What sets Bayes apart?

Models for Socio-Environmental Data

Chris Che-Castaldo, Mary B. Collins, and N. Thompson Hobbs

May 29, 2018

What sets Bayes apart?

Models for Socio-Environmental Data

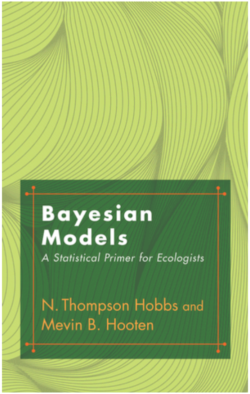
Chris Che-Castaldo, Mary B. Collins, and N. Thompson Hobbs

May 29, 2018

SESYNC

- Introductions
- GitHub for course materials
- Daily schedule
- Lecture style
- Pulling notes just in time
- Exercises

Reading



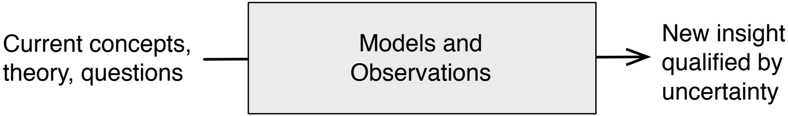
Errata:
http://www.stat.colostate.edu/~hooten/papers/pdf/Hobbs_Hooten_Bayesian_Models_2015_errata.pdf

Today

- A high elevation view of Bayesian modeling
- Goals of course
- Rules of probability
- Probability Distributions
- Moment matching

Exercise

Consider statements made by journalists, lawyers, and scientists. What do they have in common? What sets the statements of scientists apart?


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

Some notation

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

Bayesian models are stochastic.

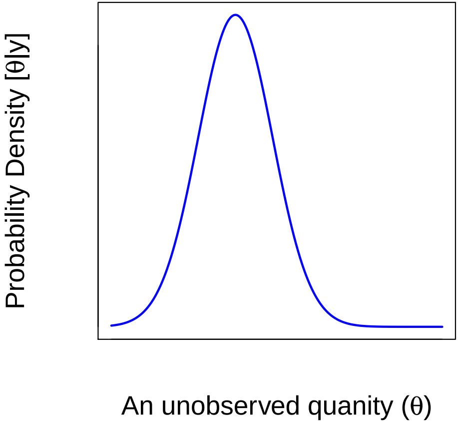
- A model is a mathematical function that returns a quantity (or quantities) given parameters and inputs.
- A deterministic model returns a scalar (or sometimes a vector or matrix) for any given set of parameters and inputs.
- A stochastic model returns a *probability distribution* for any given set 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.

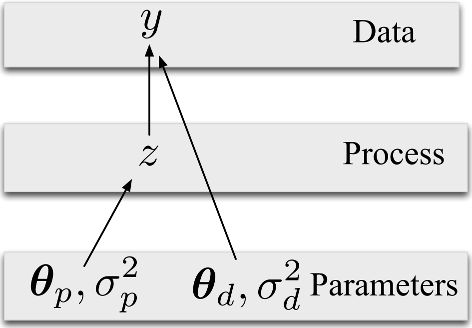
Bayesian modeling is a procedure for updating knowledge.



Current knowledge → New data → Updated knowledge

Why Bayes? One approach applies to many problems

- A deterministic model of a process
- A model of the data
- Models of parameters



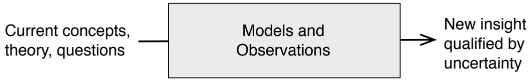
Why Bayes? You can understand it.

