also increase *0.6*."
  - More reasonable one: "payments/monthly income ratio increase *0.1(10%)*,then probability being denied will also increase *0.06(6%)*".

An Example: Mortgage applications

- What is the effect of race on the probability of *denial*, *holding constant the P/I ratio?* To keep things simple, we focus on differences between black applicants and white applicants.

$$\widehat{deny} = \underset{(0.029)}{-0.091} + \underset{(0.089)}{0.559}\ P/I\ ratio + \underset{(0.025)}{0.177} black$$

An Example: Mortgage applications

- What is the effect of race on the probability of *denial*, *holding constant the P/I ratio?* To keep things simple, we focus on differences between black applicants and white applicants.

$$\widehat{deny} = \underset{(0.029)}{-0.091} + \underset{(0.089)}{0.559}\ P/I\ ratio + \underset{(0.025)}{0.177}\, black$$

- The coefficient on black, *0.177*, indicates that an African American applicant has a *17.7%* higher probability of having a mortgage application denied than a white applicant, holding constant their payment-to-income ratio.

An Example: Mortgage applications

- What is the effect of race on the probability of *denial*, *holding constant the P/I ratio?* To keep things simple, we focus on differences between black applicants and white applicants.

$$\widehat{deny} = \underset{(0.029)}{-0.091} + \underset{(0.089)}{0.559}\, P/I\, ratio + \underset{(0.025)}{0.177} black$$

- The coefficient on black, *0.177*, indicates that an African American applicant has a *17.7%* higher probability of having a mortgage application denied than a white applicant, holding constant their payment-to-income ratio.
- This coefficient is significant at the 1% level (the t-statistic is 7.11).

LPM: shortcomings

- Always suffer heteroskedasticity.

LPM: shortcomings

- Always suffer heteroskedasticity.
  - Always use heteroskedasticity robust standard errors!

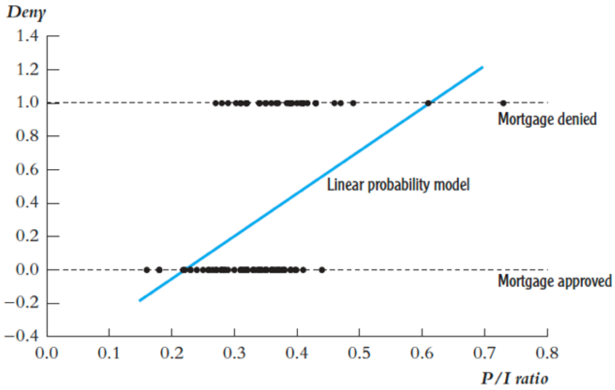
LPM: shortcomings

- Always suffer heteroskedasticity.
  - Always use heteroskedasticity robust standard errors!
- While in LPM model,the predicted probability can be below 0 or above 1!

Mortgage applications: Predicted value

FIGURE 11.1 Scatterplot of Mortgage Application Denial and the Payment-to-Income Ratio

Mortgage applicants with a high ratio of debt payments to income (*P/I ratio*) are more likely to have their application denied (*deny* = 1 if denied, *deny* = 0 if approved). The linear probability model uses a straight line to model the probability of denial, conditional on the *P/I ratio*.



Nonlinear probability model

Introduction

- Intuition: Probabilities should not be less than 0 or greater than 1

Introduction

- Intuition: Probabilities should not be less than 0 or greater than 1
- To address this problem,consider nonlinear probability models

$$\begin{aligned} Pr(Y_i = 1|X_1, ...X_k) &= G(Z) \\ &= G(\beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + ... + \beta_k X_{k,i}) \end{aligned}$$

where $Z = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + ... + \beta_k X_{k,i}$ and $0 \leq g(Z) \leq 1$

Logit and Probit

- Two types nonlinear functions

Logit and Probit

- Two types nonlinear functions
  1. Probit

$$G(Z) = \Phi(Z) = \int_{-\infty}^{z} \phi(Z)dZ$$

Logit and Probit

- Two types nonlinear functions
  1. Probit
  $$G(Z) = \Phi(Z) = \int_{-\infty}^{z} \phi(Z)dZ$$
  2. Logit
  $$G(Z) = \frac{1}{1 + e^{-Z}}$$

Probit Model

- Probit regression models the probability that $Y = 1$

$$Pr(Y_i = 1|) = \Phi(\beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + ... + \beta_k X_{k,i})$$

Probit Model

- Probit regression models the probability that $Y = 1$

$$Pr(Y_i = 1|) = \Phi(\beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + ... + \beta_k X_{k,i})$$

- Using the cumulative standard normal distribution function $\Phi(Z)$ and $0 \leq \Phi(Z) \leq 1$

Probit Model

- Probit regression models the probability that $Y = 1$

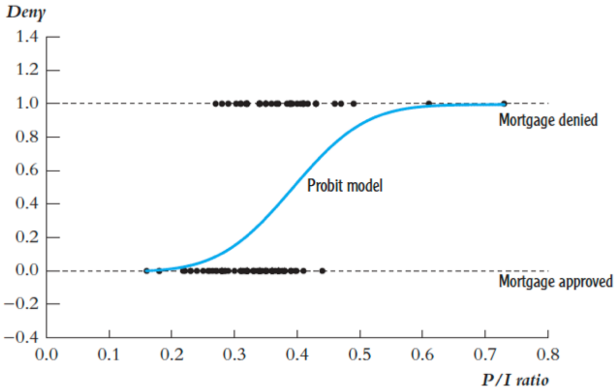
$$Pr(Y_i = 1|) = \Phi(\beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + ... + \beta_k X_{k,i})$$

- Using the cumulative standard normal distribution function $\Phi(Z)$ and $0 \leq \Phi(Z) \leq 1$
- Since $\Phi(z) = Pr(Z \leq z)$ we have that the predicted probabilities of the probit model are between 0 and 1.

