



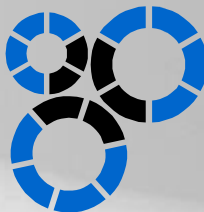
# **Machine Learning**

## **CSE - 465**

**Lecture - 02**

# Outline

- Model
  - Deterministic Model
  - Probabilistic Model
- Regression Model
- Steps for Regression Modeling
  - Types of Regression Modeling
- Linear Regression



# ◉ What is a model?

- Model
  - Used to describe relationship between variables
- Types
  - Deterministic Model
  - Probabilistic Model

# Deterministic Model

- A **deterministic model** is a system in which no randomness is involved in the development of future states of the system
- A deterministic model will always produce the same output from a given starting condition or initial state
- A **deterministic algorithm** is an algorithm which, given a particular input, will always produce the same output, with the underlying machine always passing through the same sequence of states.

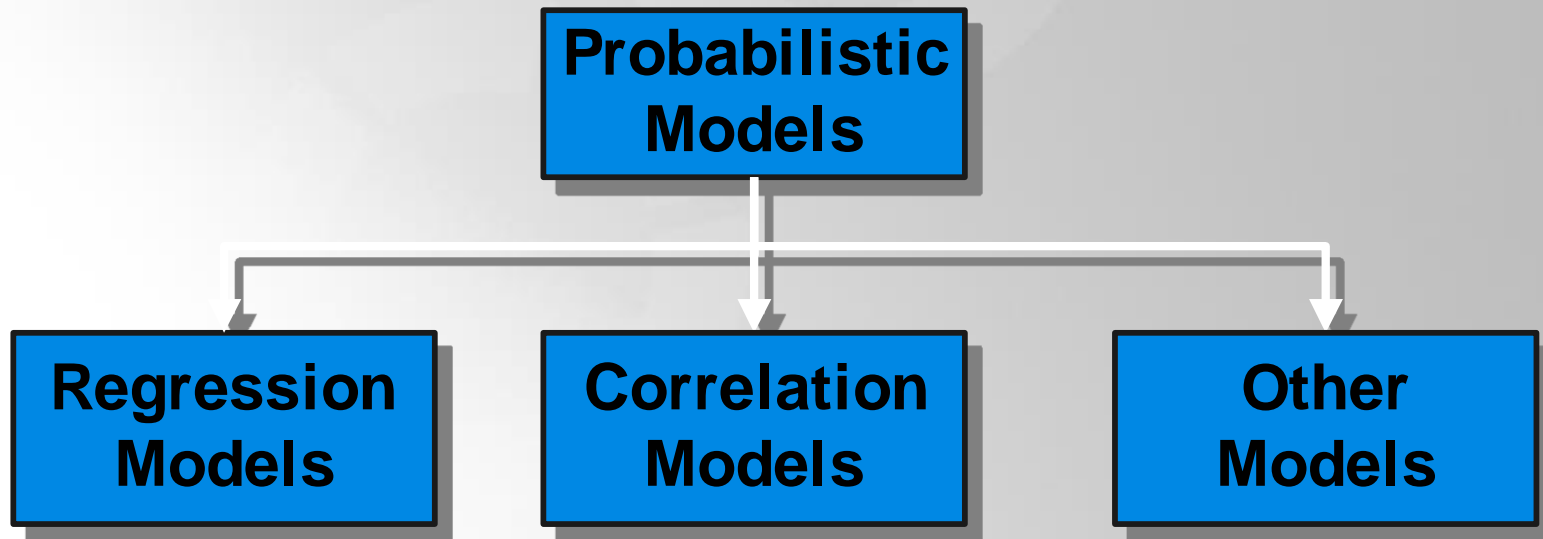
# ◉ Probabilistic Model

- Probabilistic models incorporate random variables and probability distributions into the model of an event
- While a deterministic model gives a single possible outcome for an event, a probabilistic model gives a probability distribution as a solution
- There's nearly always an element of randomness to take into account
- For example, life insurance is based on the fact we know with certainty that we *will* die, but we don't know *when*

