What Sets Bayes Apart?

Models for Socio-Environmental Data

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

June 01, 2019

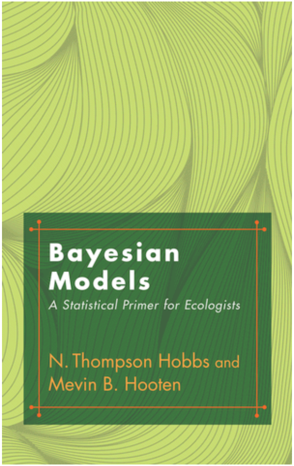
Housekeeping

- Introductions
- GitHub / Website for course materials
- Getting notes just in time
- Daily schedule
- Lecture / exercise mix
- Working in groups
- Individual modeling projects

Pace

- Challenge
- Solutions
- Questions, questions, questions
- Review if needed every day
- Advanced problems

Readings

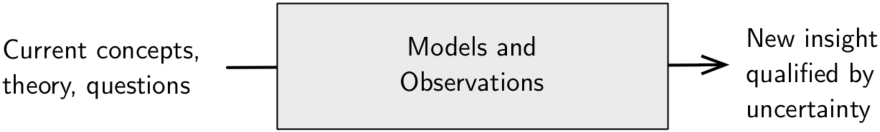


Errata and explanations can be found here

Exercise

What do statements made by journalists, attorneys, and scientists have in common? What sets the statements of scientists apart?

What is this course about?



What is this course about?

Gaining inisght about socio-ecological systems by building models

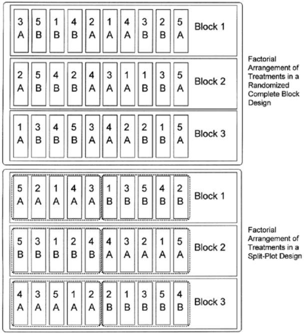
$$[z_i \mid \theta_p]$$

and fitting those models to data

$$[y_i \mid z_i, \theta_d]$$

using Bayesian methods.

Why this course?



Why this course?

Problems poorly suited to traditional approaches:

- Multiple sources of data
- Multiple sources of uncertainty
- Inference across scales
- Unobservable quantities
- Multimodal data
- Derived quantities
- Forecasting
- Synthesis

Why this course?

SESYNC is dedicated to fostering synthetic, actionable science related to the structure, functioning, and sustainability of socio-environmental systems.

Why this course?

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 ⩾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

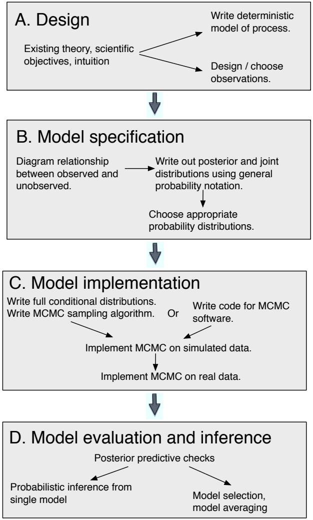
Goals

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

Learning outcomes

- Explain basic principles of Bayesian inference.
- Diagram and write out mathematically correct posterior and joint distributions for Bayesian models.
- Explain basics of the Markov chain Monte Carlo (MCMC) algorithm.
- Use software for implementing MCMC.
- Develop and implement hierarchical models.
- Evaluate model fit.
- Understand papers and proposals using Bayesian methods.

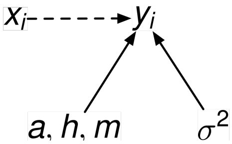
Learning outcomes



Topics

Day 1 - 2	Day 3 - 8	Day 9 - 11
Principles ▪ Rules of probability ▪ Distribution theory ▪ Likelihood ▪ Moment matching ▪ Bayes' theorem ▪ Conjugate priors	Implementation ▪ MCMC ▪ JAGS ▪ Regression ▪ Hierarchical models ▪ Model checking	Advanced topics ▪ Model selection ▪ Designed experiments ▪ Mixture models ▪ Ordinal regression ▪ Dynamic models ▪ Spatial models ▪ Individual problems

Cross cutting theme



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

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

```
model{
  a ~ dnorm(0, .0001)
  m ~ dgamma(.01, .01)
  h ~ dgamma(.01, .01)
  sigma ~ dunif(0, 5)
  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)
  }
}
```

Exercise

Describe how Bayesian analysis differs from other types of statistical analysis.

Some notation

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

Confidence envelopes

I am going to do this on the board. The board notes for this are in the lecture folder.

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., $[\theta \mid y]$.
- Those distributions quantify our uncertainty about θ.

You can understand it

- Rules of probability
  - Conditioning and independence
  - Law of total probability
  - The chain rule of probability
- Distribution theory
- Markov chain Monte Carlo