Probit Model

FIGURE 11.2 Probit Model of the Probability of Denial, Given *P/I Ratio*

The probit model uses the cumulative normal distribution function to model the probability of denial given the payment-to-income ratio or, more generally, to model $\Pr(Y = 1 | X)$. Unlike the linear probability model, the probit conditional probabilities are always between 0 and 1.



Probit Model

- evaluated at $Z = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + ... + \beta_k X_{k,i}$

Probit Model

- evaluated at $Z = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + ... + \beta_k X_{k,i}$
- The coefficient β_1 is the change in the z-value arising from a unit change in X_1, holding constant $X_2, ..., X_k$.

Probit Model

- evaluated at $Z = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + ... + \beta_k X_{k,i}$
- The coefficient β_1 is the change in the z-value arising from a unit change in X_1, holding constant $X_2, ..., X_k$.
- The effect on the predicted probability of a change in a regressor is computed by

Probit Model

- evaluated at $Z = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + ... + \beta_k X_{k,i}$
- The coefficient β_1 is the change in the z-value arising from a unit change in X_1, holding constant $X_2, ..., X_k$.
- The effect on the predicted probability of a change in a regressor is computed by

1. computing the predicted probability for the initial value of the regressors,

Probit Model

- evaluated at $Z = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + ... + \beta_k X_{k,i}$
- The coefficient β_1 is the change in the z-value arising from a unit change in X_1, holding constant $X_2, ..., X_k$.
- The effect on the predicted probability of a change in a regressor is computed by

1. computing the predicted probability for the initial value of the regressors,
2. computing the predicted probability for the new or changed value of the regressors,

Probit Model

- evaluated at $Z = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + ... + \beta_k X_{k,i}$
- The coefficient β_1 is the change in the z-value arising from a unit change in X_1, holding constant $X_2, ..., X_k$.
- The effect on the predicted probability of a change in a regressor is computed by

1. computing the predicted probability for the initial value of the regressors,
2. computing the predicted probability for the new or changed value of the regressors,
3. taking their difference.

Probit Model with one regression

- Suppose the probit population regression model with only one regressors, X_1

$$Pr(Y = 1|X_1) = \Phi(Z) = \Phi(\beta_0 + \beta_1 X_1)$$

Probit Model with one regression

- Suppose the probit population regression model with only one regressors, X_1

$$Pr(Y = 1|X_1) = \Phi(Z) = \Phi(\beta_0 + \beta_1 X_1)$$

- Suppsoe the estimate result is $Z = -2 + 3X_1$,

Probit Model with one regression

- Suppose the probit population regression model with only one regressors, X_1

$$Pr(Y = 1|X_1) = \Phi(Z) = \Phi(\beta_0 + \beta_1 X_1)$$

- Suppsoe the estimate result is $Z = -2 + 3X_1$,
- And we want to know the probability that $Y = 1$ when $X_1 = 0.4$

Probit Model with one regression

- Suppose the probit population regression model with only one regressors, X_1

$$Pr(Y = 1|X_1) = \Phi(Z) = \Phi(\beta_0 + \beta_1 X_1)$$

- Suppsoe the estimate result is $Z = -2 + 3X_1$,
- And we want to know the probability that $Y = 1$ when $X_1 = 0.4$
- Then $z = -2 + 3 \times 0.4 = -0.8$

Probit Model with one regression

- Suppose the probit population regression model with only one regressors, X_1

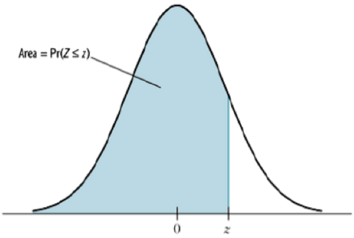
$$Pr(Y = 1|X_1) = \Phi(Z) = \Phi(\beta_0 + \beta_1 X_1)$$

- Suppsoe the estimate result is $Z = -2 + 3X_1$,
- And we want to know the probability that $Y = 1$ when $X_1 = 0.4$
- Then $z = -2 + 3 \times 0.4 = -0.8$
- So the probability $Pr(Y = 1) = Pr(z \leq -0.8) = \Phi(-0.8)$

Probit Model

- $Pr(Y = 1) = Pr(Z \leq -0.8) = \Phi(-.8) = 0.2119$

TABLE 1 The Cumulative Standard Normal Distribution Function, $\Phi(z) = \Pr(Z$ ″ $z)$



z	Second Decimal Value of z: 0	1	2	3	4	5	6	7	8	9
−2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
−2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
−0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
−0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
−0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
−0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
−0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121

