



## Workshop4. Logit & Probit Models

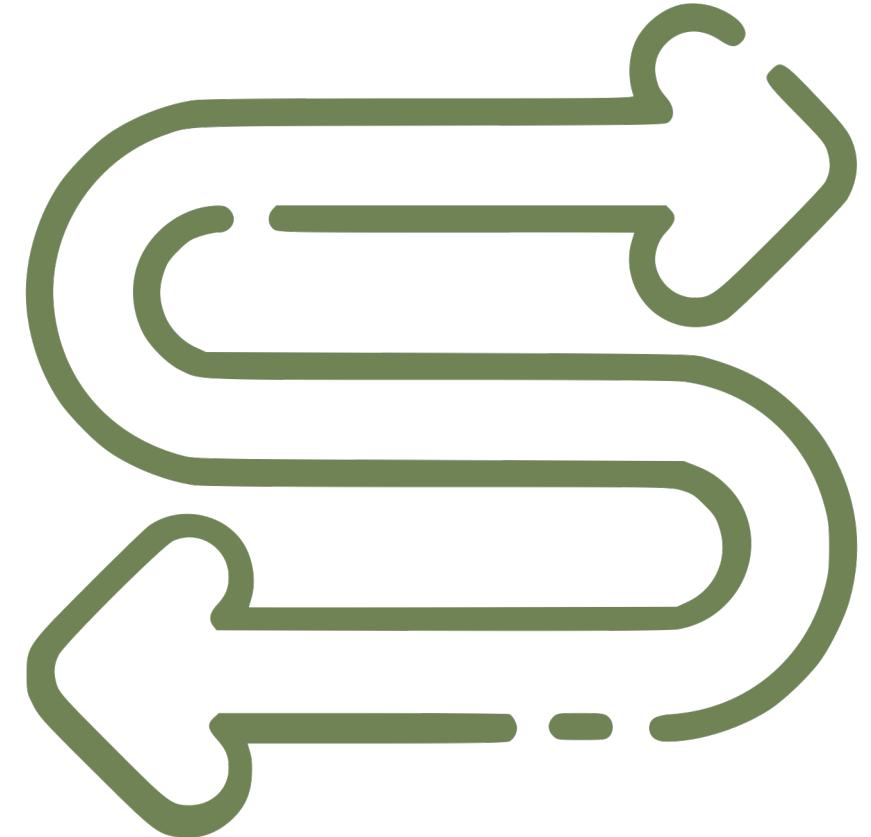
UNAM – FE Econometrics I

Esp. Humberto Acevedo  
Assistant. Emilio Sandoval

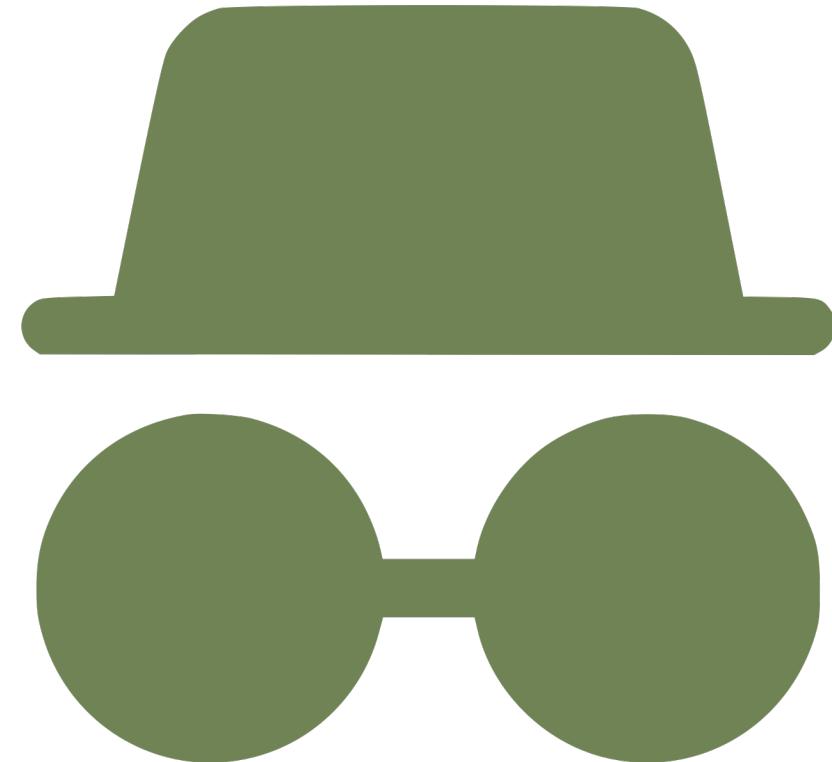
April 5, 2022

Up to now, we have studied the regression model, which means we have assumed a **unidirectional relationship exists.**

This unidirectional relationships in turn assumes there is **one dependent variables** and a **set of independent** that explicate the latest.