# Example

- The electricity that a family consumes for a period of two months and the amount it pays for this consumption are related by a deterministic relationship. If the family consumes a higher amount of electricity for a period of two months than it usually does, the amount due will be higher than usual. This is a determinist, causal relationship
- However, the linear relationship cannot describe the linear stochastic dependence of variables  $X$  and  $Y$  because if for example, if the variable  $X$  is the amount of fertilizer and variable  $Y$  is the yield of a crop, the corresponding values of  $Y$  will be different in different iterations due to factors such as temperature, rainfall, or soil quality, which also affect production

# ◉ Types of Probabilistic Model

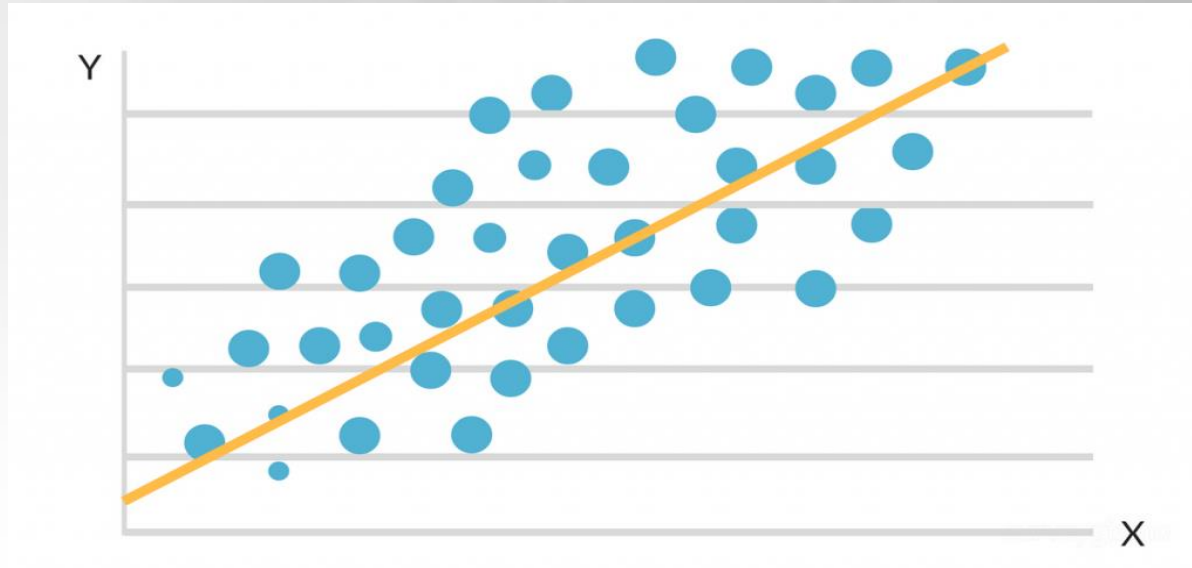


# Regression Model

- A regression model is a statistical model for estimating the relationships between a dependent variable and one or more independent (explanatory) variables
- **Dependent Variable:** This is the main factor that we're trying to understand or predict. Also called outcome variable
- **Independent Variables:** These are the factors that we hypothesize have an impact on our dependent variable. Also known as predictors, covariates or features



# Example: Regression Model



- The regression line represents the relationship between independent and dependent variable.
- The formula for a regression line might look something like  **$Y = 100 + 7X + \text{error term}$** .