Probit Model with multiple regressors

- Suppose the probit population regression model with two regressors, X_1 and X_2,

$$Pr(Y = 1|X_1, X_2) = \Phi(\beta_0 + \beta_1 X_1 + \beta_2 X_2)$$

Probit Model with multiple regressors

- Suppose the probit population regression model with two regressors, X_1 and X_2,

$$Pr(Y=1|X_1,X_2)=\Phi(\beta_0+\beta_1X_1+\beta_2X_2)$$

- Suppose $\beta_0=-1.6, \beta_1=2\ and\ \beta_2=0.5$. And If $X_1=0.4\ and\ X_2=1$ then

$$z=-1.6+2\times 0.4+0.5\times 1=-0.3$$

Probit Model with multiple regressors

- Suppose the probit population regression model with two regressors, X_1 and X_2,

$$Pr(Y = 1|X_1, X_2) = \Phi(\beta_0 + \beta_1 X_1 + \beta_2 X_2)$$

- Suppose $\beta_0 = -1.6, \beta_1 = 2 \; and \; \beta_2 = 0.5$. And If $X_1 = 0.4 \; and \; X_2 = 1$ then

$$z = -1.6 + 2 \times 0.4 + 0.5 \times 1 = -0.3$$

- So the probability $Pr(Y = 1) = Pr(z \leq -0.3) = \Phi(-0.3) = 0.38$

Example: Mortgage Applications

- Mortgage denial (deny) and the payment-toincome ratio (P/I ratio)

$$Pr(deny \widehat{= 1|P}/I\ ratio) = \Phi(-2.19 + 2.97P/I\ ratio)$$

Example: Mortgage Applications

- Mortgage denial (deny) and the payment-toincome ratio (P/I ratio)

$$\widehat{Pr(deny = 1|P/I\ ratio)} = \Phi(-2.19 + 2.97P/I\ ratio)$$

- *What is the change in the predicted probability* that an application will be deniedif P/I ratio incrreases from *0.3 to 0.4*?

Example: Mortgage Applications

- Mortgage denial (deny) and the payment-toincome ratio (P/I ratio)

$$Pr(\widehat{deny = 1|P}/I\ ratio) = \Phi(-2.19 + 2.97P/I\ ratio)$$

- *What is the change in the predicted probability* that an application will be deniedif P/I ratio incrreases from *0.3 to 0.4*?
- The probability of denial when $P/I\ ratio = 0.3$

$$\Phi(-2.19 + 2.97 \times 0.3) = \Phi(-1.3) = 0.097$$

Example: Mortgage Applications

- Mortgage denial (deny) and the payment-toincome ratio (P/I ratio)

$$\widehat{Pr(deny = 1|P}/I\ ratio) = \Phi(-2.19 + 2.97P/I\ ratio)$$

- *What is the change in the predicted probability* that an application will be deniedif P/I ratio incrreases from *0.3 to 0.4*?
- The probability of denial when $P/I\ ratio = 0.3$

$$\Phi(-2.19 + 2.97 \times 0.3) = \Phi(-1.3) = 0.097$$

- The probability of denial when $P/I\ ratio = 0.4$

$$\Phi(-2.19 + 2.97 \times 0.4) = \Phi(-1.0) = 0.159$$

Example: Mortgage Applications

- Mortgage denial (deny) and the payment-toincome ratio (P/I ratio)

$$Pr(deny \widehat{= 1|P}/I\ ratio) = \Phi(-2.19 + 2.97P/I\ ratio)$$

- *What is the change in the predicted probability* that an application will be deniedif P/I ratio incrreases from *0.3 to 0.4*?
- The probability of denial when $P/I\ ratio = 0.3$

$$\Phi(-2.19 + 2.97 \times 0.3) = \Phi(-1.3) = 0.097$$

- The probability of denial when $P/I\ ratio = 0.4$

$$\Phi(-2.19 + 2.97 \times 0.4) = \Phi(-1.0) = 0.159$$

- The estimated change in the probability of denial is $0.159 - 0.097 = 0.062$

Effect of a change in X: Marginal Effects

- For nonlinear models, the ME varies with the point of evaluation

Effect of a change in X: Marginal Effects

- For nonlinear models, the ME varies with the point of evaluation
  - *Marginal Effect at a Representative Value* (MER):ME at $X = X^*$ (at representative values of the regerssors)

Effect of a change in X: Marginal Effects

- For nonlinear models, the ME varies with the point of evaluation
  - *Marginal Effect at a Representative Value* (MER):ME at $X = X^*$ (at representative values of the regerssors)
  - *Marginal Effect at Mean* (MEM): ME at $X = \bar{X}$(at the sample mean of the regressors)

Effect of a change in X: Marginal Effects

- For nonlinear models, the ME varies with the point of evaluation
  - *Marginal Effect at a Representative Value* (MER):ME at $X = X^*$ (at representative values of the regerssors)
  - *Marginal Effect at Mean* (MEM): ME at $X = \bar{X}$(at the sample mean of the regressors)
  - *Average Marginal Effect* (AME): average of ME at each $X = X_i$ (at sample values and then average)

Example: Mortgage applications: marginal effect

- Because the probit regression function is nonlinear, the effect of a change in X depends on the starting value of X.

$$\frac{\partial Pr(deny = 1|P/I\ ratio)}{\partial P/I\ ratio} = \Phi(-2.19 + 2.97P/I\ ratio) \times 2.97$$

$$\left.\frac{\partial Pr(deny = 1|P/I\ ratio)}{\partial P/I\ ratio}\right|_{at\ mean} = \Phi(-2.19 + 2.97 \times 0.331) \times 2.97$$

Example: Mortgage applications: marginal effect

- Because the probit regression function is nonlinear, the effect of a change in X depends on the starting value of X.

$$\frac{\partial Pr(deny = 1|P/I\ ratio)}{\partial P/I\ ratio} = \Phi(-2.19 + 2.97P/I\ ratio) \times 2.97$$

- *Marginal Effect at Mean* (MEM):(at the sample mean of the regressors: $P/I\ ratio_{mean} = 0.331$

$$\frac{\partial Pr(deny = 1|P/I\ ratio)}{\partial P/I\ ratio}_{at\ mean} = \Phi(-2.19 + 2.97 \times 0.331) \times 2.97$$

Special Case: The explanatory variable is discrete.