KEY TO STATISTICAL METHODS

	Design or Purpose	Measurement Variables	Ranked Variables	Attributes
1 variable 1 sample	Examination of a single sample	Procedure for grouping a frequency distribution, Box 2.1; stem-and-leaf display, Section 2.5; testing for outliers, Section 13.4 Computing median of frequency distribution, Box 4.1 Computing arithmetic mean: unordered sample, Box 4.2; frequency distribution, Box 4.3 Computing standard deviation: unordered sample, Box 4.2; frequency distribution, Box 4.3 Setting confidence limits: mean, Box 7.2; variance, Box 7.3 Computing g_1 and g_2, Box 6.2		Confidence limits for a percentage, Section 17.1 Runs test for randomness in dichotomized data, Box 18.3
	Comparison of a single sample with an expected frequency distribution	Normal expected frequencies, Box 6.1 Goodness of fit tests: parameters from an extrinsic hypothesis, Box 17.1; from an intrinsic hypothesis, Box 17.2 Kolmogorov–Smirnov test of goodness of fit, Box 17.3 Graphic "tests" for normality: large sample sizes, Box 6.3; small sample sizes (rankit test), Box 6.4 Test of sample statistic against expected value, Box 7.4		Binomial expected frequencies, Box 5.1 Poisson expected frequencies, Box 5.2 Goodness of fit tests: parameters from an extrinsic hypothesis, Box 17.1; from an intrinsic hypothesis, Box 17.2
1 variable $\geqslant 2$ samples	Single classification	Single classification anova: unequal sample sizes, Box 9.1; equal sample sizes, Box 9.4 Planned comparison of means in anova, Box 9.8; single degree of freedom comparisons of means, Box 14.10 Unplanned comparison of means: T-method, equal sample sizes, Box 9.9; T′, GT2, and Tukey–Kramer, unequal sample sizes, Box 9.10; Welsch step-up, Box 9.11; STP test, Section 9.7; contrasts using Scheffé, T, and GT2, Box 9.12; multiple confidence limits, Section 14.10 Estimate variance components: unequal sample sizes, Box 9.2; equal sample sizes, Box 9.3 Setting confidence limits to a variance component, Box 9.3 Tests of homogeneity of variances, Box 13.1 Tests of equality of means when variances are heterogeneous, Box 13.2	Kruskal–Wallis test, Box 13.5 Unplanned comparison of means by a nonparametric STP, Box 17.5	*G*-test for homogeneity of percentages, Boxes 17.5 and 17.8 Comparison of several samples with an expected frequency distribution, Box 17.4; unplanned analysis of replicated tests of goodness of fit, Box 17.5
	Nested classification	Two-level nested anova: equal sample sizes, Box 10.1; unequal sample sizes, Box 10.4 Three-level nested anova: equal sample sizes, Box 10.3; unequal sample sizes, Box 10.5		
	Two-way or multi-way classification	Two-way anova: with replication, Box 11.1; without replication, Box 11.2; unequal but proportional subclass sizes, Box 11.4; with a single missing observation, Box 11.5 Three-way anova, Box 12.1 More-than-three-way classification, Section 12.3 and Box 12.2 Test for nonadditivity in a two-way anova, Box 13.4	Friedman's method for randomized blocks, Box 13.9	Three-way log-linear model, Box 17.9 Randomized blocks for frequency data (repeated testing of the same individuals), Box 17.11

Why Bayes? You can understand it.

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



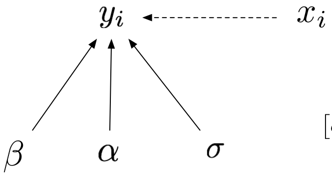
Goals

- Provide principles based understanding
- Foster collaboration
- Build a foundation for self-teaching
- Enhance intellectual satisfaction

Learning objectives

1. Understand basic principles of probability and distribution theory.
2. Explain maximum likelihood.
3. Explain key principles of Bayesian statistics.
4. Be able to diagram, write, and implement hierarchical models.
5. Explain the Markov chain Monte Carlo (MCMC) algorithm.
6. Use software for implementing MCMC methods (i.e., JAGS, R packages).
7. Understand procedures for model checking and model selection in the Bayesian framework
8. Be able to apply Bayesian methods to a broad array of analysis problems in ecology and social science research

Cross cutting theme



$$\mu_i = \frac{m x_i^a}{h^a + x_i^a}$$

$$[a, h, m, \sigma^2 \mid y] \propto \prod_{i=1}^{n} [y_i | \mu_i, \sigma^2][a][h][m][\sigma^2]$$

```
model{

for(i in 1:length(y)){
mu[i] <- (m*x[i]^a)/(h^a+x[i]^a)
y[i] ~ dgamma(mu[i]^2/sigma^2,mu[i]/sigma^2)
}
a ~ dnorm(0,.0001)
m ~ dgamma(.01,.01)
h ~ dgamma(.01,.01)
sigma ~ dunif(0,5)
}
```

Sequence

Day 1 - 3

Principles
- Rules of probability
- Distribution theory
- Moment matching
- Bayes' theorem
- Writing hierarchical models

Day 4 - 5

Implementation
- Conjugate priors
- MCMC
- JAGS

Day 6 - 10

Analysis and inference
- Multi-level regression
- Model checking and selection
- Mixture models
- Dynamic models
- Spatial models
- Ordinal regression