

# Variável Dependente Binária

**Prof. Alexandre Gori Maia**  
**Instituto de Economia - UNICAMP**



## **Ementa**

Modelo de Probabilidade Linear  
Modelo Logit  
Modelo Probit

# Variáveis Dependentes Qualitativas

- Seja a variável dependente binária:  $Y_{it} = 0$  ou  $1$
- Os principais modelos com variáveis dependentes binárias são:

- 1. Função de regressão linear:** assume relação linear entre  $Y$  e  $X$ :

$$\Pr(Y = 1|\mathbf{x}) = \mathbf{x}\boldsymbol{\beta}$$

- 2. Função de probabilidade logística binária:** assume comportamento de uma função logística de probabilidade acumulada:

$$\Pr(Y = 1|\mathbf{x}) = \Lambda(\mathbf{x}\boldsymbol{\beta})$$

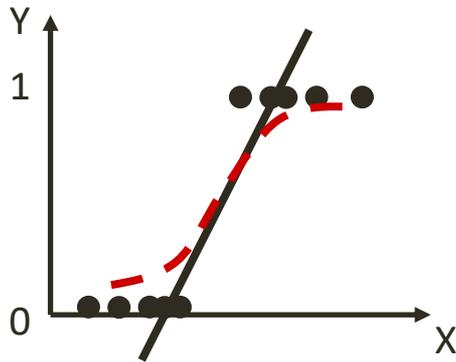
- 3. Função probabilidade normal acumulada:** assume comportamento de uma função normal acumulada;

$$\Pr(Y = 1|\mathbf{x}) = \Phi(\mathbf{x}\boldsymbol{\beta})$$

# Função de Regressão Linear

- A função de regressão linear (ou modelo de probabilidade linear) assume relação linear entre a variável binária e regressores (vetor  $\mathbf{x}$ ):

$$Y_i = \mathbf{x}_i\boldsymbol{\beta} + e_i \quad \Rightarrow \quad \Pr(Y = 1|\mathbf{x}) = \mathbf{x}\boldsymbol{\beta}$$



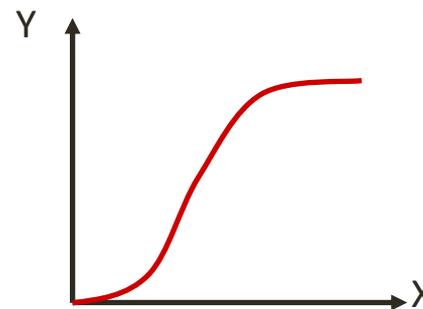
Principais limitações dos modelos de probabilidade linear:

- 1) **Ausência de normalidade na distribuição dos erros:** a distribuição dos erros em torno da reta de regressão não seguirá uma distribuição normal.
- 2) **Heterocedasticidade:** A dispersão dos erros será maior nos extremos de  $X$ .
- 3) **Forma funcional inadequada:** Assume efeitos marginais constantes e, conseqüentemente, pode haver probabilidades ajustadas com valores negativos ou superiores a 1.

# Função de Regressão Logística (logit)

- O modelo logit define uma função logística acumulada à probabilidade de  $Y_i$  assumir o valor 1:

$$P_i = \Pr(Y_i = 1 | \mathbf{x}_i) = \frac{e^{\mathbf{x}_i \boldsymbol{\beta}}}{1 + e^{\mathbf{x}_i \boldsymbol{\beta}}}$$



- Também é possível representar o modelo logit pela função linear:

$$\ln\left(\frac{P_i}{1-P_i}\right) = \mathbf{x}_i \boldsymbol{\beta}$$

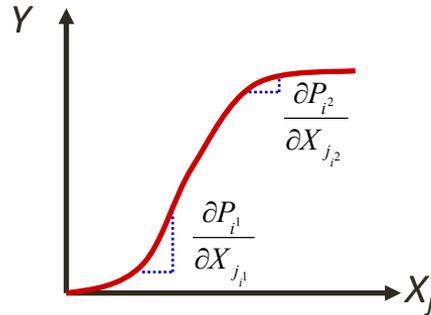
Onde :  $\frac{P_i}{1-P_i}$  representa o *odds*, ou a chance de sucesso em relação ao fracasso

$\ln\left(\frac{P_{it}}{1-P_{it}}\right)$  é chamado de **logit**.

- Uma maneira simples de interpretar os coeficientes em  $\boldsymbol{\beta}$  é como efeito marginal no logit, ou seja, na variação no log das chances (*odds*) de sucesso ( $Y=1$ );

# Efeitos Marginais – Modelo Logit

- O efeito marginal de  $X_j$  sobre a probabilidade  $P$ , ou seja,  $\partial P / \partial X_j$ , será diferente para cada ponto da regressão:



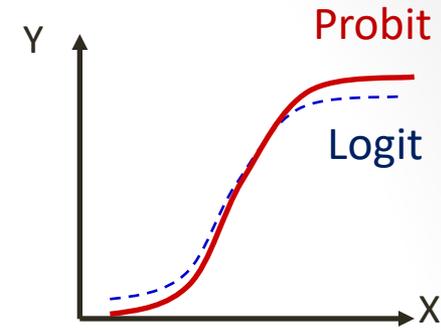
- Uma alternativa simplificadora é estimar a média dos efeitos marginais para o conjunto  $k$  de valores observados de  $X_j$  (AME);

$$\frac{\partial P}{\partial X_j} = \sum_{i=1}^k \frac{\partial P_i}{\partial X_{j_i}} / k$$

# Função Normal Acumulada - Probit

- O modelo Probit ajusta uma **função normal acumulada** à probabilidade  $P$  de  $Y$  assumir o valor 1. A especificação é mais complicada:

$$P(Y_i = 1) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\mathbf{x}_i\boldsymbol{\beta}} e^{-z^2/2} dz$$



- A compreensão do modelo Probit e a interpretação de seus coeficientes  $\boldsymbol{\beta}$  é facilitada com a representação de uma variável latente (não observável)  $Y^*$ , em que:

$$P(Y_i = 1) = \Phi(\mathbf{x}_i\boldsymbol{\beta}) \quad \Rightarrow \quad Y_i^* = \mathbf{x}_i\boldsymbol{\beta}$$

- Ou seja, os coeficientes em  $\boldsymbol{\beta}$  representam os efeitos marginais dos regressores sobre a variável latente;
- Pode-se ainda estimar o efeito marginal médio de  $X_j$  sobre a probabilidade  $P$  ( $\partial P / \partial X_j$ ), de maneira análoga ao modelo logit;

# Exercício

- 1) O arquivo *Data\_MFA.csv* contém uma amostra de domicílios com dados sobre o acesso ao programa Mas Famílian en Accion (MFA) e percepção de pobreza ([MORALES MARTINEZ, D.; MAIA, A. G. The impacts of cash transfers on subjective wellbeing and poverty: The case of Colombia. International Journal of Family and Economic Issues, v.39, p. 616-633, 2018](#)):
  - a) Estime e analise o impacto do programa MAF sobre a percepção de pobreza usando uma função de probabilidade linear;
  - b) Estime e analise o impacto do programa MAF sobre a percepção de pobreza usando uma função logit;
  - c) Estime e analise o impacto do programa MAF sobre a percepção de pobreza usando uma função probit;
  - d) Compare a qualidade dos ajustes;