- If xj is a discrete variable, then we should not rely on calculus in evaluating the effect on the response probability.

Special Case: The explanatory variable is discrete.

- If *xj* is a discrete variable, then we should not rely on calculus in evaluating the effect on the response probability.
- Assume X_2 is a dummy variable, then partial effect of X_2 changing from 0 to 1:

$$G(\beta_0+\beta_1 X_{1,i}+\beta_2\times 1+\ldots+\beta_k X_{k,i})-G(\beta_0+\beta_1 X_{1,i}+\beta_2\times 0+\ldots+\beta_k X_{k,i})$$

Example: Mortgage applications: Race

```
##
## Call:
## glm(formula = hmda_small$deny ~ hmda_small$pi_rat + hmda_sm
##     family = binomial(link = "probit"), data = hmda_small)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.1208  -0.4762  -0.4251  -0.3550   2.8799
##
## Coefficients:
##                   Estimate Std. Error z value Pr(>|z|)
## (Intercept)       -2.25879    0.13669 -16.525  < 2e-16 ***
## hmda_small$pi_rat  2.74178    0.38047   7.206 5.75e-13 ***
## hmda_small$black   0.70816    0.08335   8.496  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '
```

Example: Mortgage applications: Race

- Mortgage denial (deny) and the payment-toincome ratio (P/I ratio) and race

$$\widehat{Pr(deny = 1|P}/I\ ratio) = \Phi(-2.26 + 2.74P/I\ ratio + 0.71black)$$

Example: Mortgage applications: Race

- Mortgage denial (deny) and the payment-toincome ratio (P/I ratio) and race

$$\widehat{Pr(deny = 1|P}/I\ ratio) = \Phi(-2.26 + 2.74P/I\ ratio + 0.71black)$$

- The probability of denial when $black = 0$,thus whites(non-blacks) is

$$\Phi(-2.26 + 2.74 \times 0.3 + 0.71 \times 0) = \Phi(-1.43) = 0.075$$

Example: Mortgage applications: Race

- Mortgage denial (deny) and the payment-toincome ratio (P/I ratio) and race

$$\widehat{Pr(deny = 1|P}/I\ ratio) = \Phi(-2.26 + 2.74P/I\ ratio + 0.71black)$$

- The probability of denial when $black = 0$,thus whites(non-blacks) is

$$\Phi(-2.26 + 2.74 \times 0.3 + 0.71 \times 0) = \Phi(-1.43) = 0.075$$

- The probability of denial when $black = 1$,thus blacks is

$$\Phi(-2.26 + 2.74 \times 0.3 + 0.71 \times 1) = \Phi(-0.73) = 0.233$$

Example: Mortgage applications: Race

- Mortgage denial (deny) and the payment-toincome ratio (P/I ratio) and race

$$\widehat{Pr(deny = 1|P/I\ ratio)} = \Phi(-2.26 + 2.74 P/I\ ratio + 0.71 black)$$

- The probability of denial when $black = 0$,thus whites(non-blacks) is

$$\Phi(-2.26 + 2.74 \times 0.3 + 0.71 \times 0) = \Phi(-1.43) = 0.075$$

- The probability of denial when $black = 1$,thus blacks is

$$\Phi(-2.26 + 2.74 \times 0.3 + 0.71 \times 1) = \Phi(-0.73) = 0.233$$

- so the difference between whites and blacks at $P/Iratio = 0.3$ is $0.233 - 0.075 = 0.158$, which means probability of denial for blacks is 15.8% higher than that for whites.

Logit Model

- Logit regression models the probability that $Y = 1$

Logit Model

- Logit regression models the probability that Y $= 1$
- Using the cumulative standard logistic distribution function

$$Pr(Y_i = 1|Z) = \frac{1}{1 + e^{-Z}}$$

Logit Model

- Logit regression models the probability that Y $= 1$
- Using the cumulative standard logistic distribution function

$$Pr(Y_i = 1|Z) = \frac{1}{1 + e^{-Z}}$$

- evaluated at $Z = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + ... + \beta_k X_{k,i}$

Logit Model

- Logit regression models the probability that Y $= 1$
- Using the cumulative standard logistic distribution function

$$Pr(Y_i = 1|Z) = \frac{1}{1 + e^{-Z}}$$

- evaluated at $Z = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + ... + \beta_k X_{k,i}$
- since $F(z) = Pr(Z \leq z)$ we have that the predicted probabilities of the probit model are between 0 and 1.