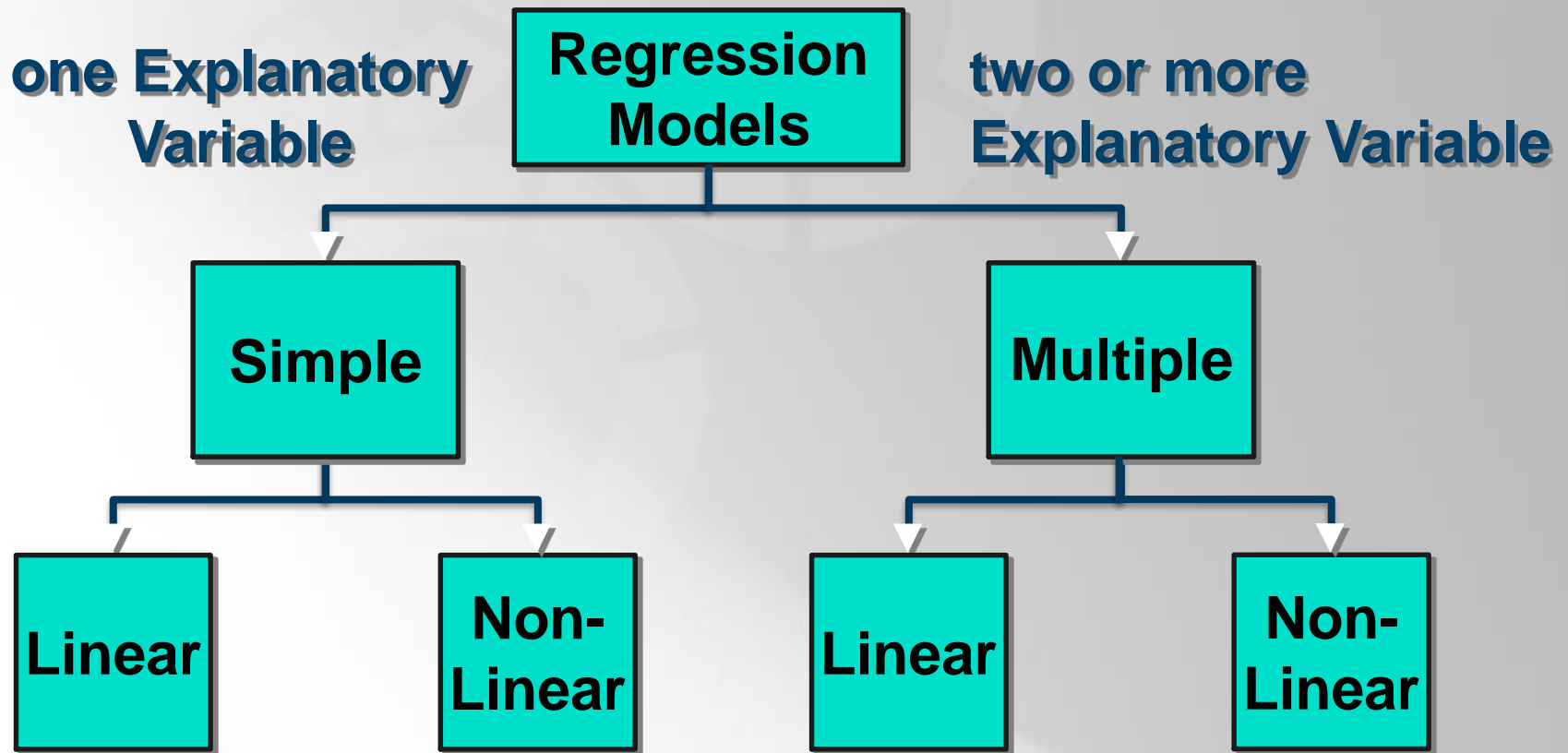
# ◉ Example: Regression Model

- It tells us that if there is no “X”, then  $Y = 100$ . If X is our increase in ticket price, this informs us that if there is no increase in ticket price, event satisfaction will still increase by 100 points.
- Regression lines always consider an error term because in reality, independent variables are never precisely perfect predictors of dependent variables. This makes sense while looking at the impact of ticket prices on event satisfaction — there are clearly other variables that are contributing to event satisfaction outside of price

# ◉ Steps for Regression Modeling

1. Hypothesize Deterministic Component
  - Estimate Unknown Parameters
2. Specify Probability Distribution of Random Error Term
  - Estimate Standard Deviation of Error
3. Evaluate the fitted Model
4. Use Model for Prediction & Estimation

# ○ Types of Regression Model



# Linear Regression Model

- Relationship between variables is a linear function

The diagram illustrates the Linear Regression Model equation:  $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ . Each term in the equation is labeled with a blue arrow pointing to it:

- Population Y-Intercept** points to  $\beta_0$ .
- Population Slope** points to  $\beta_1$ .
- Random Error** points to  $\varepsilon_i$ .
- Dependent Variable** points to  $Y_i$ .
- Independent Variable** points to  $X_i$ .

