Estimating demographic parameters from samples of unmarked individuals

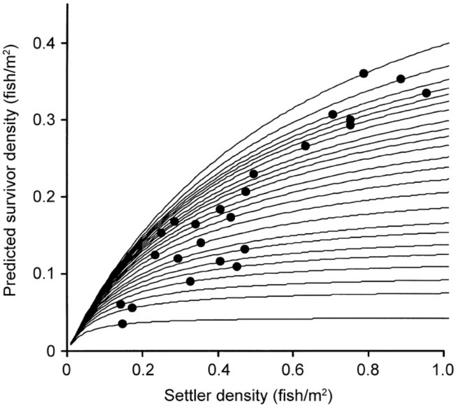
Ben Bolker

Departments of Mathematics & Statistics and Biology, McMaster University

8 March 2011

Outline

Cryptic density-dependence



(Wilson & Osenberg 2002, Shima & Osenberg 2003)

- Environmental variation
- + correlated variation in density ...
- = **cryptic density-dependence**
- best possible estimates of demographic parameters?

Data

- Repeated samples of individual animals
- . . . Multiple times, observers, locations
- Presence and size (imperfect)
- Estimate demographic parameters:
  - birth/immigration
  - death/emigration
  - change (growth)
- . . . as a function of size, (population density)

Data

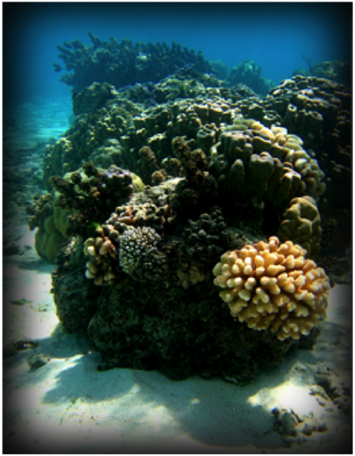
- Repeated samples of individual animals
- . . . Multiple times, observers, locations
- Presence and size (imperfect)
- Estimate demographic parameters:
  - birth/immigration
  - death/emigration
  - change (growth)
- . . . as a function of size, (population density)

Natural history

photo: Leonard Low, Flickr via `species.wikimedia.org`

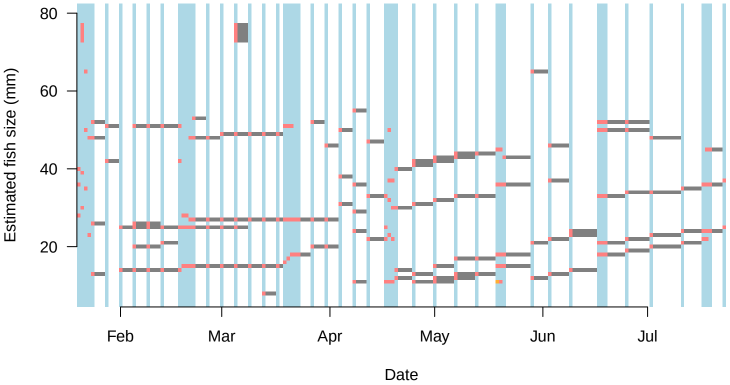
- **Thalassoma hardwicke** (sixbar wrasse)
- **Settlement**: $\approx$ 5–10 mm
- **Growth**: a few mm per month
- **Death**: by predator (competition for refuges)
- **Immigration/emigration**: rare below 40–50 mm

Where they live

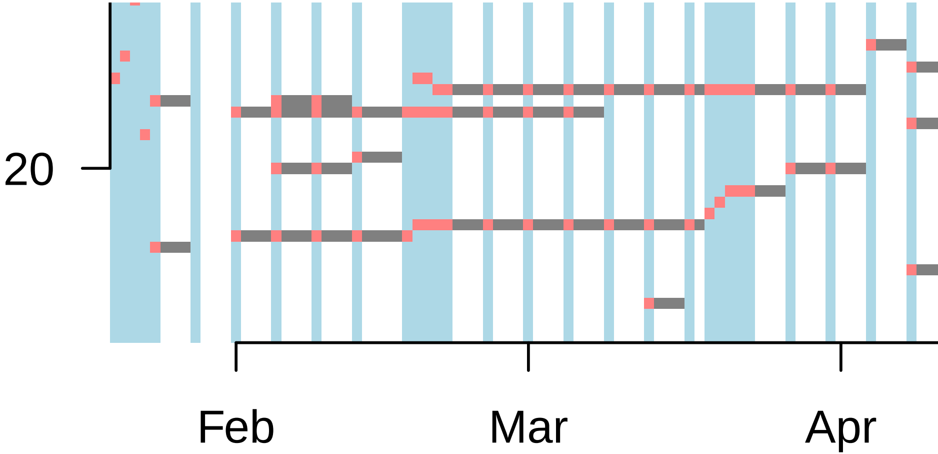


Patch reefs, French Polynesia (and throughout the south Pacific)

