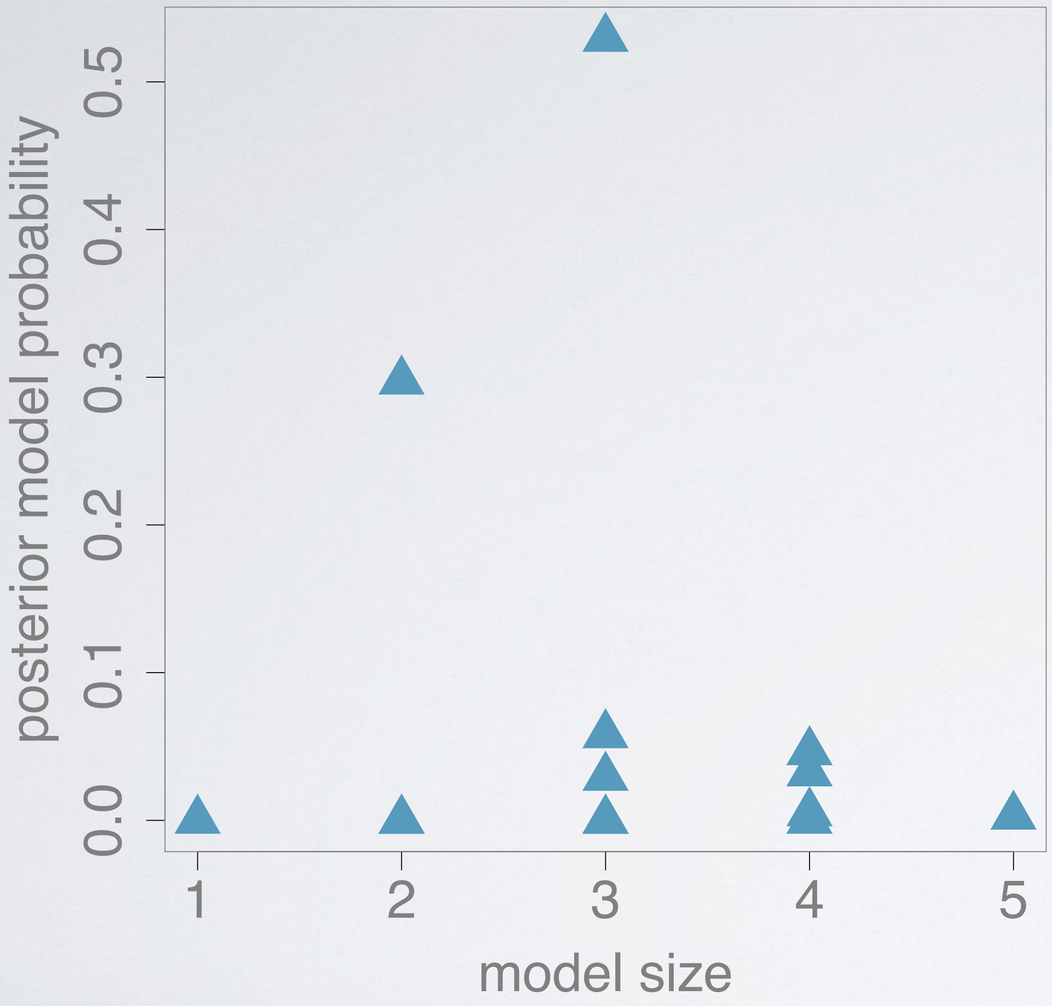
stochastic exploration

Dr. Merlise Clyde

Duke UNIVERSITY

sampling models



- sample models with probability equal to the posterior probability of a model
- estimate quantities by relative frequency

$$P(\mathcal{M}_m \mid \text{data}) \approx \sum_i^I \frac{I(\mathcal{M}_i = \mathcal{M}_m)}{I}$$

- what if there are too many models to enumerate?

Markov Chain Monte Carlo sampling

start at $\mathcal{M}^{(0)}$ for $i = 1, \ldots, I$

1. randomly select $\mathcal{M}^{*(i+1)}$
2. if

$$R \equiv \frac{p(\mathcal{M}^{*(i+1)} \mid \text{data})}{p(\mathcal{M}^{(i)} \mid \text{data})}$$

$$= BF[\mathcal{M}^{*(i+1)} : \mathcal{M}^{(i)}] \times O[\mathcal{M}^{*(i+1)} : \mathcal{M}^{(i)}] > 1$$