Logit Model

- Suppose we have only one regressor X and $Z = -2 + 3X_1$

Logit Model

- Suppose we have only one regressor X and $Z = -2 + 3X_1$
- We want to know the probability that $Y = 1$ when $X_1 = 0.4$

Logit Model

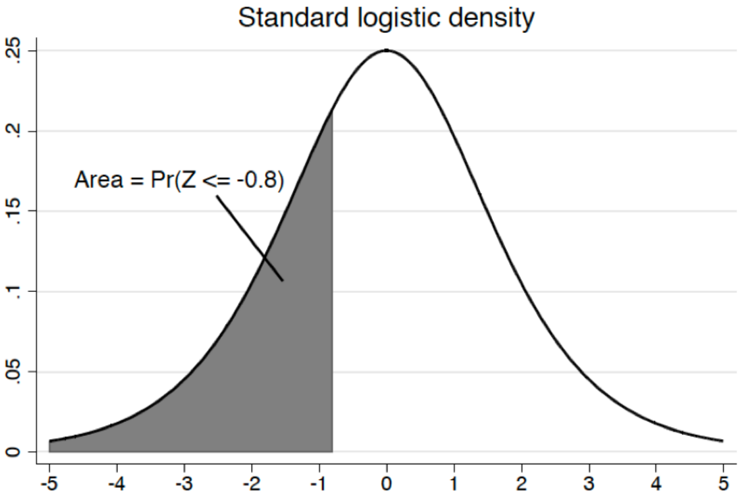
- Suppose we have only one regressor X and $Z = -2 + 3X_1$
- We want to know the probability that $Y = 1$ when $X_1 = 0.4$
- Then $Z = -2 + 3 \times 0.4 = -0.8$

Logit Model

- Suppose we have only one regressor X and $Z = -2 + 3X_1$
- We want to know the probability that $Y = 1$ when $X_1 = 0.4$
- Then $Z = -2 + 3 \times 0.4 = -0.8$
- So the probability $Pr(Y = 1) = Pr(Z \leq -0.8) = F(-0.8)$

Logit Model

- $Pr(Y=1) = Pr(Z \leq -0.8) = \frac{1}{1+e^{0.8}} = 0.31$



Example: Mortgage applications

- Logit Model: Mortgage denial (deny) and the payment-toincome ratio (P/I ratio) and race

$$\widehat{Pr(deny = 1|P/I\ ratio)} = F(\underset{(0.35)}{-4.13} + \underset{(0.96)}{5.37} P/I\ ratio + \underset{(0.15)}{1.27} black)$$

Example: Mortgage applications: Race

- The predicted denial probability of a white applicant with $P/I\ ratio = 0.3$ is

$$\frac{1}{1 + e^{-(-4.13+5.37\times 0.3+1.27\times 0)}} = 0.074$$

Example: Mortgage applications: Race

- The predicted denial probability of a white applicant with $P/I\ ratio = 0.3$ is

$$\frac{1}{1+e^{-(-4.13+5.37\times 0.3+1.27\times 0)}} = 0.074$$

- The predicted denial probability of a black applicant with $P/I\ ratio = 0.3$ is

$$\frac{1}{1+e^{-(-4.13+5.37\times 0.3+1.27\times 1)}} = 0.222$$

Example: Mortgage applications: Race

- The predicted denial probability of a white applicant with $P/I\ ratio = 0.3$ is

$$\frac{1}{1+e^{-(-4.13+5.37\times 0.3+1.27\times 0)}} = 0.074$$

- The predicted denial probability of a black applicant with $P/I\ ratio = 0.3$ is

$$\frac{1}{1+e^{-(-4.13+5.37\times 0.3+1.27\times 1)}} = 0.222$$

- the difference is

$$0.222 - 0.074 = 0.148 = 14.8\%$$

How to estimate Logit and Probit models

- nonlinear in the independent variables(X).

How to estimate Logit and Probit models

- nonlinear in the independent variables(X).
  - these models can be estimated by OLS

How to estimate Logit and Probit models

- nonlinear in the independent variables(X).
  - these models can be estimated by OLS
- While Logit and Probit models are nonlinear in the coefficients $\beta_0, \beta_1, ..., \beta_k$

How to estimate Logit and Probit models

- nonlinear in the independent variables(X).
    - these models can be estimated by OLS
- While Logit and Probit models are nonlinear in the coefficients $\beta_0, \beta_1, ..., \beta_k$
    - these models can NOT be estimated by OLS

How to estimate Logit and Probit models

- nonlinear in the independent variables(X).
  - these models can be estimated by OLS
- While Logit and Probit models are nonlinear in the coefficients $\beta_0, \beta_1, ..., \beta_k$
  - these models can NOT be estimated by OLS
- The method used to estimate logit and probit models is **Maximum Likelihood Estimation** (MLE).

MLE estimator in practice

- There is no simple formula for the probit and logit MLE, the maximization must be done using **numerical algorithm** on a computer.

MLE estimator in practice

- There is no simple formula for the probit and logit MLE, the maximization must be done using **numerical algorithm** on a computer.
- Because regression software commonly computes the MLE of the estimate coefficients, this estimator is easy to use in practice.

MLE estimator in practice

- There is no simple formula for the probit and logit MLE, the maximization must be done using **numerical algorithm** on a computer.
- Because regression software commonly computes the MLE of the estimate coefficients, this estimator is easy to use in practice.
- The MLE is consistent and normally distributed in large samples.

Statistical inference based on the MLE

- Because the MLE is normally distributed in large samples, statistical inference about the probit and logit coefficients based on the MLE proceeds in the same way as inference about the linear regression function coefficients based on the OLS estimator.

$$F_{stat} \longrightarrow \frac{\chi^2_q}{q}$$

where q is the number of restrictions being tested.

Statistical inference based on the MLE

- Because the MLE is normally distributed in large samples, statistical inference about the probit and logit coefficients based on the MLE proceeds in the same way as inference about the linear regression function coefficients based on the OLS estimator.
- That is, hypothesis tests are performed using the **t-statistic** and **95% confidence intervals** are formed as 1.96 standard errors.

$$F_{stat} \longrightarrow \frac{\chi^2_q}{q}$$

where q is the number of restrictions being tested.

