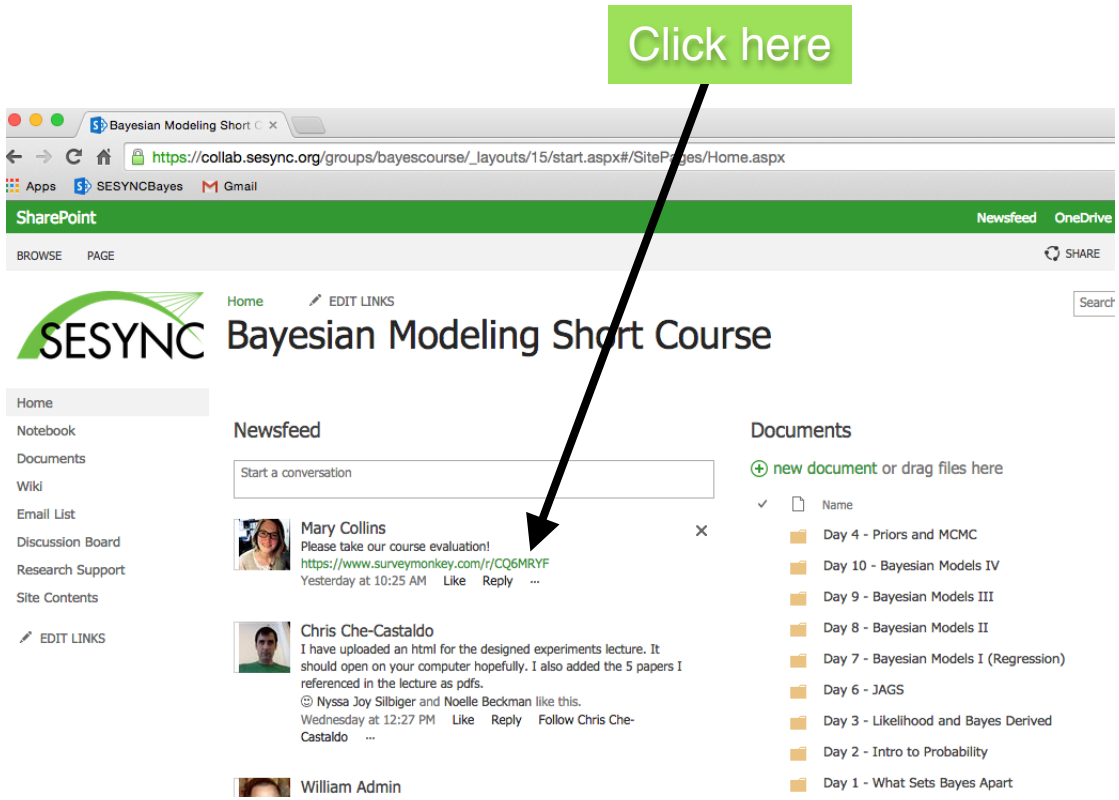
Ordinal categorical outcomes?