Implicitly, there is one more assumption:



Have you noticed the data type  
for dependent variable?



Which kind of type is it?

Dependent variable has always  
been quantitative!



If independent variable is  
**quantitative**...

Objective: to estimate expected value  
given the regressors

Examples:

- Simple Lineal Regression
- Multiple Linear Regression
  - MC2E

If independent variable is  
**qualitative**...

Objective: *to find probability that an event happens*

Examples:

- Linear Probability Model (LPM)
  - LOGIT
  - PROBIT
  - TOBIT



## Differences

This models (qualitative dependent variable) allow to explain decisions taken by an individual from a set of independent variables

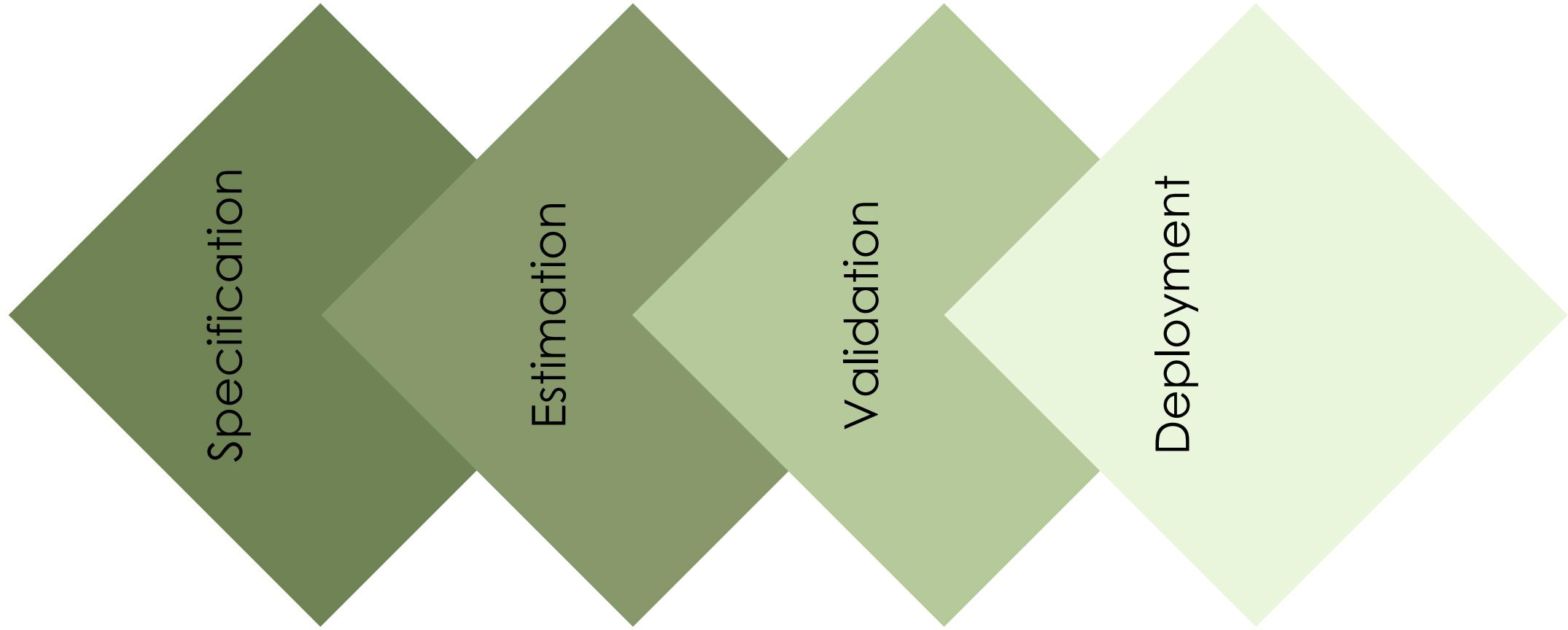
Given the number of alternative decisions that an individual can take, we differentiate between two types of models:

Binary election

Multiple election

**Daniel L. McFadden**, American economist and cowinner (with James J. Heckman) of the 2000 Nobel Prize in Economic Sciences for his development of theory and methods used in the analysis of individual or household behaviour, such as understanding how people choose where to work, where to live, or when to marry