Statistical inference based on the MLE

- Because the MLE is normally distributed in large samples, statistical inference about the probit and logit coefficients based on the MLE proceeds in the same way as inference about the linear regression function coefficients based on the OLS estimator.
- That is, hypothesis tests are performed using the **t-statistic** and **95% confidence intervals** are formed as 1.96 standard errors.
- Tests of joint hypotheses on multiple coefficients use the **F-statistic** in a way similar to that discussed for the linear regression model.

$$F_{stat} \longrightarrow \frac{\chi_q^2}{q}$$

where q is the number of restrictions being tested.

Statistical inference based on the MLE

- Because the MLE is normally distributed in large samples, statistical inference about the probit and logit coefficients based on the MLE proceeds in the same way as inference about the linear regression function coefficients based on the OLS estimator.
- That is, hypothesis tests are performed using the **t-statistic** and **95% confidence intervals** are formed as 1.96 standard errors.
- Tests of joint hypotheses on multiple coefficients use the **F-statistic** in a way similar to that discussed for the linear regression model.
- F-statistic and Chi-squared stistic

$$F_{stat} \longrightarrow \frac{\chi_q^2}{q}$$

where q is the number of restrictions being tested.

Measures of Fit

- R2 is a poor measure of fit for the linear probability model. This is also true for probit and logit regression.

Measures of Fit

- R2 is a poor measure of fit for the linear probability model. This is also true for probit and logit regression.
- Two measures of fit for models with binary dependent variables

Measures of Fit

- R2 is a poor measure of fit for the linear probability model. This is also true for probit and logit regression.
- Two measures of fit for models with binary dependent variables

1. *fraction correctly predicted*

Measures of Fit

- R2 is a poor measure of fit for the linear probability model. This is also true for probit and logit regression.
- Two measures of fit for models with binary dependent variables

1. *fraction correctly predicted*

- If $Y_i = 1$ and the predicted probability exceeds 50% or if $Y_i = 0$ and the predicted probability is less than 50%, then Y_i is said to be correctly predicted.

Measures of Fit

2. **The pseudo-R2**

Measures of Fit

2. **The pseudo-R2**

- The $pseudo - R^2$ compares the value of the likelihood of the estimated model to the value of the likelihood when none of the Xs are included as regressors.

$$pseudo - R^2 = 1 - \frac{ln(f_{probit}^{max})}{ln(f_{bernoulli}^{max})}$$

Measures of Fit

2. **The pseudo-R2**

- The $pseudo - R^2$ compares the value of the likelihood of the estimated model to the value of the likelihood when none of the Xs are included as regressors.

$$pseudo - R^2 = 1 - \frac{ln(f^{max}_{probit})}{ln(f^{max}_{bernoulli})}$$

- f^{max}_{probit} is the value of the maximized probit likelihood (which includes the X's)

Measures of Fit

2. **The pseudo-R2**

- The $pseudo-R^2$ compares the value of the likelihood of the estimated model to the value of the likelihood when none of the Xs are included as regressors.

$$pseudo-R^2 = 1 - \frac{ln(f^{max}_{probit})}{ln(f^{max}_{bernoulli})}$$

- f^{max}_{probit} is the value of the maximized probit likelihood (which includes the X's)
- $f^{max}_{bernoulli}$ is the value of the maximized Bernoulli likelihood (the probit model excluding all the X's).

Comparing the LPM,Probit and Logit

- All three models: *linear probability, probit, and logit* are just approximations to the unknown population regression function $E(Y|X) = Pr(Y = 1|X)$.

Comparing the LPM,Probit and Logit

- All three models: *linear probability, probit, and logit* are just approximations to the unknown population regression function $E(Y|X) = Pr(Y = 1|X)$.
    - LPM is easiest to use and to interpret, but it cannot capture the nonlinear nature of the true population regression function.

Comparing the LPM,Probit and Logit

- All three models: *linear probability, probit, and logit* are just approximations to the unknown population regression function $E(Y|X) = Pr(Y = 1|X)$.
  - LPM is easiest to use and to interpret, but it cannot capture the nonlinear nature of the true population regression function.
  - Probit and logit regressions model this nonlinearity in the probabilities, but their regression coefficients are more difficult to interpret.

Comparing the LPM,Probit and Logit

- All three models: *linear probability, probit, and logit* are just approximations to the unknown population regression function $E(Y|X) = Pr(Y = 1|X)$.
    - LPM is easiest to use and to interpret, but it cannot capture the nonlinear nature of the true population regression function.
    - Probit and logit regressions model this nonlinearity in the probabilities, but their regression coefficients are more difficult to interpret.
- So which should you use in practice?

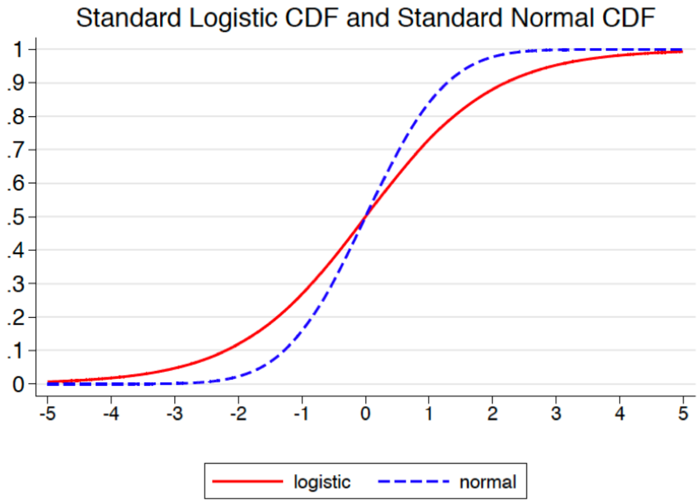
Comparing the LPM,Probit and Logit

- All three models: *linear probability, probit, and logit* are just approximations to the unknown population regression function $E(Y|X) = Pr(Y = 1|X)$.
  - LPM is easiest to use and to interpret, but it cannot capture the nonlinear nature of the true population regression function.
  - Probit and logit regressions model this nonlinearity in the probabilities, but their regression coefficients are more difficult to interpret.
- So which should you use in practice?
  - *There is no one right answer*, and different researchers use different models.