# Population & Sample Regression Models

## Population

**Unknown  
Relationship** 😊

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

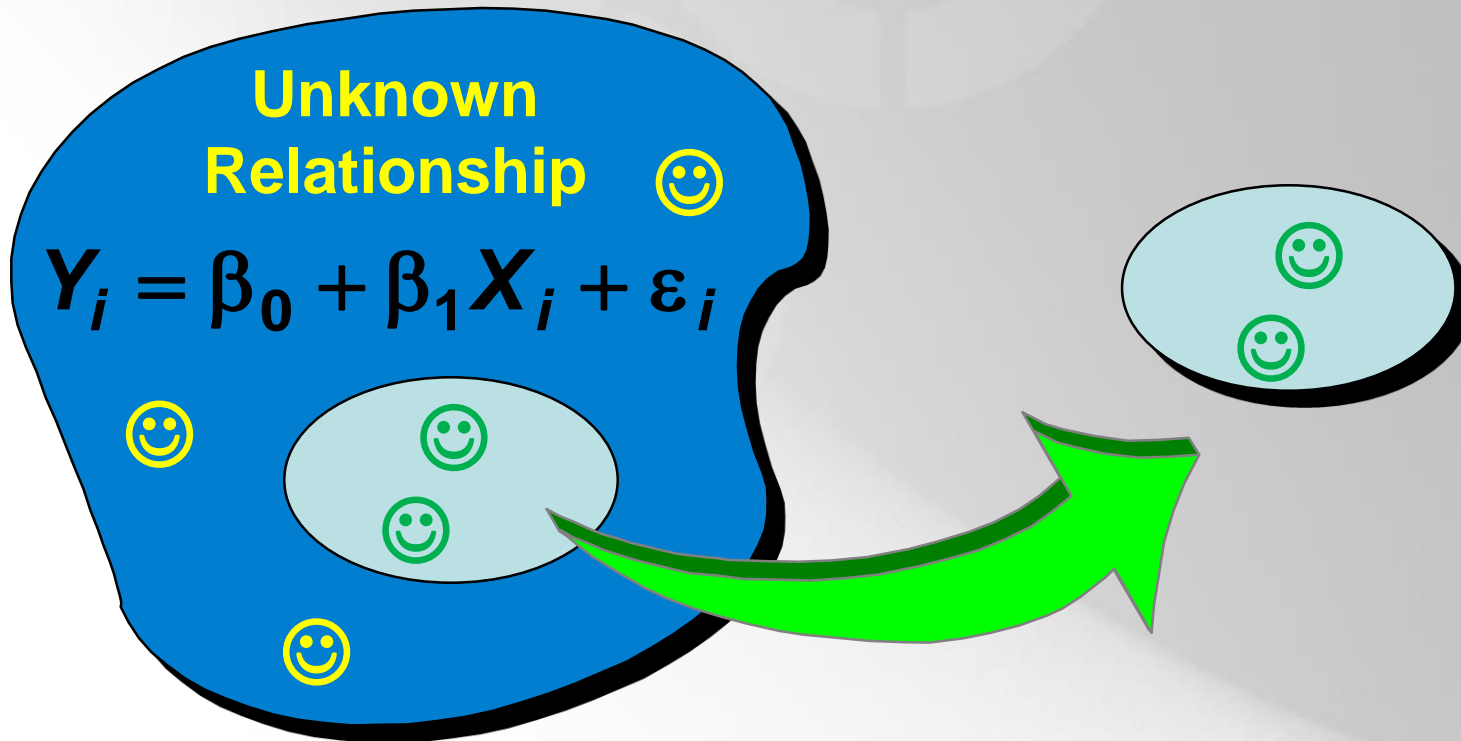




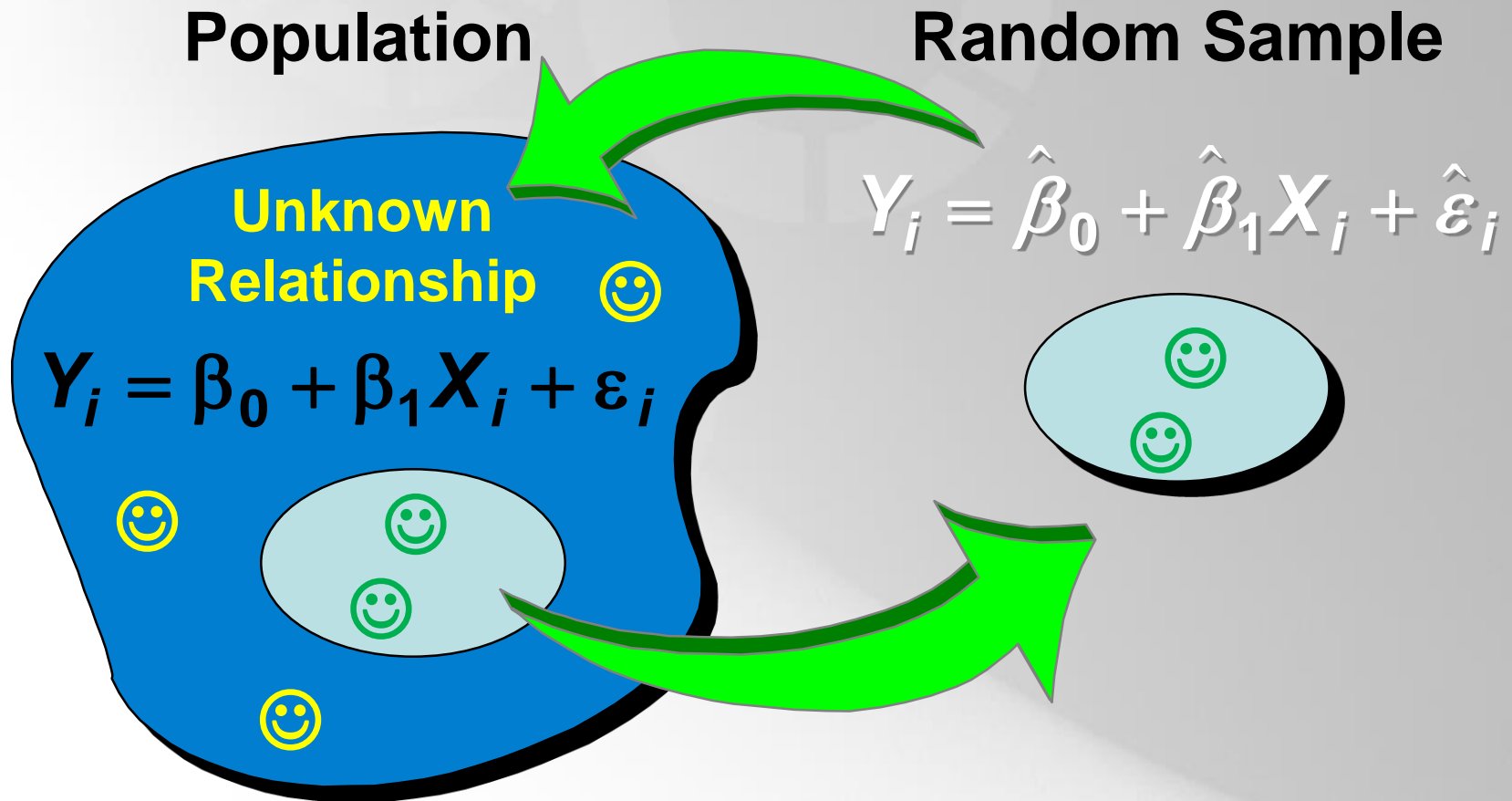
# Population & Sample Regression Models