Differences



## Differences

### Specification

Definition of endogenous variable and functional form

### Estimation

### Validation

### Deployment

Differences

Specification

Estimation

Calculation of  
parameters

Validation

Deployment

## Differences

Specification

Estimation

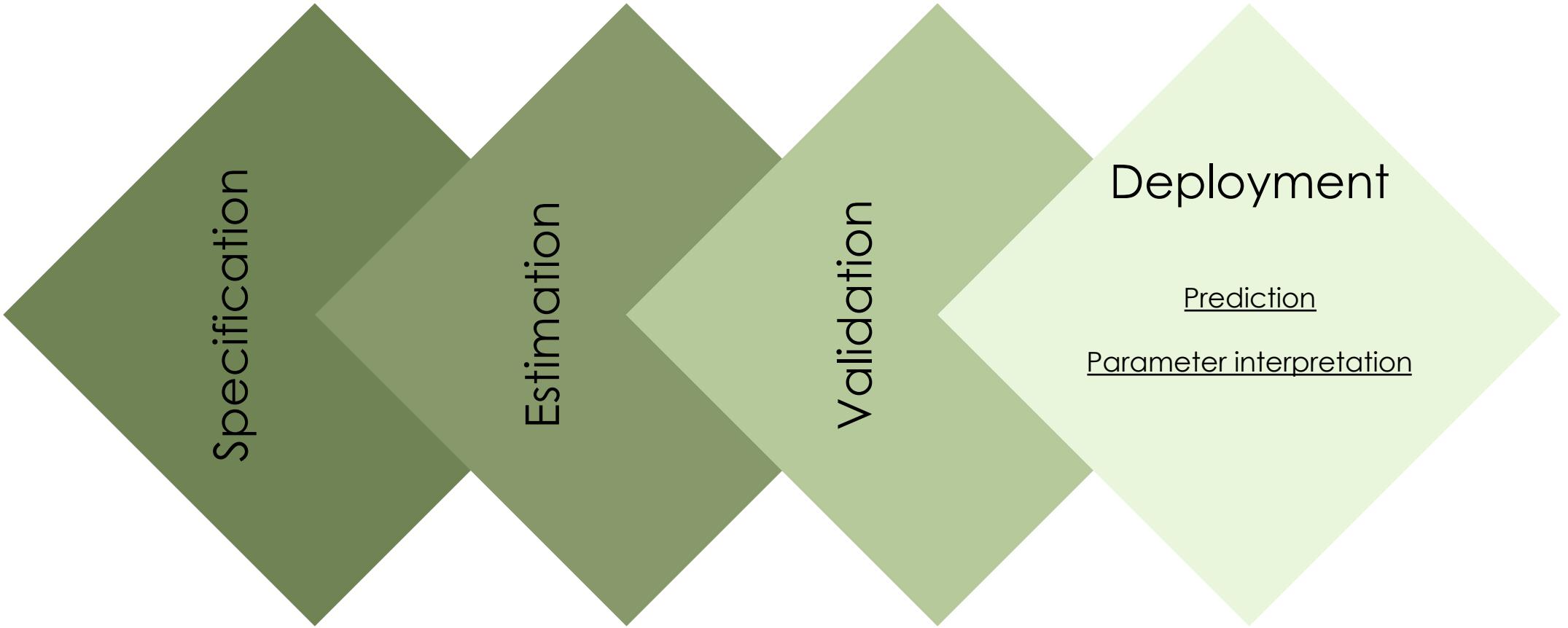
Validation

Individual: check for  
significative variables

Joint: check if full model is  
acceptable

Deployment

Differences



## Binary

Binary election models are those which dependent variable is dichotomous: it can take just two possible outcomes, no less no more

Individuals can choose between two possible and alternative options.

These options are **mutually exclusive!**

$$Y_i = \begin{cases} 1 & \text{yes} \\ 0 & \text{no} \end{cases}$$

## Multiple

Multiple election models are those in which dependent variable can refer to more than categories

Individuals can choose between two or more possible options.

