Bayesian inference

Arnaud Dufays
Université Laval, département d'économique

Bayesian inference stands for the Bayesian statistics method for estimating parameters of a model. It is an alternative to classical inference. While the latter measures the frequency of an event, the former relies on subjective statistics as formalized below. To operate, Bayesian inference extensively applies Bayes' theorem that we first shortly review.

Bayes' theorem

Bayes' theorem is mathematically stated as

$$P(H|E) = \frac{P(E|H)P(H)}{P(E)}. \tag{1}$$

Each term of the theorem should read as follows :

- $P(H|E)$ is the probability of an hypothesis conditional to observe the evidence E. This probability is commonly called the "posterior probability" of H given the event E.
- $P(E|H)$ denotes the probability of evidence E to occur conditional to the fact that the hypothesis is true.
- $P(H)$ stands for the probability of the hypothesis regardless of the evidence E. This probability is typically called the "prior probability" as it reflects the state of knowledge about hypothesis H before observing the evidence E.
- $P(E)$ is the probability of event E regardless the hypothesis H.

Bayes' theorem stands for a rational rule for updating our belief about an hypothesis or for uncovering hidden causes from a set of consequences. To make the theorem intuitive enough, let us assume that you have been tested positive for a rare disease. Concerned, you decide to quantify your probability of being truly sick. To do so, you first look into the medical test and learn that the medical trial is 100% reliable if you are ill which is not comforting. More hopefully, the test is 99% reliable if you are in good health, meaning that over 100 healthy individuals, one test ends up positive on average. In addition to that, you question the disease extent and determine that 0.1% individuals of the population are affected. Bayes' theorem produces the precise probability of being sick. In fact, this probability of being sick (i.e. hypothesis H) conditional to a positive medical test (i.e. evidence E) stands exactly for the Bayes' theorem output. To compute it, we need the probability of a positive test if the individual has the disease ($P(E|H) = 1$), the probability of catching the disease ($P(H) = 0.001$) and the probability of getting a positive test for a randomly chosen individual in the population ($P(E) = 1(0.001) + 0.01(0.999)$). Putting altogether, Bayes' theorem tells us that the probability of being truly sick only amounts to 9.1%.

Bayesian inference

Bayesian inference consists in inferring any (statistical) abstract quantities using Bayes' theorem. It extends the scope of Bayes' theorem to any non-random value. In that framework, the hypothesis H can stand for any discrete or continuous parameters. For instance, a Bayesian statistician will use the Bayes' rule to assess a dice fairness or to quantify the correlation between the U.S. inflation and the U.S. unemployment rate.
Unwin (2004) illustrates very well the Bayesian paradigm. In his book, he suggested to infer the odds of a monotheistic God existence using Bayes' theorem. To make a parallel with the above equation, the hypothesis H stands for the existence of God and is a discrete variable. The event E denotes several observed evidence that are compatible or not with the existence of God. For instance, we observe miracles which is a clear evidence in favor of God. On the contrary, terrorism can be seen as an evidence of God non-existence. Improving Pascal's wager, Dr. Unwin ends up with a precise probability of the existence of God thanks to Bayes' theorem.

Bayesian statisticians use this rational thinking everyday to infer parameters. A clear advantage of the Bayesian inference stands for the systematic application of the Bayes' rule. At this stage, it is worth mentionning that classical statisticians find the Bayesian statistics controversial for two reasons : i) the randomness of the inferred parameter and ii) the prior probability.

- i Classical statistics assume that an unknown but fixed parameter value exists. As opposed to the Schroedinger's cat, God exists or does not exist. So, what does it mean that God exists with, let say, 60% of probability ? For Bayesian statisticians, this probability highlights that the data are not fully informative about the inferred quantity. Consequently, one can argue that there is no paradox as the posterior probability over the parameter quantifies the data uncertainty over the true value.

- ii Applying the Bayes' formula to infer the probability over a parameter value requires the prior probability of the parameter (i.e. $P(H)$). This represents the statistician's state of knowledge about the parameter value before observing the data. In his book, Unwin (2004) fixes this quantity to 50% to reflect his complete ignorance about God existence. However, a zealot or an atheist would have chosen very different prior odds. On one hand, this probability is a weakness due to its subjectivity. In fact, every prior probability delivers a different posterior probability. On the other hand, the state of knowledge about a parameter value or an hypothesis can be naturally included into the prior probability.

The main drawback of the Bayesian approach lies in the computation of the posterior probability. In particular, computing analytically the posterior probability is a complex mathematical problem for almost any useful application. This feature has limited Bayesian statistics for years. Starting from the early 1990's, Bayesian statistics boomed as new Bayesian inference methods have emerged. These new tools rely on the computational power to sample from the posterior probability rather than evaluating it. Roughly speaking, sampling from the posterior can be understood as generating the most likely values of the distribution. The prominent simulation method is called Markov-chain Monte Carlo and is abbreviated as MCMC. Nowadays other very powerful methods exist such as the Sequential Monte Carlo approach and the Approximate Bayesian Computation method. All these methods have contributed to the emergence of the Bayesian statistics. As Bayesian research is sky-rocketing, it is expected that the Bayesian inference will remain a standard tool to infer parameters in the future.
Bayesian principle are literary discussed in Stone (2013). Mathematical-oriented expositions

can be found in Koop (2003) and Gelman, Carlin, Stern, Dunson, Vehtari, and Rubin (2013).

References

Gelman, A., J. Carlin, H. Stern, D. Dunson, A. Vehtari, and D. Rubin (2013): *Bayesian Data Analysis, Third Edition*, Chapman & Hall/CRC Texts in Statistical Science. Taylor & Francis.

Koop, G. (2003): *Bayesian Econometrics*. John Wiley and Sons Ltd.

Stone, J. (2013): *Bayes' Rule: A Tutorial Introduction to Bayesian Analysis*. Sebtel Press.

Unwin, S. D. (2004): *The Probability of God: A Simple Calculation that Proves the Ultimate Truth*. Three Rivers Press.