**Population**

**Random Sample**

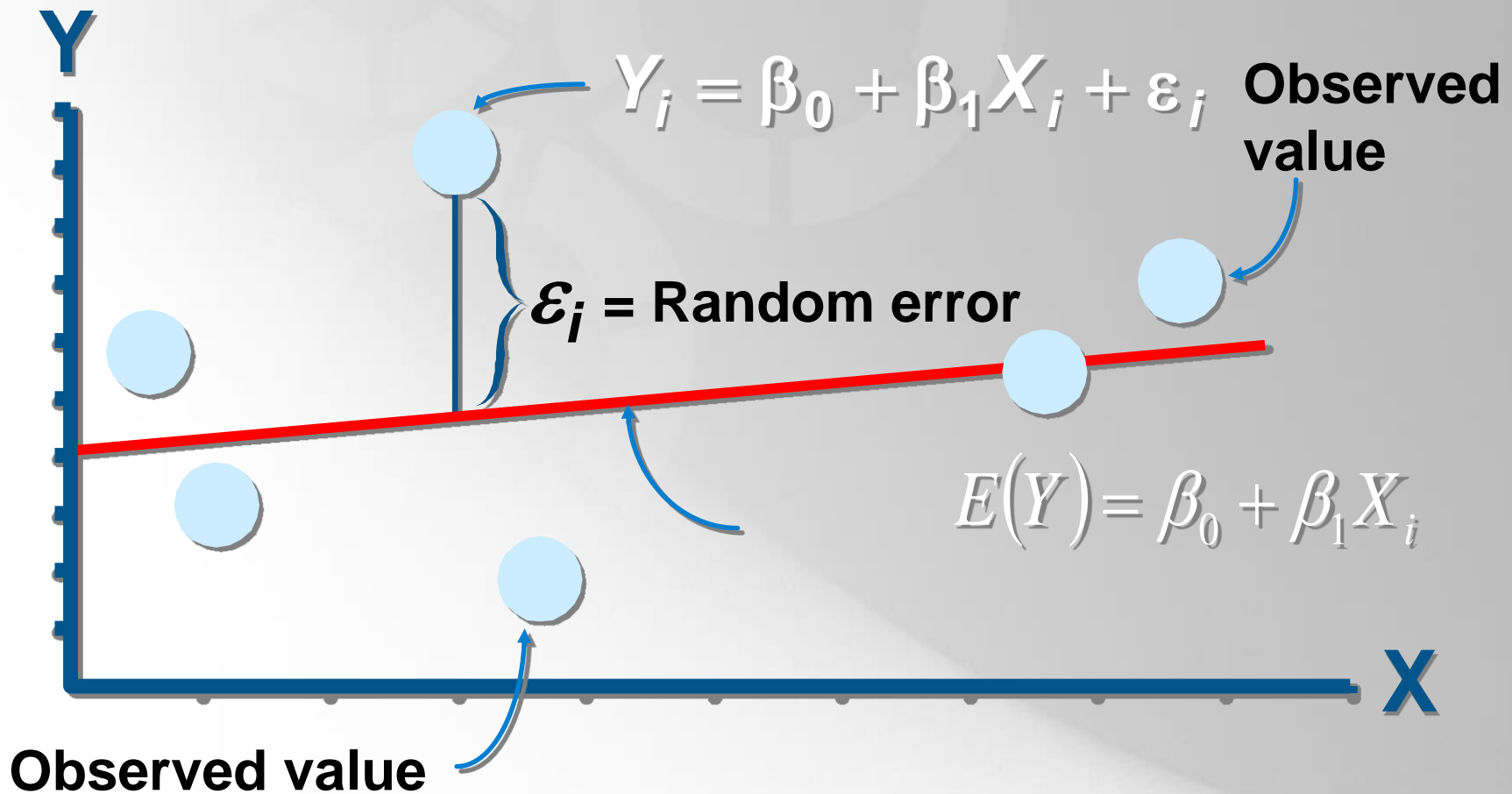


# Population & Sample Regression Models

