Symmetry and Sufficiency: Practical Assumptions for Data Analysis

Terris Becker, Dr. Bee Leng Lee

November 22, 2019

Abstract

Thomas Bayes laid the foundation for Bayesian inference by making judgments about the structure of observed data rather than claiming that data must initially conform to a particular distribution. Since then, Bayesian statisticians have found necessary and sufficient conditions, which are based on symmetry and sufficient statistics, under which the joint probability distribution of observed data can be represented as mixtures of independent random variables from named distributions. Exchangeability is one example of a particular form of data symmetry that leads to Bruno de Finetti's 0-1 representation theorem. Exponential, geometric, and Gaussian representation theorems are also discussed in detail as well as their applications to business and economics. Most of the theoretical notions of symmetry are shown to have easy, real-world interpretations so that statisticians can rely on their judgment and proceed with analyses. The sufficient condition of spherical symmetry, which leads to the Gaussian representation, is often too difficult to justify using subjective assessments. In this case, a Monte Carlo algorithm is presented that checks observations for the presence of spherical symmetry. Additional conditions about posterior expected value and experimental design are shown to pin down prior distributions to specific conjugate families.