Comparing the LPM,Probit and Logit

- All three models: *linear probability, probit, and logit* are just approximations to the unknown population regression function $E(Y|X) = Pr(Y = 1|X)$.
  - LPM is easiest to use and to interpret, but it cannot capture the nonlinear nature of the true population regression function.
  - Probit and logit regressions model this nonlinearity in the probabilities, but their regression coefficients are more difficult to interpret.
- So which should you use in practice?
  - *There is no one right answer*, and different researchers use different models.
  - *Probit and logit regressions frequently produce similar results*.

Logit v.s. Probit



Comparing the LPM,Probit and Logit

- The marginal effects and predicted probabilities are much more similar across models.

Comparing the LPM,Probit and Logit

- The marginal effects and predicted probabilities are much more similar across models.
- Coefficients can be compared across models, using the following rough conversion factors (Amemiya 1981)

$$\hat{\beta}_{logit} \simeq 4\hat{\beta}_{ols}$$
$$\hat{\beta}_{probit} \simeq 2.5\hat{\beta}_{ols}$$
$$\hat{\beta}_{logit} \simeq 1.6\hat{\beta}_{probit}$$

Example: Mortgage Applications(short regression)

Dependent variable: *deny* = 1 if mortgage application is denied, = 0 if accepted			
regression model	LPM	Probit	Logit
black	0.177***	0.71***	1.27***
	(0.025)	(0.083)	(0.15)
P/I ratio	0.559***	2.74***	5.37***
	(0.089)	(0.44)	(0.96)
constant	-0.091***	-2.26***	-4.13***
	(0.029)	(0.16)	(0.35)
difference Pr(*deny*=1) between black and white applicant when *P/I ratio*=0.3	17.7%	15.8%	14.8%

Example: Mortgage Applications(long regression)

TABLE 11.1 Variables Included in Regression Models of Mortgage Decisions

Variable	Definition	Sample Average
Financial Variables		
P/I ratio	Ratio of total monthly debt payments to total monthly income	0.331
housing expense-to-income ratio	Ratio of monthly housing expenses to total monthly income	0.255
loan-to-value ratio	Ratio of size of loan to assessed value of property	0.738
consumer credit score	1 if no "slow" payments or delinquencies 2 if one or two slow payments or delinquencies 3 if more than two slow payments 4 if insufficient credit history for determination 5 if delinquent credit history with payments 60 days overdue 6 if delinquent credit history with payments 90 days overdue	2.1
mortgage credit score	1 if no late mortgage payments 2 if no mortgage payment history 3 if one or two late mortgage payments 4 if more than two late mortgage payments	1.7
public bad credit record	1 if any public record of credit problems (bankruptcy, charge-offs, collection actions) 0 otherwise	0.074

Example: Mortgage Applications(long regression)

Additional Applicant Characteristics

denied mortgage insurance	1 if applicant applied for mortgage insurance and was denied, 0 otherwise	0.020
self-employed	1 if self-employed, 0 otherwise	0.116
single	1 if applicant reported being single, 0 otherwise	0.393
high school diploma	1 if applicant graduated from high school, 0 otherwise	0.984
unemployment rate	1989 Massachusetts unemployment rate in the applicant's industry	3.8
condominium	1 if unit is a condominium, 0 otherwise	0.288
black	1 if applicant is black, 0 if white	0.142
deny	1 if mortgage application denied, 0 otherwise	0.120

图 2: pic

Example: Mortgage Applications(long regression)

TABLE 11.2 Mortgage Denial Regressions Using the Boston HMDA Data

Dependent variable: *deny* = 1 if mortgage application is denied, = 0 if accepted; 2380 observations.

Regression Model	LPM	Logit	Probit	Probit	Probit	Probit
Regressor	(1)	(2)	(3)	(4)	(5)	(6)
black	0.084** (0.023)	0.688** (0.182)	0.389** (0.098)	0.371** (0.099)	0.363** (0.100)	0.246 (0.448)
P/I ratio	0.449** (0.114)	4.76** (1.33)	2.44** (0.61)	2.46** (0.60)	2.62** (0.61)	2.57** (0.66)
housing expense-to-income ratio	−0.048 (0.110)	−0.11 (1.29)	−0.18 (0.68)	−0.30 (0.68)	−0.50 (0.70)	−0.54 (0.74)
medium loan-to-value ratio (0.80 ≤ loan-value ratio ≤ 0.95)	0.031* (0.013)	0.46** (0.16)	0.21** (0.08)	0.22** (0.08)	0.22** (0.08)	0.22** (0.08)
high loan-to-value ratio (loan-value ratio > 0.95)	0.189** (0.050)	1.49** (0.32)	0.79** (0.18)	0.79** (0.18)	0.84** (0.18)	0.79** (0.18)
consumer credit score	0.031** (0.005)	0.29** (0.04)	0.15** (0.02)	0.16** (0.02)	0.34** (0.11)	0.16** (0.02)
mortgage credit score	0.021 (0.011)	0.28* (0.14)	0.15* (0.07)	0.11 (0.08)	0.16 (0.10)	0.11 (0.08)
public bad credit record	0.197** (0.035)	1.23** (0.20)	0.70** (0.12)	0.70** (0.12)	0.72** (0.12)	0.70** (0.12)
denied mortgage insurance	0.702** (0.045)	4.55** (0.57)	2.56** (0.30)	2.59** (0.29)	2.59** (0.30)	2.59** (0.29)
self-employed	0.060** (0.021)	0.67** (0.21)	0.36** (0.11)	0.35** (0.11)	0.34** (0.11)	0.35** (0.11)