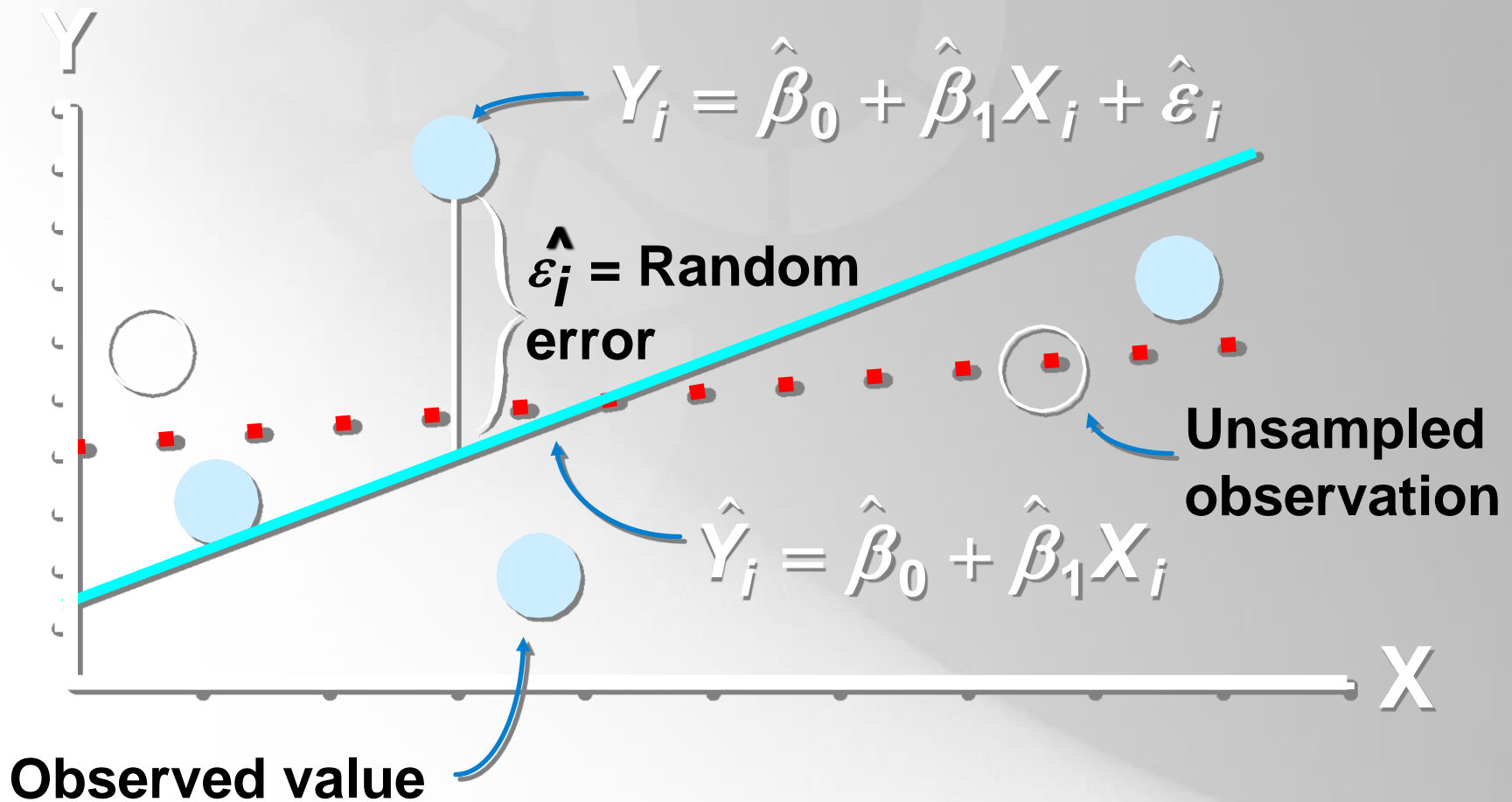




# Population & Sample Regression Models



# Population & Sample Regression Models





# Thank You

