What sets Bayes apart?

ESS 575 Models for Ecological Data

N. Thompson Hobbs

January 18, 2017

What sets Bayes apart?

ESS 575 Models for Ecological Data

N. Thompson Hobbs

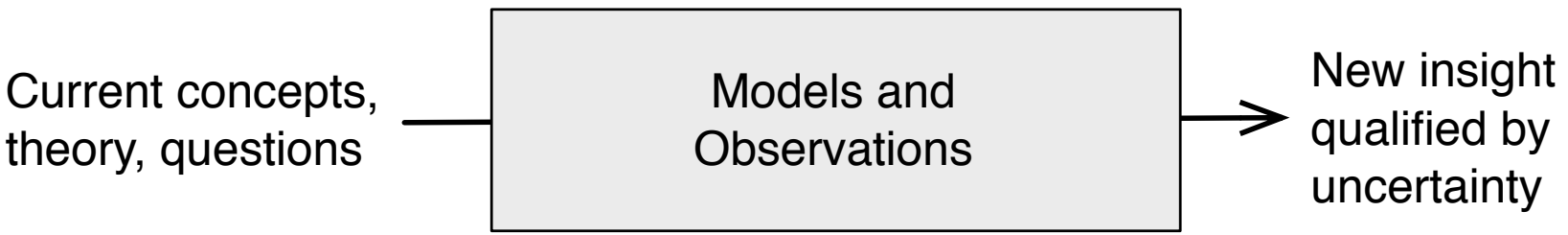
January 18, 2017

Today

- Some motivation for learning
- A high elevation view of Bayesian modeling

What sets statements of scientists apart from statements made by journalists, lawyers, and logicians?

Goals


Current concepts,
theory, questions
Models and
Observations
New insight
qualified by
uncertainty

Exercise

Write out the definition of a frequentist, 95% confidence interval on a parameter of interest, θ.

Frequentist confidence interval

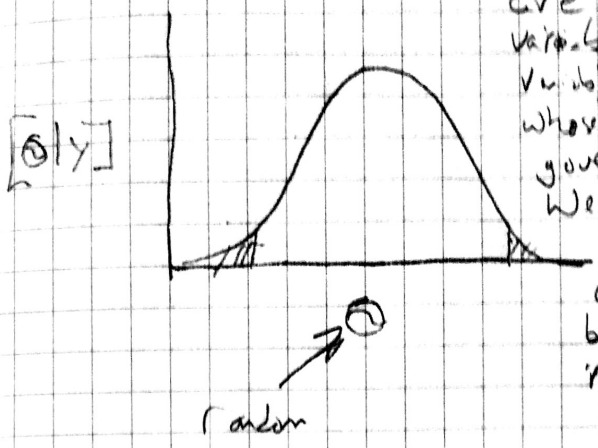
1. In frequentist statistics, a 95% confidence interval represents an interval such that if the experiment were repeated 100 times, 95% of the resulting confidence intervals (e.g.. average + or - 1.96 * standard error) would contain the true, fixed parameter value.
2. In a narrower sense, a CI for a population parameter is an interval with an associated proportion p that is generated from a random sample of an underlying population such that if the sampling was (sic) repeated numerous times and the confidence interval recalculated from each sample according to the same method, a proportion p of the confidence intervals would contain the population parameter in question.

Some notation

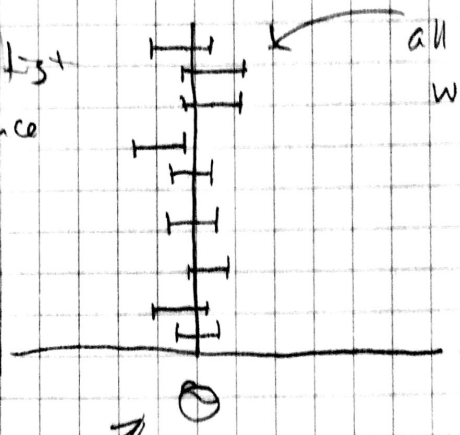
- y data
- θ a parameter or other unknown quantity of interest
- $[y|\theta]$ The probability distribution of y conditional on θ
- $[\theta|y]$ The probability distribution of θ conditional on y
- $P(y|\theta) = p(y|\theta) = [y|\theta] = f(y|\theta) = f(y, \theta)$, different notation that means the same thing.