Example: Mortgage Applications(long regression)

single				0.23** (0.08)	0.23** (0.08)	0.23** (0.08)
high school diploma				–0.61** (0.23)	–0.60* (0.24)	–0.62** (0.23)
unemployment rate				0.03 (0.02)	0.03 (0.02)	0.03 (0.02)
condominium					–0.05 (0.09)	
black × *P/I ratio*						–0.58 (1.47)
black × *housing expense-to-income ratio*						1.23 (1.69)
additional credit rating indicator variables	no	no	no	no	yes	no
constant	–0.183** (0.028)	–5.71** (0.48)	–3.04** (0.23)	–2.57** (0.34)	–2.90** (0.39)	–2.54** (0.35)

Example: Mortgage Applications(long regression)

(Table 11.2 continued)

F-Statistics and p-Values Testing Exclusion of Groups of Variables

	(1)	(2)	(3)	(4)	(5)	(6)
applicant single; high school diploma; industry unemployment rate				5.85 (< 0.001)	5.22 (0.001)	5.79 (< 0.001)
additional credit rating indicator variables					1.22 (0.291)	
race interactions and black						4.96 (0.002)
race interactions only						0.27 (0.766)
difference in predicted probability of denial, white vs. black (percentage points)	8.4%	6.0%	7.1%	6.6%	6.3%	6.5%

These regressions were estimated using the $n = 2380$ observations in the Boston HMDA data set described in Appendix 11.1. The linear probability model was estimated by OLS, and probit and logit regressions were estimated by maximum likelihood. Standard errors are given in parentheses under the coefficients, and p-values are given in parentheses under the F-statistics. The change in predicted probability in the final row was computed for a hypothetical applicant whose values of the regressors, other than race, equal the sample mean. Individual coefficients are statistically significant at the *5% or **1% level.

the last evaluation

- Both for the LPM models as for the Probit & Logit models we have to consider threats to the **Internal validity**

the last evaluation

- Both for the LPM models as for the Probit & Logit models we have to consider threats to the **Internal validity**
  - Is there omitted variable bias?

the last evaluation

- Both for the LPM models as for the Probit & Logit models we have to consider threats to the **Internal validity**
    - Is there omitted variable bias?
    - Is the functional form correct?

the last evaluation

- Both for the LPM models as for the Probit & Logit models we have to consider threats to the **Internal validity**
  - Is there omitted variable bias?
  - Is the functional form correct?
  - Is there measurement error?

the last evaluation

- Both for the LPM models as for the Probit & Logit models we have to consider threats to the **Internal validity**
  - Is there omitted variable bias?
  - Is the functional form correct?
  - Is there measurement error?
  - Is there sample selection bias?

the last evaluation

- Both for the LPM models as for the Probit & Logit models we have to consider threats to the **Internal validity**
  - Is there omitted variable bias?
  - Is the functional form correct?
  - Is there measurement error?
  - Is there sample selection bias?
  - is there a problem of simultaneous causality?

More Extension: Categoried and Limited Dependent Variables families

- Binary outcomes: LPM, logit and probit

More Extension: Categoried and Limited Dependent Variables families

- Binary outcomes: LPM, logit and probit
- Mulitnomial outcomes: No order, sucha as (multi-logit,mprobit)

More Extension: Categoried and Limited Dependent Variables families

- Binary outcomes: LPM, logit and probit
- Mulitnomial outcomes: No order, sucha as (multi-logit,mprobit)
- Ordered outcomes: Ordered Response Models(order probit and logit)

More Extension: Categoried and Limited Dependent Variables families

- Binary outcomes: LPM, logit and probit
- Mulitnomial outcomes: No order, sucha as (multi-logit,mprobit)
- Ordered outcomes: Ordered Response Models(order probit and logit)
- Count outcomes: The outcomes is a nonnegative integer or a count. (possion model)

More Extension: Categoried and Limited Dependent Variables families

- Binary outcomes: LPM, logit and probit
- Mulitnomial outcomes: No order, sucha as (multi-logit,mprobit)
- Ordered outcomes: Ordered Response Models(order probit and logit)
- Count outcomes: The outcomes is a nonnegative integer or a count. (possion model)
- Limited Dependent Varaible(Censored, Tobit and Selection Models)

More Extension: Categoried and Limited Dependent Variables families

- Binary outcomes: LPM, logit and probit
- Mulitnomial outcomes: No order, sucha as (multi-logit,mprobit)
- Ordered outcomes: Ordered Response Models(order probit and logit)
- Count outcomes: The outcomes is a nonnegative integer or a count. (possion model)
- Limited Dependent Varaible(Censored, Tobit and Selection Models)
- Time: (Duration Model)