These options are mutually exclusive too!

$$Y_i = \begin{cases} 0 & \text{if car } A \text{ is bought} \\ 1 & \text{if car } B \text{ is bought} \\ 2 & \text{if car } C \text{ is bought} \\ 3 & \text{if no car is bought} \end{cases}$$

Linear Probability Model (LPM) is the simplest model we can use. Nevertheless, given its simplicity, it has many disadvantages

It assumes that the relationship among variables is linear

$$Y_i = \beta_1 + \beta_2 X_2 + \cdots + \beta_k X_{ki} + u_i, \quad i = 1, \dots, N$$

LOGIT and PROBIT models are the same, but quite unique in distribution

LOGIT

Accumulative Logistic Distribution

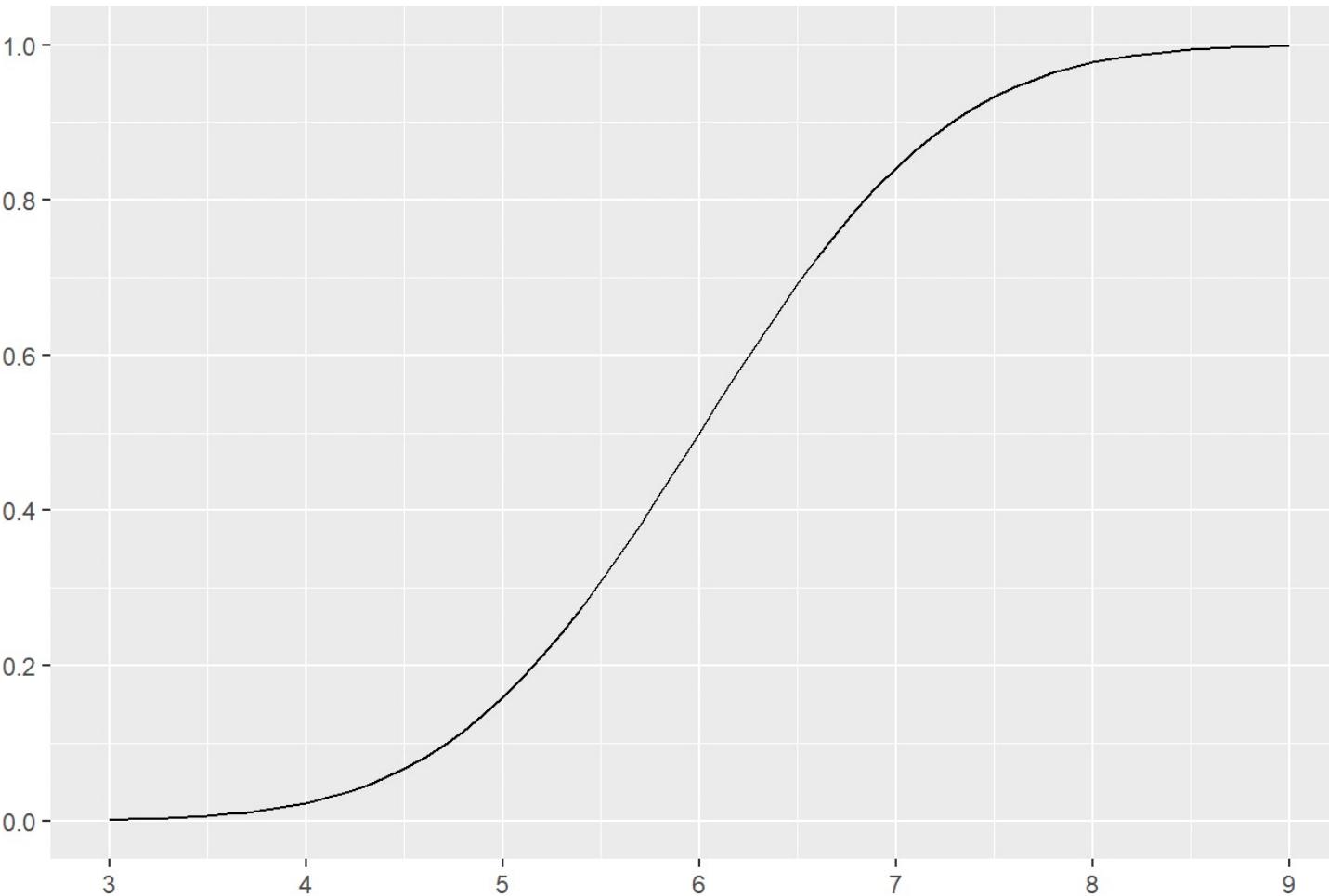
$$Y_i = \int_{-\infty}^{\alpha+\beta X} \frac{1}{(2\pi)^{1/2}} e^{-\frac{s^2}{2}} ds + u_i$$