Bayesian Credible Interval

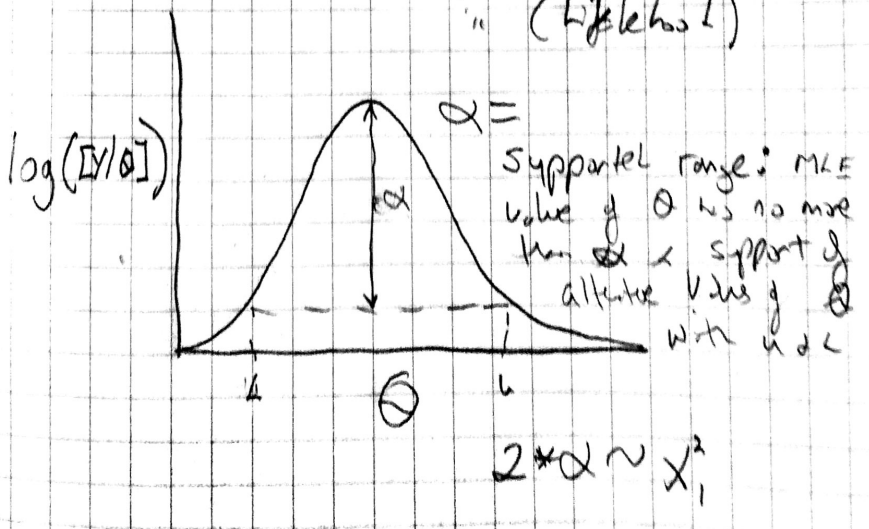
All unobserved quantities are treated as random variables. A random variable is a quantity whose behavior is governed by chance. We seek to understand the probability distribution controlling the behavior of these random variables



Frequentist Confidence Interval



Profile Confidence Interval (Likelihood)



Bayesian models are stochastic.

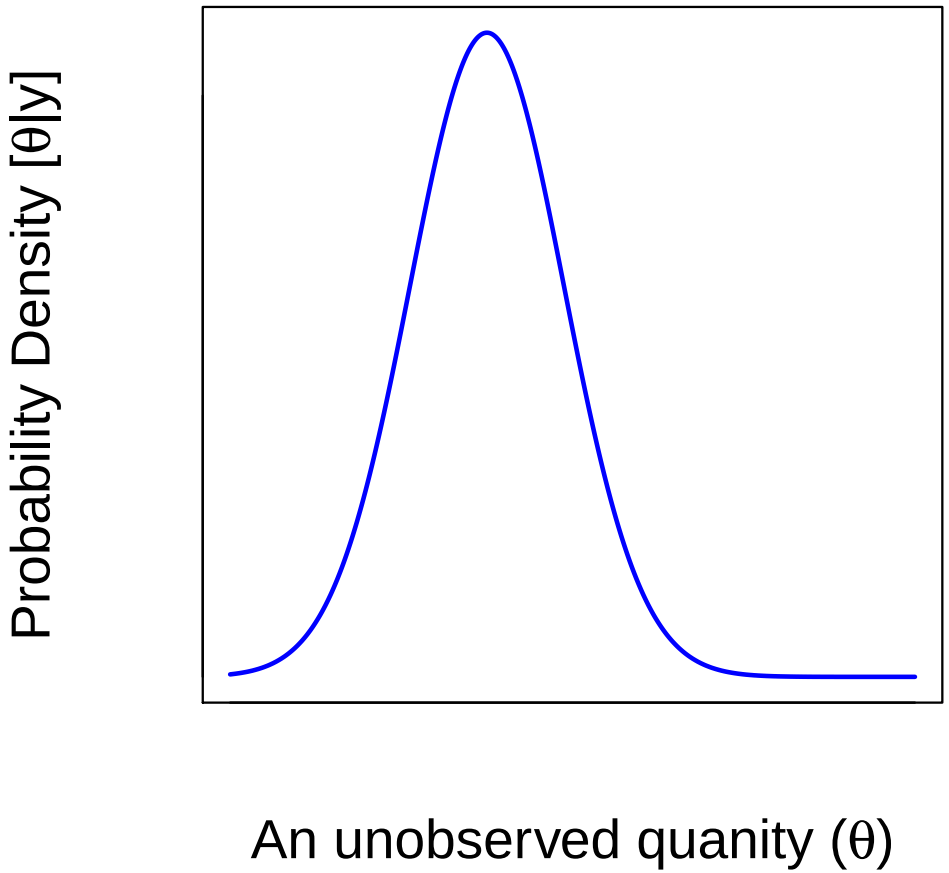
- A model is a mathematical function that returns a quantity (or quantities) given parameters and inputs.
- A deterministic model returns a scalar (or sometimes a vector or matrix) for any given set of parameters and inputs.
- A stochastic model returns a *probability distribution* for any given set of parameters and inputs.
- Probability distributions characterize the behavior of *random variables.*[1].
- In Bayesian analysis, we seek to understand the probability distributions of random variables of interest using data, models, and prior information (including limited prior information).

