

# Notes on Statistical Estimation v. Identification in Causal Inference

by Carlos Rubio

While statistical estimation and causal identification are terms and important aspects of causal effect estimation, they refer to entirely different concepts.

1. **Statistical estimation** (also known as *statistical inference*) is the process of approximating the parameters of a population through the use of statistical methods; namely, *study design, probability theory and mathematical distributions*. Through *statistical estimation*, a person **can deduce or at least approximate certain truths about the population**: what characteristics the individuals possess, and the tendencies and differences among them. It can be used to determine further truths and relationships about those characteristics. For example, it can help determine the degree of association between a subset of characteristics, or in other words,
2. **Identification in causal inference** is both technique and process used to estimate the actual causal effect of one characteristic of the population (namely, a cause, sometimes named the treatment) over another characteristic (called the outcome). In this process, a **subset of characteristics is identified and condition upon** in order to account and remove the bias introduced by its elements. Such subset is the set of all confounding variables that generate indirect association between the cause and the outcome and generate confusion about the actual effect or impact the cause has over the effect.

Both techniques are used in combination in order to **estimate the causal effect of an independent variable and an outcome**. Through *statistical estimation* we can **find a numerical value that is representative of the effect**, while **identification ensures that such numerical value is truly representative**.