Bayesian estimation of model parameter value

In this exercise you will use Bayesian statistics to learn about model parameters. The example is the same that was previously examined using maximum likelihood.

DATA:

Coin flipped N times, heads came up h times (actual values are on next page)

MODEL:

In any flip of the coin, there is probability **p** of getting heads and **(1-p)** of getting tails. The probability of getting **h** heads out of **N** flips is given by:

$$P(h) = \frac{N!}{h!(N-h)!} p^h (1-p)^{(N-h)}$$

The first part of this expression gives the number of ways one can get h heads in N flips (the "binomial coefficient"). The last part gives the probability of getting h heads, each with probability p, and (N-h) tails each with probability (1-p).

Recall that N! (the factorial of N) is defined as follows:

N! = N x (N-1) x (N-2) x ... x 1

For instance 4! = 4 x 3 x 2 x 1 = 24.

GOAL:

We want to estimate p from the observed data (see below) using Bayesian statistics. To simplify matters a bit, we will work in a limited parameter space where p can only have one of the following nine discrete values:

p=0.1, p=0.2, p=0.3, p=0.4, p=0.5, p=0.6, p=0.7, p=0.8, p=0.9

GENERAL INSTRUCTIONS:

On the last page of this hand-out you will find an empty table that should be used for entering the results of your computations. Below the table there is space for making a sketch of the resulting probability distributions (see instructions below).

Exercise

(1) Assess your prior probabilities:

Based on your previous experiences with coin flipping, place your own, personal (subjective) prior probability on each of the possible p-values. Enter the values in column 2 of the table (labeled "Prior probability"). Make sure that the probabilities sum to one.

Presumably you will want to place the highest probabilities on the values of p that are near p=0.5 (although this will of course depend on the sort of company you usually keep). Note that I have already dictated that the prior probabilities of p=0.1 and p=0.9 should be zero. This means that I believe it is absolutely impossible to construct a coin that is this biased.

(2) Observe data:

We will now pretend that you performed an experiment with N=10 flips of the coin, resulting in h=7 heads. You therefore got heads in 70% of the cases.

(3) Compute likelihoods:

Using the equation on the previous page, compute the likelihood of each possible model. Recall that the possible models are p=0.1, p=0.2, ..., p=0.9. The likelihood of one of these models is the probability of getting the observed data (N=10, h=7) given the model. Thus, the likelihood of p=0.3 is P(h=7|p=0.3). Enter the results in column 3 (labeled "Likelihood"). Also compute the sum of the likelihoods and enter the value below the column.

Based on this column, what is the maximum likelihood estimate of p? What is the sum of the likelihoods? Consider whether this sum is meaningful in itself.

(4) Compute the product of priors and likelihoods:

For each row in table 1, compute the product of the prior and the likelihood for that value of p (multiply column 2 by column 3). Enter the value in column 4 (labeled "Prior x Likelihood"). Compute the sum of column 4 and enter the value.

(5) Compute posterior probabilities:

For each row in table 1, compute the posterior probability by dividing the number in column 4 (product of prior and likelihood) by the sum of column 4. Enter the values in column 5 (labeled "Posterior probabilities"). Also compute the sum of this column.

Exercise

(6) Examine prior, likelihood, and posterior distributions

In the space beneath the table, make a rough sketch of these three distributions: (a) prior probability, (b) likelihood, and (c) posterior probability. Remember that the prior distribution shows your beliefs about different parameter values before seeing any data. The likelihood essentially shows what the observed data tells about the probability of different parameter values. The posterior probability combines the information in these two distributions, and shows your updated beliefs about parameter values after having seen the data.

Notice how the prior affects the posterior. What happens when the prior is zero?

(7) Use posterior distribution to learn about parameter value

In frequentist statistics, the goal is to obtain good point estimates of parameter values.
In Bayesian statistics, the goal is instead to obtain a probability distribution over all possible parameter values. This is the posterior probability distribution. It shows how uncertain we are about the parameter value, and can be used as the basis for asking many different questions. Answer the following questions:

(a) What is the "maximum a posteriori" (MAP) estimate of the parameter value?
(The MAP is simply the value that has the highest posterior probability).

(b) What is the probability that $p=0.5$?

(c) What is the probability that $p<0.5$

(d) What is the probability that $p>0.5$

(e) What is the probability $0.3<p<0.7$

(f) Find a 95% credible set: Order the p's according to decreasing posterior probability. (The p with the highest probability is number 1, etc.). Include p's until the cumulated posterior probability is $>=95\%$. The values of p that were included make up a 95% credible set - the Bayesian version of a confidence interval.

Exercise, results

$$\frac{10!}{7! \times 3!} = \_\_\_\_\_\_\_\_\_\_$$

Model (p)	Prior probability	Likelihood	Prior x Likelihood	Posterior probability
0.1	0			
0.2				
0.3				
0.4				
0.5				
0.6				
0.7				
0.8				
0.9	0			
Sum	1			