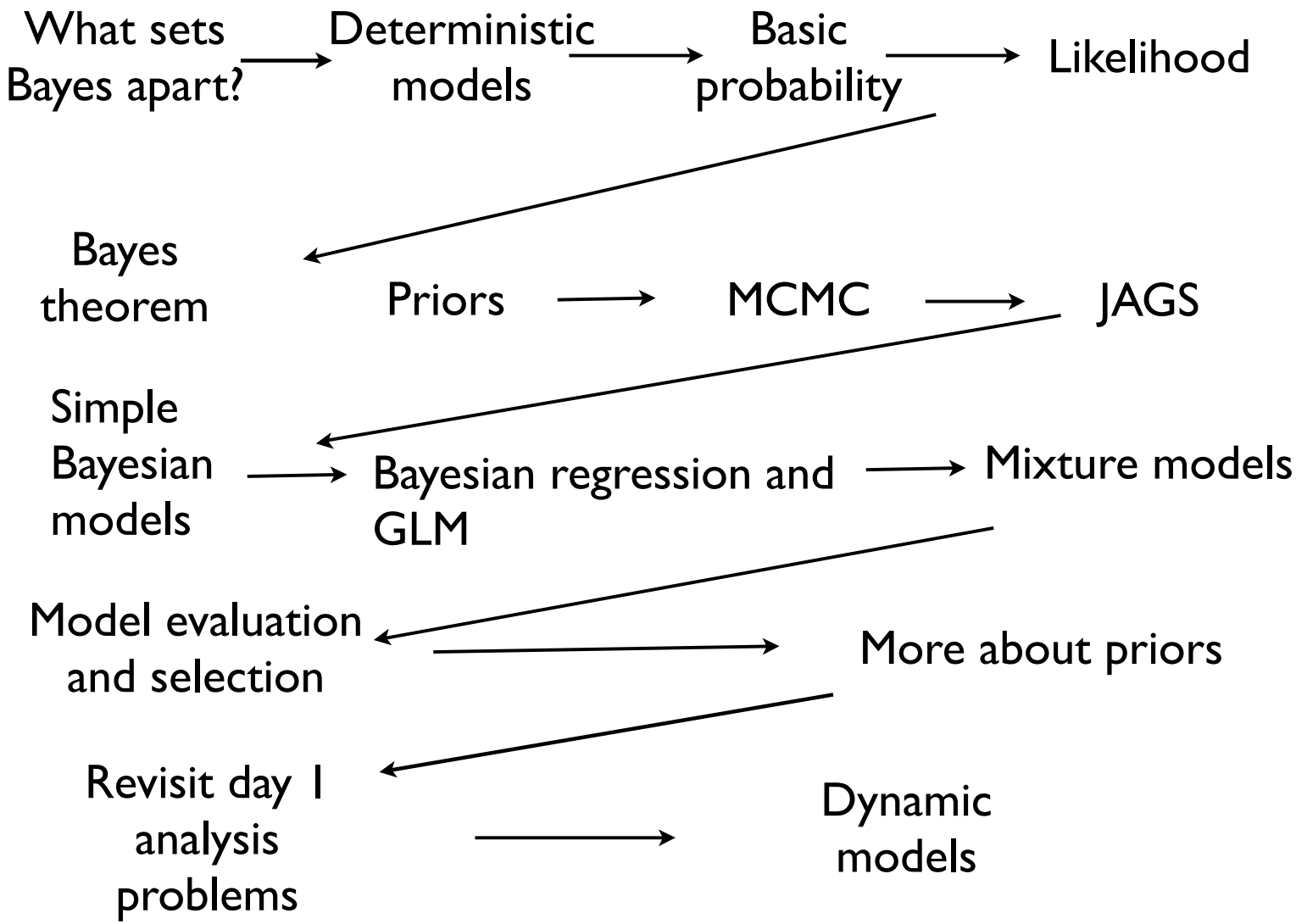
Bayesian Modeling for Socio-Environmental Data

Mary B. Collins

August 28, 2015

Please fill out the course survey




What sets Bayes apart?
Deterministic models
Basic probability
Likelihood
Bayes theorem
Priors
MCMC
JAGS
Simple Bayesian models
Bayesian regression and GLM
Mixture models
Model evaluation and selection
More about priors
Revisit day 1 analysis problems
Dynamic models

Considering dependent variables that are ordinal:

- Likert scales (very dissatisfied to very satisfied)
- Size categories (small, medium, large)
- Acheivement groupings (some high school to graduate degree)
- etc.

Ordinal-level quantities are sometimes called *polychotomous*.

Besides Hobbs and Hooten, 2015...

...I also rely heavily on

1. Gelman, Andrew & Jennifer Hill. 2007. *Data Analyis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.
2. Jackman, Simon. 2009. *Bayesian Analysis for the Social Sciences*. John Wiley & Sons, Ltd.

Ordinary regression analysis is inappropriate.

- Regression is an incoherant model for discrete variables (makes continuous predictions that are impossible for the data).
- We can extend the GLM for binary data to model ordinal data with predictors.
- "Cumulative link GLM" (McKelvey & Zavoina, 1975; Aitchison & Silvey 1957). Logistic regression can be extended to multiple categories (ordered or unordered).
- Caveat for 7+ ordinal categories: treat as continuous.