[1] A random variable is a quantity whose behavior is governed by chance.

What do we do in Bayesian modeling?

- We divide the world into things that are observed (y) and things that unobserved (θ).
- The unobserved quantities (θ) are random variables . The data are random variables before they are observed and fixed after they have been observed.
- We seek to understand the probability distribution of θ using fixed observations, i.e., $[\theta|y]$.
- Those distributions quantify our uncertainty about θ.

Bayesian modeling is a procedure for updating knowledge.

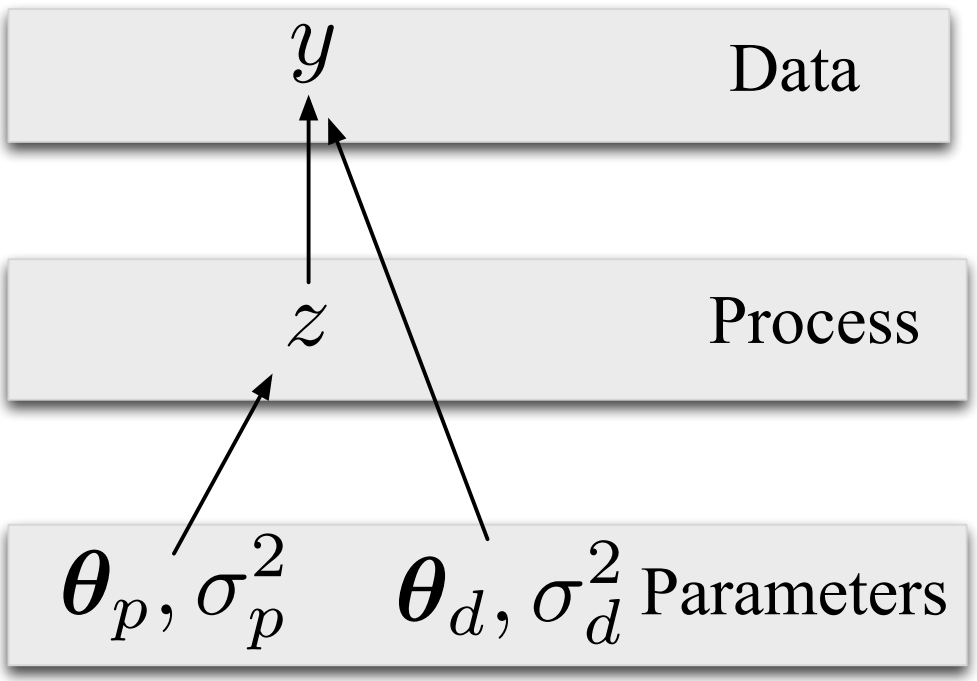


Updating knowledge

Show updating process in R

One approach applies to many problems

- An unobservable state of interest, z
- A deterministic model of a process, $g(\theta, x)$, controlling the state.
- A model of the data
- Models of parameters



You can understand it.

- Rules of probability
  - Conditioning and independence
  - Law of total probability
  - Factoring joint probabilities
- Distribution theory
- Markov chain Monte Carlo