What they look like to me



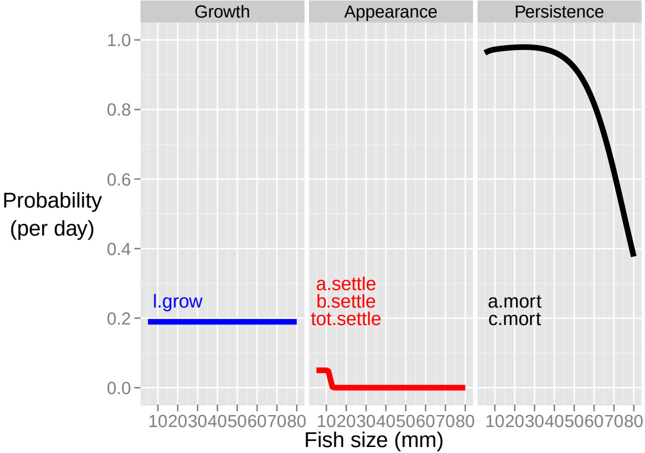
What they look like to me



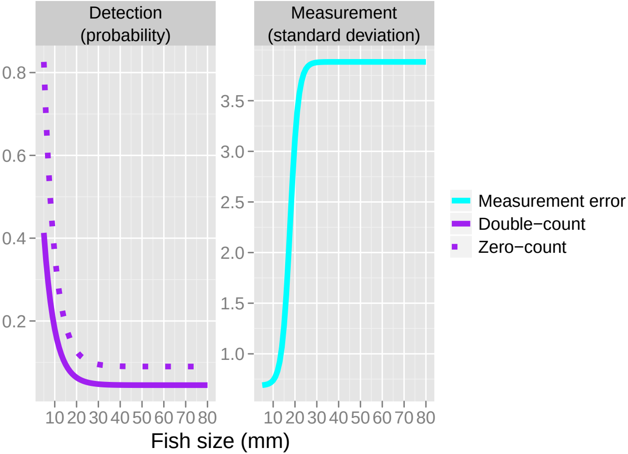
Outline

Outline

Demographic parameters as f(size)



Observation model

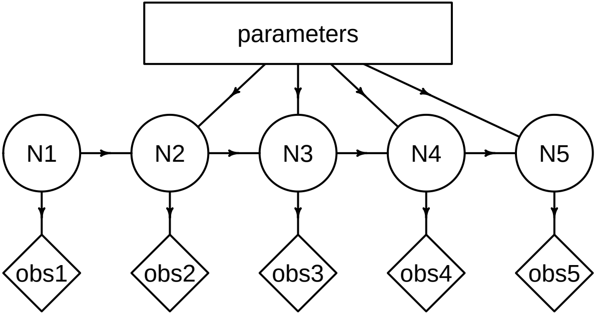


Outline

State-space models

- **Problem**: estimating parameters of systems with un- or imperfectly-observed states?
- Easy if no feedback (measurement error models)
- With feedback (dynamic models), brute force approaches become infeasible . . .

Directed acyclic graph (DAG)



Bayesian statistics, in 30 seconds

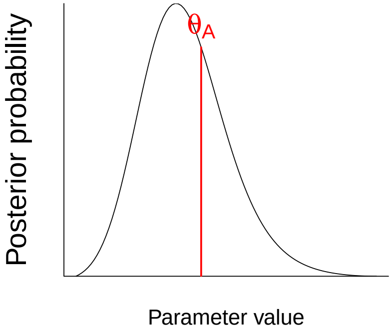
- **computational** (**convenience**) Bayesian statistics

  (forget all that stuff about philosophy of statistics, subjectivity, incorporating prior information ...)
- have a **likelihood**, $L(\boldsymbol{\theta}) = \text{Prob}(\text{data}|\boldsymbol{\theta})$
- want a **posterior probability**, $P_{\text{post}}(\boldsymbol{\theta}) = \text{Prob}(\boldsymbol{\theta}|\text{data})$
- Bayes' rule:

$$P_{\text{post}}(\boldsymbol{\theta}) = \frac{L(\boldsymbol{\theta}) \cdot \text{prior}(\boldsymbol{\theta})}{\iiint L(\boldsymbol{\theta}') \cdot \text{prior}(\boldsymbol{\theta}')\, d\boldsymbol{\theta}'}$$

