Statistical Thinking

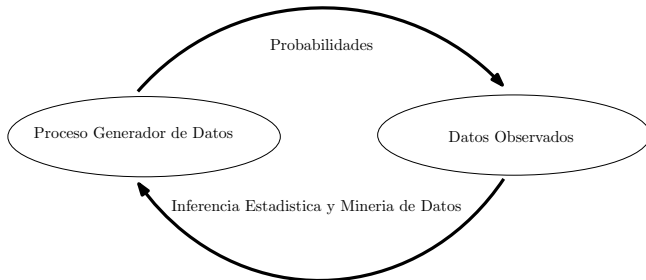Felipe José Bravo Márquez

October 1, 2020

Motivation

- Collections of data are stored on a daily basis. For example: The Web, e-commerce, transactional data.
- Computers are becoming cheaper and more powerful.
- Analyzing this data allows to find hidden patterns.
- A good use of data can bring business benefits. Ex: customer segmentation, demand prediction.

Probability and Statistics

- Probability is the language of uncertainty that is also the basis for statistical inference.
- The problem studied in probabilities is: given a data generating process, which are the properties of the outputs?
- The problem studied in statistical inference, data mining and machine learning is: given the outputs, what can we say about the process that generates the observed data?

Bilbiografía I

Venables, William N., David M. Smith, and R Development Core Team. *An introduction to R.*, 2002.