

Bayesian Inference

Chris Lauver

NPS, Southern Colorado Plateau Network

In making natural resource management decisions, managers often need to evaluate the potential outcomes from among a series of alternative management actions. But predicting these outcomes with confidence can be difficult because of numerous sources of uncertainty, most of which are beyond the managers' control. These uncertainties result from sampling and measurement errors of data used to estimate populations or to describe current ecosystem states, environmental variability of populations and ecosystem dynamics, and lack of control over the management actions (Dorazio and Johnson 2003, Prato 2005).

Bayesian statistics can help address the difficulties arising from uncertainty in predicting outcomes. It provides a way to express uncertainty (using probabilities) in parameter estimates, and to determine the probability that a hypothesis (or model) of interest is true given a set of data (Ellison 1996). For natural resource managers, Bayesian statistics offers a theoretical and coherent framework for solving management problems (Dorazio and Johnson 2003). Bayesian analysis is the only statistical approach that can be used to explicitly produce probabilities of hypotheses. In other words, it incorporates what is known (data or beliefs) about population or ecosystem dynamics with hypotheses about the consequences of management actions and generates statements about the probabilities of each outcome.

Bayesian analysis, applied iteratively, is analogous to the cycle of adaptive management. It allows you to learn from previous data, experiments, and expert opinion. After forming a specific hypothesis from existing information or beliefs, Bayesian methods can estimate the probability that the hypothesis is true and update that probability as a result of new data. When contrasted with traditional statistics, Bayesian inference has been described as well suited for, and a central ingredient in, implementing adaptive management (Ellison 2004, Prato 2005).

OVERVIEW



What is Bayesian statistics?

Although Bayes' theorem is not new, the use and application of Bayesian statistics is new to most biologists and ecologists. Bayesian methods are based on Bayes' theorem (Bayes 1763), which expressed the probability of one event given (or conditional on) another event. In applying the theorem to statistical inference for parameter estimation, observed data are used to calculate the probability of a parameter value. For hypothesis testing, Bayesian methods determine the probability of a hypothesis being true given the observed data.

Performing a Bayesian inference starts with quantifying information (known or unknown) about a parameter or hypothesis before conducting a survey or experiment. This quantity is called the prior (or prior probability distribution). Prior information can be from other surveys or experiments, models, published literature, or expert opinion. Next, sampling data are collected and a likelihood function is used to describe the probability of the data given (conditional on) a particular model or hypothesis. Finally, a posterior (or posterior probability) is calculated from the prior and the likelihood function, and reflects the probability that a model or hypothesis is true given the observed data. The posterior is the updated estimate of the prior (see Fig. 1). For decision analysis in natural resource management, Bayesian methods can be used to estimate the consequences of a decision (as posterior probabilities) based on uncertainty (the prior) and observed data (likelihood functions) (Ellison 1996).

Overview - Bayesian Inference

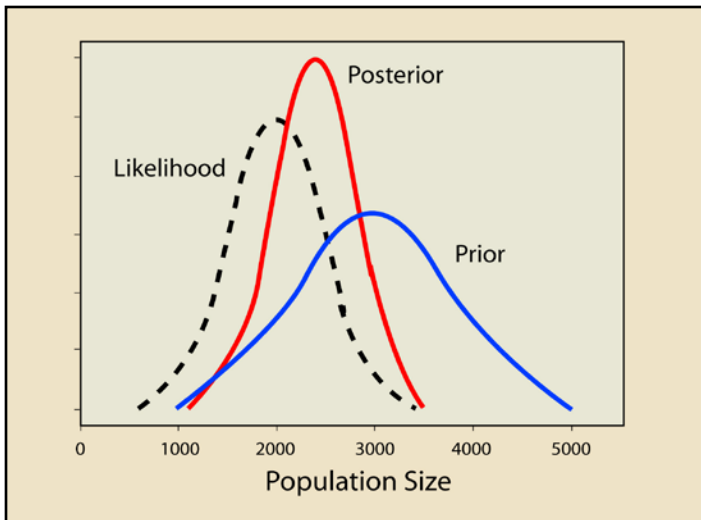


Fig. 1. An example of a Bayesian analysis (adapted from Wade 2000) where a survey was conducted, resulting in an initial abundance estimate of a given population (the prior distribution). The likelihood function describes the probability of observing these initial estimates with a particular model, such as centering the true population size at 2000. The posterior distribution is the new abundance estimate, and is intermediate to the prior and the likelihood.

an alternative source of evidence, and may provide a more intuitive approach to estimating parameters and evaluating hypotheses.

Bayesian methods are different from the standard statistics (also called frequentist statistics) that most of us have been taught. Traditional hypothesis testing produces a statement about the probability of the observed data given the null hypothesis, whereas Bayesian methods provide a direct assessment of how likely the null hypothesis is true, given

the data. For estimating parameters, frequentists rely on only the sampled data while Bayesian methods combine prior knowledge (data or beliefs) with the sampled data to provide an updated estimate. Fundamental differences between these two methods are shown in Table 1. In deciding whether to use Bayesian methods, consider these contrasting views of scientific inference in light of the available data, hypotheses, and models of interest. For more detail on comparing these two approaches, see Ellison (1996, 2004), Wade (2000), and Prato (2005).