set $\mathcal{M}^{(i+1)} \leftarrow \mathcal{M}^{*(i+1)}$

else {

with probability R set $\mathcal{M}^{(i+1)} \leftarrow \mathcal{M}^{*(i+1)}$

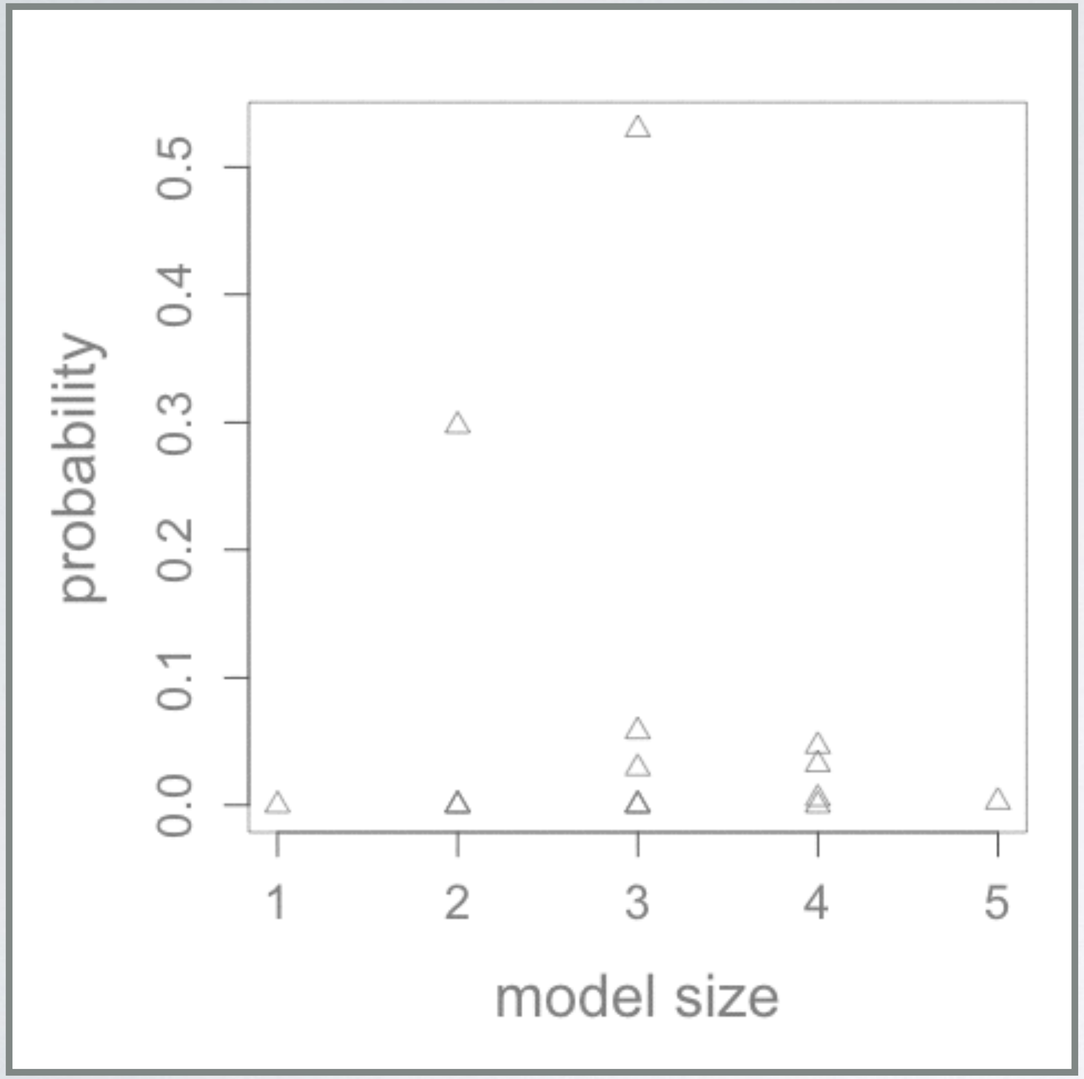
otherwise $\mathcal{M}^{(i+1)} \leftarrow \mathcal{M}^{(i)}$ }

3. $i \leftarrow i + 1$

moving around

- propose to add a variable or delete a variable (symmetric random walk)
- propose to swap a variable
- other moves
- adjust for bias in proposal so that as $I \to \infty$ Monte Carlo frequencies converge to $p(\mathcal{M}_m \mid \mathrm{data})$

moving around



summary

- use MCMC to explore problems that can not be enumerated
- estimates based on the Monte Carlo samples or discover models
- biased stochastic search

next:

- other prior distributions and generalizations