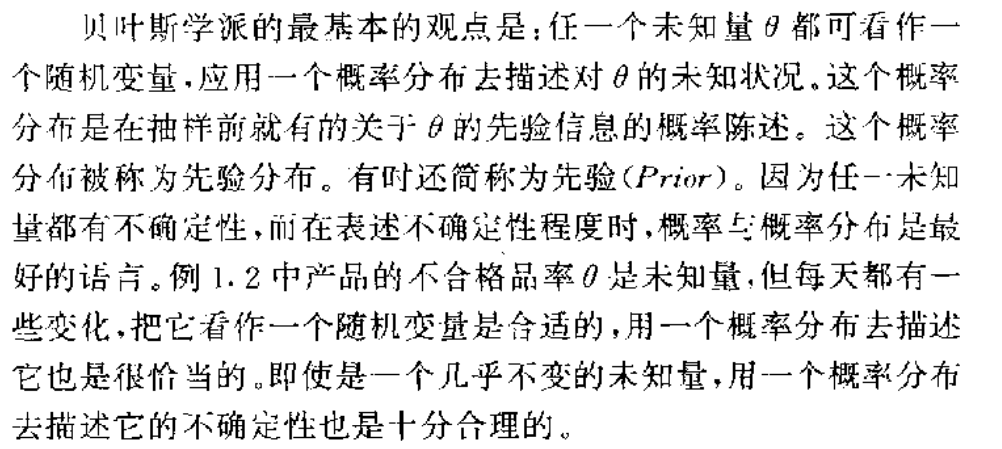
**Based on overall information, sample information and prior information.**

****

**Prior probability**: the prior probability of a random event or an uncertain proposition is the unconditional probability that is assigned before any relevant evidence is taken into account.

**Posterior probability**: the posterior probability of a random event or an uncertain proposition is the conditional probability that is assigned after the relevant evidence or background is taken into account.

**Bayes’ rule:**

P(A|B) = P(A and B) / P(B)

**Credible interval**: an interval in the domain of a posterior probability distribution or predictive distribution used for interval estimation.

**P-value**: probability of the observed or more extreme (to the direction of the alternative hypothesis) outcome given that the null hypothesis is true.

**Inference for a proportion: Bayesian approach**

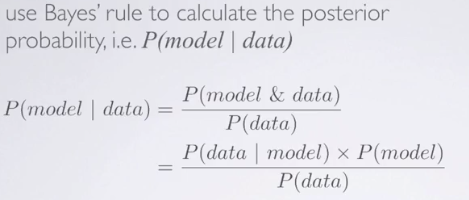
Model (p)

Prior probability (model)

Likelihood probability (data | model)

P(data | model) × P(model)

Posterior probability (model | data)



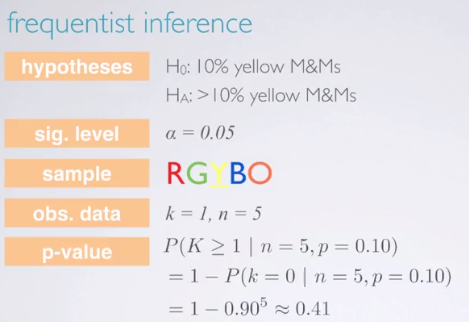
Find out which model has the highest posterior probability.

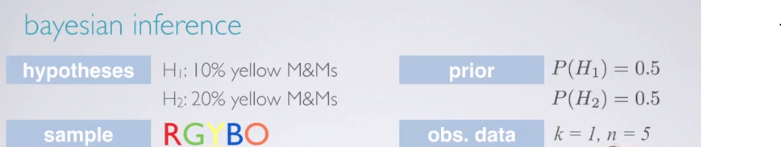
Data ‘at least as extreme as observed’ plays no part in the Bayesian paradigm.

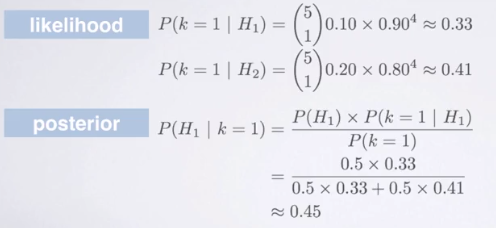
Bayesian paradigm allows us to make direct probability statements about our models.

When sample size increases, the likelihood dominates the prior. (when we don't put a zero probability mass on any of the model in the prior.)

Frequentist vs. Bayesian inference





With equal priors on the two models, and a low sample size, it is difficult to tell with strong confidence which model is more likely given the observed data. If we have to make a decision, we would pick hypothesis 2 because it has higher posterior probability.

The frequentist method is highly sensitive to the null hypothesis, while for the Bayesian method, our results would be the same regardless of which order we evaluate our models.



Summary:

1. continuous random variables can take any value in a range

2. the probability that a continuous random variable takes a specific value is zero.

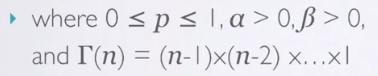
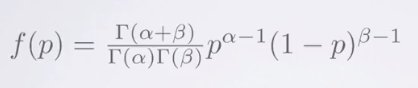
3. its probabilities are determined by a pdf, which is non-negative and the area under the curve is equal to one.

4. the probability that it lies between c and d is the area under the pdf between c and d.

Elicitation

Beta distribution

The pdf of a beta distribution is specified by two parameters, α and β.

****

**Summary**

Bayesians express uncertainty through probability distributions.

One can self-elicit a probability distribution that reflects your personal probability.

Personal probability should change as new data are observed.

The beta family of distribution can flexibly express many possible beliefs about p.

Conjugacy

Conjugacy occurs when your new belief that is your posterior distribution is in the same family as your prior belief but with new parameter values.

# Notes of Bayesian Methods (Thomas Leonard)

## Concepts

1. three main types of probabilities:

**Classical prob**. : this is defined by an ‘m over k’ rule and is appropriate whenever s = {e1, e2, e3, …., ek} possesses k outcomes that are judged to be ‘equally likely’, and when an event A consists of m of these k outcomes.

**Frequency prob**. : this will be defined by the following equation. The frequency prob. of an event is the long-run proportion of times the event occurs in a large number of replications of the experiment.

For any event A contained in S:

m replicates; relative frequency rm(A);

**Subjective prob**. : this measures an individual’s uncertainty in an event and may vary from individual to individual.

## 1.1 Sampling models and likelihoods

Modeling requires substantial inductive thought, while inference requires deduction, that is, the calculation of mathematical conclusions, given that the functional form of the model is assumed true.