The use of Bayesian statistics is not without its critics. A common concern is with the subjective nature of defining prior probabilities. The methods allow for use of either subjective priors (from previous data or expert knowledge) or non-informative (objective) priors that assign equal probabilities to all potential values. Some argue that using personal beliefs should not be classified as science (Dennis 1996, Quinn and Keough 2002), where others reason that it makes little sense to ignore valuable existing knowledge (Hilborn and Mangel 1997, Ellison 2004, Prato 2005). For biologists, treating parameters as random variables may be another concern. Although mean values of biological elements (e.g., the density of a population of interest) may be unknown, treating these means as fixed quantities may make more sense (Quinn and Keough 2002).

Examples of Bayesian inference

Several recent papers in ecology contrast the use of Bayesian methods with traditional statistics. For more on this, see the example papers of Anderson 1998 and Ward 2000.

Table 1. Some fundamental differences between Bayesian and traditional (frequentist) statistical inference (from Ellison 1996, 2004).

<i>Bayesian inference</i>	<i>Frequentist inference</i>
Probabilities defined as an individual's degree of belief in the likelihood of an event.	Probabilities defined by a long-run or infinite series of trials that determine relative frequencies of events.
Uses prior knowledge and sample data.	Uses only the sample data.
Treats model parameters as random variables.	Treats model parameters as estimates of fixed, 'true' quantities.
Provides a quantitative measure of the probability of a hypothesis being true in light of the available data.	Estimates the probability of the data having occurred given a particular hypothesis.

Intermountain Region

Other studies have used Bayesian belief networks (or conditional probability networks) to provide frameworks for environmental monitoring (e.g., Kiiveri et al. 2001) and to model the effects of management alternatives on wildlife populations (e.g., Lehmkuhl et al 2001, Marcot et al 2001). These networks are easy to understand, presented as an illustration of linked variables that allow the calculation of probabilities based on prior information. For more on Bayesian networks, see the example papers of Anderson 1998 and Lehmkuhl et al. 2001.

How can I learn more about conducting a Bayesian analysis?

Software programs are available to conduct Bayesian analysis, including packages in R and WinBUGS (Windows Bayesian inference using Gibbs sampling). These and other software programs allow for Bayesian analysis of simple and complex statistical models using Markov chain Monte Carlo (MCMC) methods. In WinBUGS, you begin by specifying the prior and build up the likelihood probabilities with data, then the program computes the posterior by running a Gibbs sampling algorithm. For access to the software and more information on WinBUGS, see this web site: <http://www.mrc-bsu.cam.ac.uk/bugs/>. The computational requirements for running a Bayesian analysis can be substantial, so computers with advanced computing power are recommended.

Literature Cited

Anderson, J.L. 1998. Embracing uncertainty: the interface of Bayesian statistics and cognitive psychology. *Conservation Ecology* [online] 2(1): 2. Available from the Internet. URL: <http://www.consecol.org/vol2/iss1/art2/>

Bayes, T. 1763. An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions, Royal Society of London*. 53: 370-418.

Dennis, B. 1996. Discussion: should ecologists become Bayesians? *Ecological Applications* 6: 1095-1103.

Dorazio, R.M. and F.A. Johnson. 2003. Bayesian inference and decision theory – a framework for decision making in natural resource management. *Ecological Applications* 13: 556-563.

Ellison, A.M. 1996. An introduction to Bayesian inference for ecological research and environmental decision-making. *Ecological Applications* 6: 1036-1046.

Ellison, A.M. 2004. Bayesian inference in ecology. *Ecology Letters* 7: 509-520.

Hilborn, R. and M. Mangel. 1997. *The ecological detective: confronting models with data*. Princeton University Press. Princeton, New Jersey.

Kiiveri, H.T., P. Caccetta, and F. Evans. 2001. Use of conditional probability networks for environmental monitoring. *International Journal of Remote Sensing* 22: 1173-1190.

Lehmkuhl, J.F., J.G. Kie, L.C. Bender, G. Servheen, and H. Nyberg. 2001. Evaluating the effects of ecosystem management alternatives on elk, mule deer, and white-tailed deer in the interior Columbia River basin, USA. *Forest Ecology and Management* 153: 89-104.

Marcot, B.G., R.S. Holthausen, M.G. Raphael, M.M. Rowland, and M.J. Wisdom. 2001. Using Bayesian belief networks to evaluate fish and wildlife population viability under land management alternatives from an environmental impact statement. *Forest Ecology and Management* 153: 29-42.

Prato, T. 2005. Bayesian adaptive management of ecosystems. *Ecological Modelling* 183: 147-156.

Quinn, G.P. and M.J. Keough. 2002. *Experimental design and data analysis for biologists*. Cambridge University Press. Cambridge, United Kingdom.

Wade, P.R. 2000. Bayesian methods in conservation biology. *Conservation Biology* 14: 1308-1316.