Every Bayesian computation is an integral (or a sum)

Matthew Stephens *University of Chicago*

See here for a PDF version of this vignette.

Prerequisites

Be familiar with basic probability and Bayesian calculations.

Overview

The goal here is simply to point out that everything you want to compute in Bayesian calculations is an integral.

Examples

Consider inference for a parameter θ from data D.

The posterior distribution of θ is given by Bayes Theorem

$$p(\theta|D) = p(\theta)p(D|\theta)/p(D)$$

First note that the denominator $p(D)$ is an integral:

$$p(D) = \int p(D|\theta)p(\theta)d\theta$$

.

Now suppose we want to estimate θ by its posterior mean. This is

$$E(\theta|D) = \int \theta p(\theta|D)d\theta.$$

And if we want to find a 90% posterior credible interval for θ then we want to find A and B such that $\Pr(\theta \in [A, B]|D) = 0.9$. Note that the LHS of this is

$$\Pr(\theta \in [A, B]|D) = \int I(\theta \in [A, B])p(\theta|D)d\theta,$$

where $I(E)$ denotes the indicator function for the event E, which takes the value 1 if E is true and 0 otherwise.

Examples: discrete

Of course, if θ is discrete then the integrals above all become sums.

For example

$$E(\theta|D) = \sum_n \theta_n \Pr(\theta = \theta_n|D)$$

where $\theta_1, \theta_2, \ldots$ are the possible values for θ.

Summary

Pretty much all the things you want to compute when doing Bayesian inference are integrals (or sums) of one kind or another…

If you are computing 1-dimensional integrals then numerical methods are often useful. For example, Simpsons Rule, Gaussian Quadrature. These can also work in 2-dimensions, and maybe even 3 or 4.

Other simple methods that can work for low dimensions: naive Monte Carlo, and Importance Sampling. Also Laplace approximation.

For higher dimensions we usually resort to Markov Chain Monte Carlo.