# Statistical inference and asymptotic theory

Juan Felipe Acevedo Pérez

Econometría Avanzada I
Maestría en Ciencias Económicas
Universidad Nacional de Colombia

1S2023

# Table of Contents

# Tabla de contenido

# Statistical inference

One of the fundamental concepts in econometrics is statistical inference, which involves using data to make inferences about population parameters. In other words, we use a sample of data to estimate characteristics of a larger population. This is important in econometrics because we often cannot observe the entire population, so we have to rely on a sample of data to draw conclusions about the population as a whole.

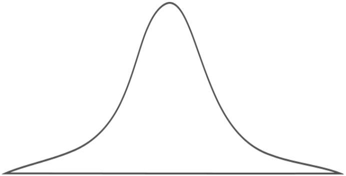
**Keeping, 1995**

In everyday life we are again and again faced with the necessity of making decisions. Many of these are trivial, some may be serious, but almost always there is an element of uncertainty about the wisdom of the decision we make. Scientists in their regular work have a similar problem.

# Basics of Probability

Probability theory is the mathematical study of uncertainty. It provides a framework for understanding the likelihood of events occurring, and it is a fundamental concept in many fields, including econometrics.

**A probability distribution** is a function that assigns probabilities to each possible value of a random variable. The most common probability distribution is the normal distribution, which is a bell-shaped curve that describes many natural phenomena.



In econometrics, we often use probability theory to *estimate parameters of a model and make predictions about future economic trends*. For example, we might use a probability distribution to model the behavior of stock prices or interest rates.

By understanding the basics of probability theory, we can make more informed decisions based on the likelihood of different outcomes occurring. We can also use probability theory to develop more accurate models of economic behavior and make more accurate predictions about future economic trends.

A **sample space S** is the set of all possible outcomes of a random experiment. The elements that make up the sample space are called events or outcomes.

A **random variable (RV)** is a function whose domain consists of the elements of the sample space, and whose range is the set of all real numbers. In other words, it is a function from the sample space to the real numbers. Alternatively, X is a random variable if for every real number x there exists a probability such that the value assumed by the random variable does not exceed x, denoted by $P(X \leq x)$ or $F_X(x)$, and called the cumulative distribution function (CDF) of X.

**The cumulative distribution function (CDF)** is an important concept in statistics and econometrics, as it describes the probability that a random variable takes on a value less than or equal to a given point. The CDF has several mathematical properties, which are useful for understanding its behavior and properties.

1. Monotonicity: The CDF is a non-decreasing function, meaning that as x increases, F(x) also increases.

$$F(x_1) \leq F(x_2) \quad \text{if} \quad x_1 \leq x_2$$

2. Bounds: The CDF is bounded between 0 and 1, inclusive.

$$0 \leq F(x) \leq 1$$

3. Continuity: The CDF is a continuous function.

$$\lim_{x \to c} F(x) = F(c)$$

4. Limits: The CDF has limits at positive and negative infinity.

$$\lim_{x \to -\infty} F(x) = 0 \quad \text{and} \quad \lim_{x \to \infty} F(x) = 1$$

The probability density function (PDF) is a concept in statistics and econometrics that describes the likelihood of a continuous random variable taking on a particular value. The PDF has several mathematical properties, which are useful for understanding its behavior and properties.

1. Non-negativity: The PDF is a non-negative function.

$$f(x) \geq 0$$

2. Normalization: The integral of the PDF over its entire support is equal to one.

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

The derivative of the cumulative distribution function of $X$, denoted by $f_X(x)$, is a non-negative function called the probability density function (PDF) of $X$. Thus, when $X$ is a continuous random variable:

$$F_X(x) = \int_{-\infty}^{x} f_X(u)du,$$
$$f_X(x) = \frac{d}{dx}F_X(x) = F'_X(x).$$

Moreover, the integral of the PDF over its entire support is equal to one:

$$\int_{-\infty}^{\infty} f_X(x)dx = 1.$$

3. Local maxima: The PDF has local maxima at the points where the PDF is equal to zero.

$$f'(x_0) = 0 \quad \text{and} \quad f''(x_0) < 0$$

4. Monotonicity: The PDF is a non-increasing function, meaning that as x increases, f(x) decreases.

$$f(x_1) \geq f(x_2) \quad \text{if} \quad x_1 \leq x_2$$

## Examples

Let $X$ be the random variable "duration" (in units of 100 hours) of a certain electronic device. Suppose that $X$ is a continuous random variable and that the probability density function (PDF) $f$ is given by:

$$f(x) = \begin{cases} 2e^{-2x}, & x > 0, \\ 0, & \text{otherwise.} \end{cases}$$

The probability that such a device lasts more than one unit of time (100 hours) is:

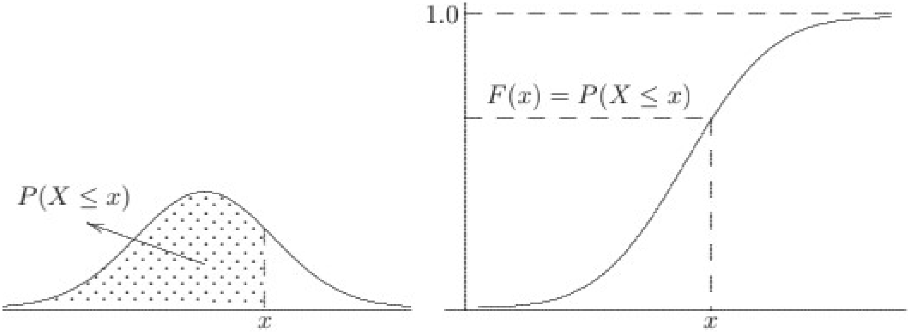
$$\Pr(X > 1) = \ ?$$

# Gaussian distribution

The normal distribution, also known as the Gaussian distribution, is a probability distribution that is widely used in statistics and econometrics to model many natural phenomena. It is a bell-shaped curve that is symmetric around the mean, and it is characterized by two parameters: the mean ($\mu$) and the standard deviation ($\sigma$).

The probability density function of the normal distribution is given by:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



Función de densidad normal