

Stat 136 (Bayesian Statistics)

Lesson 1.4: Elements of Bayesian Models

Elements of a Bayesian model

Bayes' formula extends naturally to statistical models. A Bayesian model is a parametric model in the classical (or frequentist) sense, but with the addition of a prior probability distribution for the model parameter, which is treated as a random variable rather than an unknown constant. The basic elements of a Bayesian model may be listed as:

1. The **parameter** of interest, say θ . Note that this is completely general, since θ may be vector valued. So θ might be a binomial parameter, or the mean and variance of a Normal distribution, or an odds ratio, or a set of regression coefficients, etc. The parameter of interest is sometimes usefully thought of as the “true state of nature”.
2. The **prior distribution** of θ , $f(\theta)$. This prior distribution summarizes what is known about θ before the experiment is carried out. It is “subjective”, so may vary from investigator to investigator.
3. The **likelihood function**, $f(y|\theta)$. The likelihood function provides the distribution of the data, y , given the parameter value θ . So it may be the binomial likelihood, a normal likelihood, a likelihood from a regression equation with associated normal residual variance, logistic regression model, etc.
4. The **posterior distribution**, $f(\theta|y)$. The posterior distribution summarizes the information in the data, y , together with the information in the prior distribution, $f(\theta)$. Thus, it summarizes what is known about the parameter of interest θ after the data are collected.
5. **Bayes Theorem**. This theorem relates the above quantities:

$$\text{posterior distribution} = \frac{\text{likelihood of the data} \times \text{prior distribution}}{\text{normalizing constant}}$$

or

$$f(\theta|y) = \frac{f(y|\theta) \times f(\theta)}{f(y)}$$

where:

$$f(y) = \begin{cases} \sum_{\theta} f(\theta) \times f(y|\theta), & \text{if } \theta \text{ is discrete} \\ \int f(\theta) \times f(y|\theta) d\theta, & \text{if } \theta \text{ is continuous} \end{cases}$$

Ignoring the normalizing constant, we get

$$f(\theta|y) \propto f(y|\theta) \times f(\theta)$$

Thus we “update” the prior distribution to a posterior distribution after seeing the data via Bayes Theorem.

6. The **action**, a . The action is the decision or action that is taken after the analysis is completed. For example, one may decide to treat a patient with *Drug 1* or *Drug 2*, depending on the data collected in a clinical trial. Thus our action will either be to use *Drug 1* (so that $a = 1$) or *Drug 2* (so that $a = 2$).
7. The **loss function**, $L(\theta, a)$. Each time we choose an action, there is some loss we incur, which depends on what the true state of nature is, and what action we decide to take. For example, if the true state of nature is that *Drug 1* is in fact superior to *Drug 2*, then choosing action $a = 1$ will incur a smaller loss than choosing $a = 2$. Now, the usual problem is that we do not know the true state of nature, we only have data that lets us make probabilistic statements about it (ie, we have a posterior distribution for θ , but do not usually know the exact value of θ). Also, we rarely make decisions before seeing the data, so that in general, $a = a(y)$ is a function of the data. Note that while we will refer to these as “losses”, we could equally well use “gains”.
8. **Expected Bayes Loss (Bayes Risk)**: We do not know the true value of θ , but we do have a posterior distribution once the data are known, $f(\theta|y)$. Hence, to make a “coherent” Bayesian decision, we minimize the Expected Bayesian Loss, defined by:

$$EBL = \int L(\theta, a(y)) f(\theta|y) d\theta$$

In other words, we choose the action $a(y)$ such that the EBL is minimized.

The first five elements in the above list comprise a non-decision theoretic Bayesian approach to statistical inference. This type of analysis (ie, nondecision theoretic) is what most of us are used to seeing in the medical literature. However, many Bayesians argue that the main reason we carry out any statistical analyses is to help in making decisions, so that elements 6, 7, and 8 are crucial. There is little doubt that we will see more such analyses in the near

future, but it remains to be seen how popular the decision theoretic framework will become in medicine. The main problem is to specify the loss functions, since there are so many possible consequences (main outcomes, side effects, costs, etc.) to medical decisions, and it is difficult to combine these into a single loss function. My guess is that much work will have to be done on developing loss functions before the decision theoretic approach becomes mainstream. This course, therefore, will focus on elements 1 through 5.

Steps of Bayesian Data Analysis

1. **Identify/Collect the data** required to answer the research questions. As a general recommendation, it is helpful to visualize the data to get a sense of how the data look, as well as to inspect for any potential anomalies in the data collection.
2. **Choose a statistical model** for the data in relation to the research questions. The model should have good theoretical justification and have parameters that are meaningful for the research questions.
3. **Specify prior distributions** for the model parameters. Although this is a subjective endeavor, the priors chosen should be at least sensible to audience who are skeptical.
4. **Obtain the posterior distributions** for the model parameters. As described below, this can be obtained by analytical or various mathematical approximations. For mathematical approximations, one should check the algorithms for convergence to make sure the results closely mimic the target posterior distributions
5. **Conduct a posterior predictive check** to examine the fit between the model and the data, i.e., whether the chosen model with the estimated parameters generate predictions that deviate from the data being analyzed on important features. If the model does not fit the data, one should go back to step 2 to specify a different model
6. If the fit between the model and the data is deemed satisfactory, one can proceed to **interpret the results** in the context of the research questions. It is also important to **visualize the results** in ways that are meaningful for the analysis.