Bayes’ rule: continuous case

\[ f(\theta | D) = \frac{f(D | \theta) f(\theta)}{\int f(D | \theta) f(\theta) d\theta} \]

If you had to guess...

Not knowing anything about my archery abilities, draw a curve representing your view of the chances of my arrow landing a distance \( d \) from the center of the target (if it helps, I’m standing 50 meters away from the target).

Case 1: assume I have talent

An informative prior (low dispersion = low variance) that says most of my arrows will fall within 10 cm of the center (thanks for your confidence!)

Case 2: assume I have a talent for missing the target!

Also an informative prior, but one that says most of my arrows will fall within a narrow range just outside the entire target!
Case 3: assume I have no talent

This is a vague prior: its high variance reflects nearly total ignorance of my abilities, saying that my arrows could land nearly anywhere!

Integration of densities

A probability density curve is scaled so that the value of this integral (i.e. the total area) equals 1.0

Integration of a probability density yields a probability

Area under the density curve from 0 to 2 is the probability that \( \theta \) is less than 2.
Archery Priors Revisited

These density curves are all variations of a gamma probability distribution. We could have used a gamma distribution to specify each of the prior probability distributions for the archery example. Note that higher variance means less informative.

Archery example

Usually there are many parameters...

\[ f(\theta, \phi | D) = \frac{f(D | \theta, \phi) f(\theta) f(\phi)}{\int_\theta \int_\phi f(D | \theta) f(\theta) f(\phi) d\theta d\phi} \]

Posterior probability density

Marginal probability of data

Markov chain Monte Carlo (MCMC)

For more complex problems, we might settle for a good approximation to the posterior distribution.

MCMC robot’s rules

An analysis of 100 sequences under the simplest model (JC69) requires 197 branch length parameters. The denominator is a 197-fold integral in this case! Now consider summing over all possible tree topologies! It would thus be nice to avoid having to calculate the marginal probability of the data...
(Actual) MCMC robot rules

Uphill steps are always accepted because $R > 1$

Slightly downhill steps are usually accepted because $R$ is near 1

Drastic "off the cliff" downhill steps are almost never accepted because $R$ is near 0

The robot takes a step if it draws a Uniform(0,1) random deviate that is less than or equal to $R$.