Bayesian Model Selection-Bayes Factor

Donald E. Brown

School of Data Science
University of Virginia
Charlottesville, VA 22904

Bayesian Model Evaluation

D.E. Brown

Given an indexed set of models $M_1, \ldots, M_m$, and priors for each model, $p(M_i)$, the model posterior probability

$$p(M_i|\mathcal{D}) = \frac{p(\mathcal{D}|M_i)p(M_i)}{p(\mathcal{D})}$$

where the unconditional likelihood of the data $\mathcal{D}$ is

$$p(\mathcal{D}) = \sum_{i=1}^{m} p(\mathcal{D}|M_i)p(M_i)$$

Model M_i is parameterised by θ_i, and the model likelihood is

$$p(\mathcal{D}|M_i) = \int p(\mathcal{D}|\theta_i, M_i)p(\theta_i|M_i)d\theta_i$$

In discrete parameter spaces, the integral is replaced with summation. Note that the number of parameters $\dim(\theta_i)$ need not be the same for each model.

Bayes Factor

Comparing two competing model hypotheses M_i and M_j only requires the Bayes Factor:

$$\frac{p(M_i|\mathcal{D})}{p(M_j|\mathcal{D})} = \underbrace{\frac{p(\mathcal{D}|M_i)}{p(\mathcal{D}|M_j)}}_{\text{Bayes' Factor}} \frac{p(M_i)}{p(M_j)}$$

which does not require integration/summation over all possible models.

Caveat

$p(M_i|\mathcal{D})$ only refers to the probability relative to the set of models specified $M_1, \ldots, M_m$. This is not the *absolute* probability that model M fits 'well'.

Suggestions Regarding Bayes Factor Values

Bayes Factor	Strength of Evidence
1-3	Anecdotal
3-10	Moderate
10-30	Strong
30-100	Very Strong
> 100	Extreme

Example: Fair or Biased coin?

Two models:

M_{fair} : The coin is fair, M_{biased} : The coin is biased

For simplicity we assume $\mathrm{dom}(\theta) = \{0.1, 0.2, \ldots, 0.9\}$.

$p(\theta|M)$

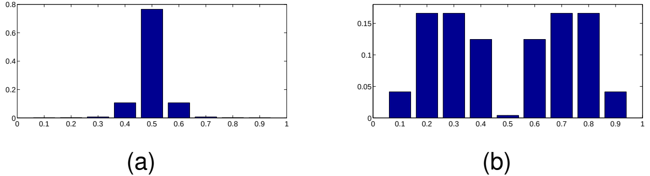

Figure: **(a)**: Discrete prior model of a 'fair' coin $p(\theta|M_{fair})$. **(b)**: Prior for a biased 'unfair' coin $p(\theta|M_{biased})$. In both cases, we are making explicit choices here about what we consider to be a "fair" and "unfair."

Example: Fair or Biased coin?

The model likelihood

For each model M, the likelihood is

$$p(\mathcal{D}|M) = \sum_{\theta} p(\mathcal{D}|\theta, M)p(\theta|M) = \sum_{\theta} \theta^{N_H} (1-\theta)^{N_T} p(\theta|M)$$

This gives

$$0.1^{N_H} (1-0.1)^{N_T} p(\theta = 0.1|M) + \ldots \\ + 0.9^{N_H} (1-0.9)^{N_T} p(\theta = 0.9|M)$$

Bayes factor

Assuming that $p(M_{fair}) = p(M_{biased})$ the Bayes' factor is

$$\frac{p(M_{fair}|\mathcal{D})}{p(M_{biased}|\mathcal{D})} = \frac{p(\mathcal{D}|M_{fair})}{p(\mathcal{D}|M_{biased})}$$

Example: Fair or Biased coin?

5 Heads and 2 Tails

Here $p(\mathcal{D}|M_{fair}) = 0.00786$ and $p(\mathcal{D}|M_{biased}) = 0.0072$. The Bayes factor is

$$\frac{p(M_{fair}|\mathcal{D})}{p(M_{biased}|\mathcal{D})} = 1.09$$

indicating that there is little to choose between the two models.

50 Heads and 20 Tails

Here $p(\mathcal{D}|M_{fair}) = 1.5 \times 10^{-20}$ and $p(\mathcal{D}|M_{biased}) = 1.4 \times 10^{-19}$. The Bayes factor is

$$\frac{p(M_{fair}|\mathcal{D})}{p(M_{biased}|\mathcal{D})} = 0.109$$

indicating that have around 10 times the belief in the biased model as opposed to the fair model.