Bayes Workshop 1

Bayesian Probability and the Likelihood Function

Corson N. Areshenkoff

University of Victoria

March 15, 2016

The need for Bayesian models

Setup:

- Police force in a city of 1,000,000 people with 100 criminals
- Lie detector with 99% accuracy
- H_0: Person is innocent

Experiment:

- Test one person and find guilty.
- p-value?
  - 1% false positive rate, so $p = 0.01$.
- Probability that person is actually guilty?

$$
\begin{aligned}
P(H_1|D) &= \frac{P(D|H_1)P(H_1)}{P(D)} \\
&= \frac{0.99 \cdot 0.0001}{0.99 \cdot 0.0001 + 0.01 \cdot 0.9999} \\
&\approx 0.001
\end{aligned}
$$

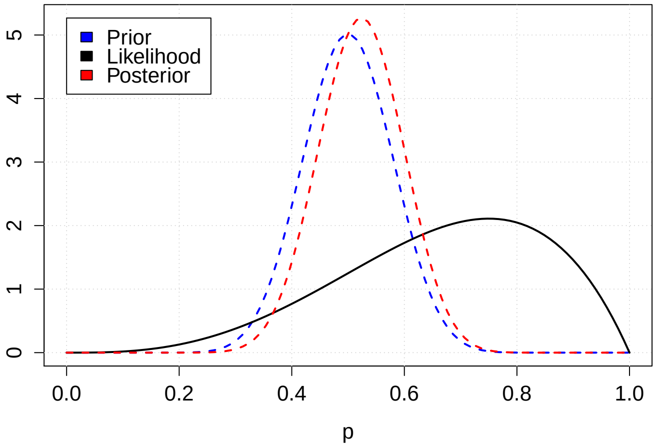

Figure: Binomial posterior for $N = 4$ and $k = 3$.

The Problem

- Generally, we cannot compute the posterior distribution, so we have to approximate it numerically.
- How?
  - Suppose we want to compute the mean m of a normal distribution.
  - Can compute it analytically:

$$m = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{\frac{(x-\mu)^2}{2\sigma^2}} \mathrm{d}x$$

  - Or, can draw lots of samples from a normal distribution and compute the empirical mean (numerical approach).
  - Or, if we want to see the full distribution, we can plot a histogram of the samples.

The Problem

- Another problem:
  - If we want to draw samples from a normal distribution, we can use e.g. the `rnorm()` function in R.
  - In general, there is no way to draw samples directly from a posterior distribution.
- The solution:
  - Several algorithms (e.g. MCMC) can generate sequences of samples which converge in probability to the posterior distribution.
  - i.e. The algorithm generates samples $(x_1, x_2, \ldots, x_n, \ldots)$, and the distribution of each observation gets closer to the posterior distribution as $n \to \infty$.

Working with MCMC

General MCMC procedure:

1. Specify a model + prior.
2. Use an MCMC "chain" to generate random samples from the posterior.
3. How to tell if MCMC works?

Working with MCMC

Problem Chains take a while to converge to the posterior distribution.

- **Solution** Let the chain "burn-in"; discard the first few samples.

Problem Hard to tell when convergence has taken place for a single chain.

- **Solution** Run multiple chains and verify that they generate similar samples.
- **Solution** Summary statistics (e.g. $\hat{R}$)

Problem Samples are not independent, so 100 samples from MCMC does "contain" 100 samples worth of information about the posterior.

- **Solution** Examine autocorrelation
- **Solution** Effective sample size n_{eff}.

Sample Model

- Consider the problem of estimating a sample mean, where the sample is assumed to be normally distributed with known variance $\sigma^2 = 1$.
- We assume that the mean is close to zero, so we use a standard normal prior.
- Let $X = (x_1, x_2, \ldots, x_N)$ be our sample. Then the model is written as

$$
\begin{aligned}
x &\sim \mathrm{N}(\mu, 1) \\
\mu &\sim \mathrm{N}(0, 1)
\end{aligned}
$$

- In this case, we know that the posterior is

$$\mathrm{N}\left(\frac{1}{N+1}\sum x_i, \frac{1}{N+1}\right)$$

so we can compare it with the results from MCMC.