PROBIT

Accumulative Standard Normal Distribution

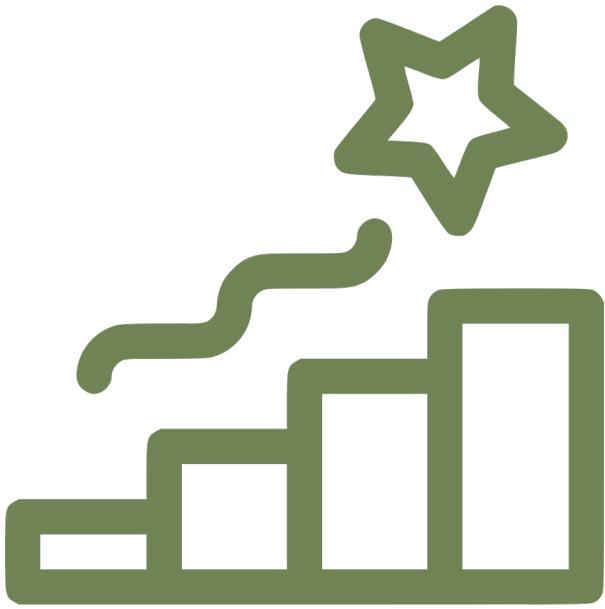
$$Y_i = \frac{1}{1 + e^{-\alpha - \beta_k X_{ki}}} + u_i = \frac{e^{\alpha - \beta_k X_{ki}}}{1 + e^{\alpha + \beta_k X_{ki}}} + u_i$$

## Logit/Probit

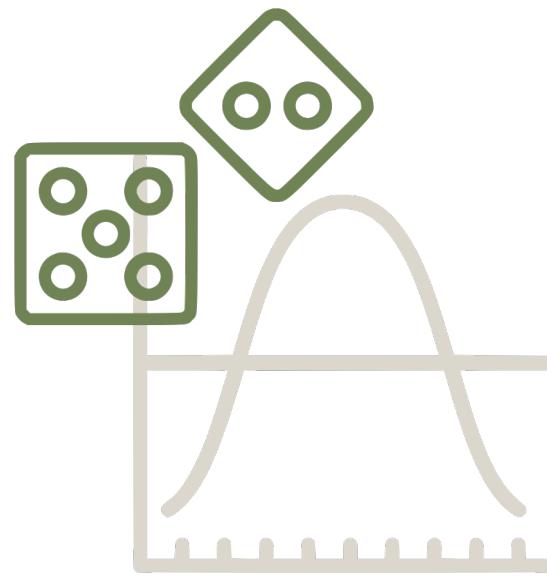


These models take into account **conditional probability**, (occurrence of  $Y$  given  $X$ ) so output must be between 0 and 1

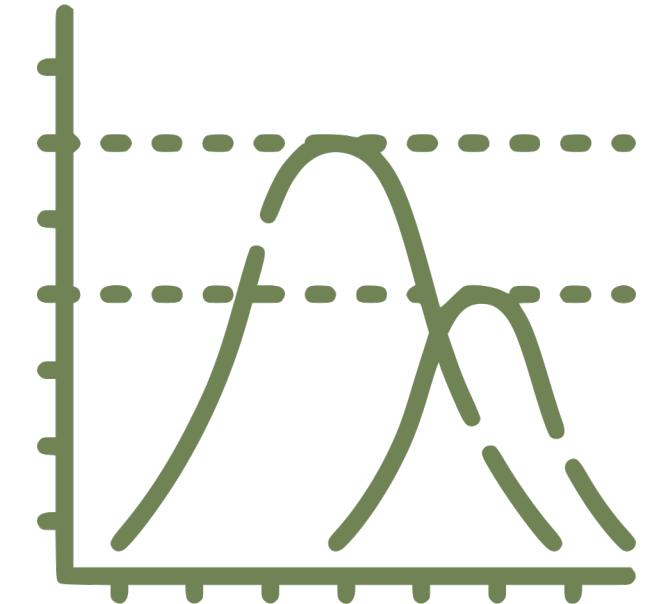
## Logit



Logit is the most used and simplest discrete model. Its popularity is based on its equation and its explainability



MCO techniques are not longer suitable (probs must be between 0 and 1, not just 1 as usual)



We use Maximum Likelihood

## References

- **Salvatore, D., & Sarmiento, J. C.** (1983). *Econometría* (No. HB141 S39). McGraw-Hill.
- **Gujarati, D. N.** (2009). Basic econometrics. Tata McGraw-Hill Education.
- **Wooldridge, J.M.** (2016). *Introductory Econometrics*, Cengage Learning, 6<sup>th</sup> edition.
- **CFA Institute** (2020), “Level I, Volume 1, 2020, Ethical and Professional Standards and Quantitative Methods; Reading 7: Statistical Concepts and Market Returns”, pp. 422-430