The ordered multinomial logit model

Consider a categorical outcome y that can take the values $1, 2, ..., K$

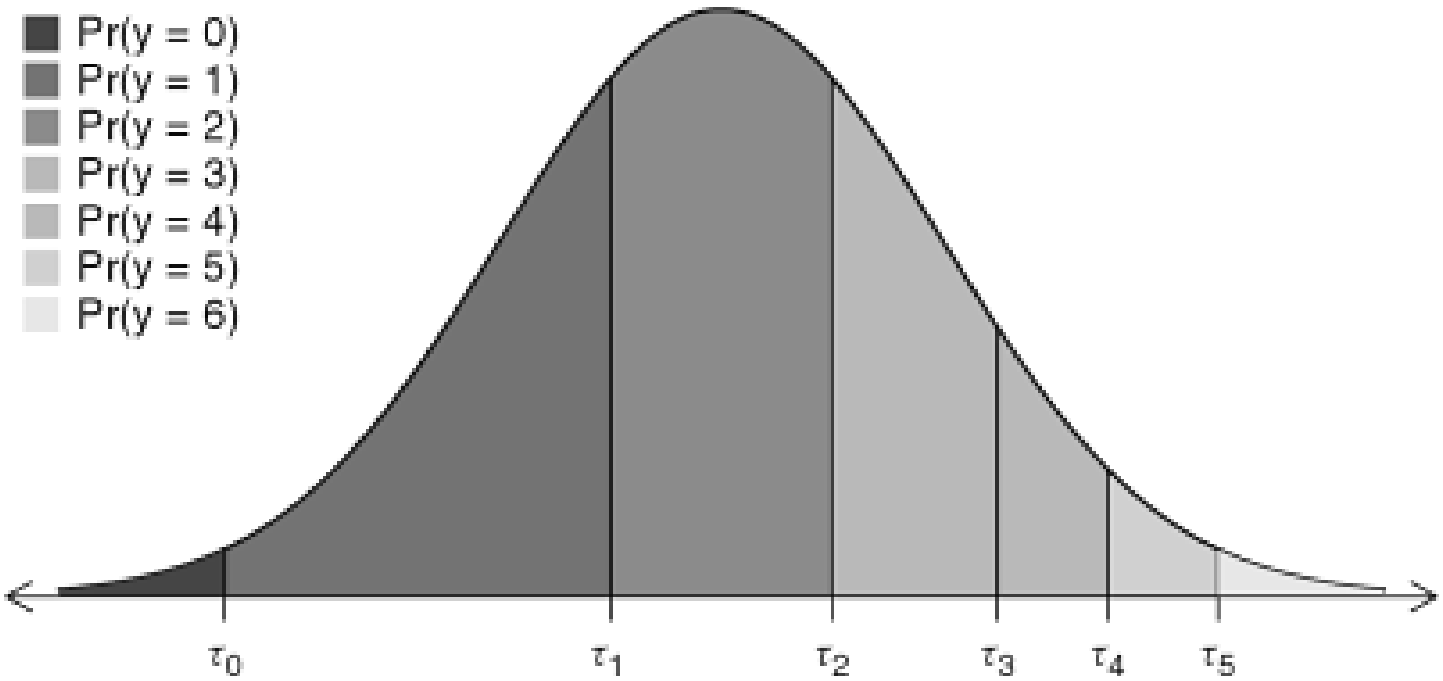
$$\Pr(y > 1) = \text{logit}^{-1}(X\beta) \tag{1}$$
$$\Pr(y > 2) = \text{logit}^{-1}(X\beta - c_2) \tag{2}$$
$$\Pr(y > 3) = \text{logit}^{-1}(X\beta - c_3) \tag{3}$$
$$... \tag{4}$$
$$\Pr(y > K - 1) = \text{logit}^{-1}(X\beta - c_{K-1}) \tag{5}$$

In this model c are thresholds or *cutpoints*

Probabilities of individual outcomes:

$$\Pr(y = k) = \Pr(y > k - 1) - \Pr(y > k) \tag{6}$$

$$= \text{logit}^{-1}(X\beta - c_{k-1}) - \text{logit}^{-1}(X\beta - c_k) \tag{7}$$



Fracking and risk perception example.

Data background: 2500 US respondents took a ~30 web-based environmental risk perception survey. Due to survey length, not all respondents were asked all questions.

Dependent variable: 4-point ordinal acceptability of fracking.

Independent variables: political ideology (liberal – conservative), 3-class political party identification, environmental values index (NEP battery), income, gender, race (white vs. non-white).

DAG