- may want mean, mode, confidence intervals, ... denominator **very** high-dimensional

Markov chain Monte Carlo

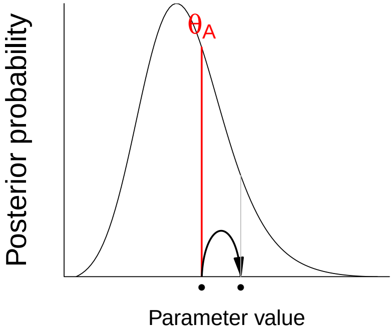


MCMC rule:

$$\text{if } \frac{P(\theta_A)}{P(\theta_B)} = \frac{J(\theta_B \to \theta_A)}{J(\theta_A \to \theta_B)}$$

then stationary distribution = posterior probability (provided chain is irreducible etc.)

Markov chain Monte Carlo

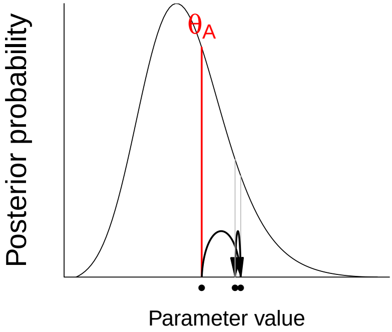


MCMC rule:

$$\text{if } \frac{P(\theta_A)}{P(\theta_B)} = \frac{J(\theta_B \to \theta_A)}{J(\theta_A \to \theta_B)}$$

then stationary distribution = posterior probability (provided chain is irreducible etc.)

Markov chain Monte Carlo

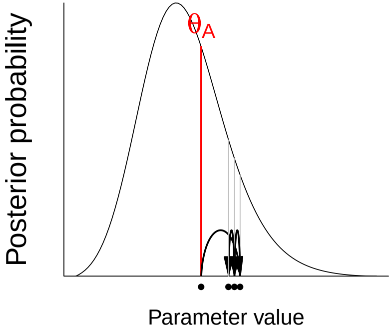


MCMC rule:

$$\text{if } \frac{P(\theta_A)}{P(\theta_B)} = \frac{J(\theta_B \to \theta_A)}{J(\theta_A \to \theta_B)}$$

then stationary distribution = posterior probability (provided chain is irreducible etc.)

Markov chain Monte Carlo

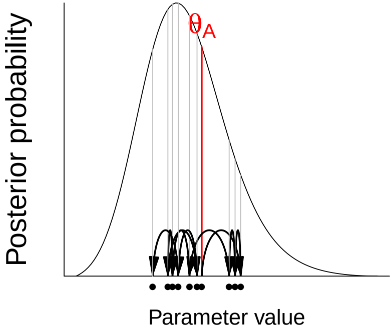


MCMC rule:

$$\text{if } \frac{P(\theta_A)}{P(\theta_B)} = \frac{J(\theta_B \to \theta_A)}{J(\theta_A \to \theta_B)}$$

then stationary distribution = posterior probability
(provided chain is irreducible etc.)

Markov chain Monte Carlo

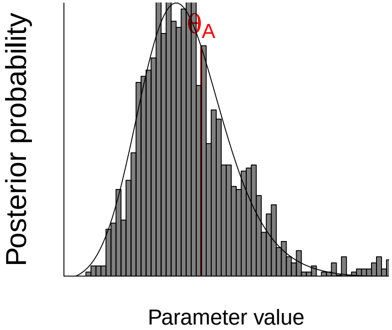


MCMC rule:

$$\text{if } \frac{P(\theta_A)}{P(\theta_B)} = \frac{J(\theta_B \to \theta_A)}{J(\theta_A \to \theta_B)}$$

then stationary distribution = posterior probability (provided chain is irreducible etc.)

Markov chain Monte Carlo

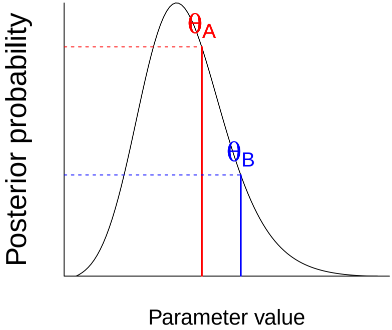


MCMC rule:

$$\text{if } \frac{P(\boldsymbol{\theta}_A)}{P(\boldsymbol{\theta}_B)} = \frac{J(\boldsymbol{\theta}_B \to \boldsymbol{\theta}_A)}{J(\boldsymbol{\theta}_A \to \boldsymbol{\theta}_B)}$$

then stationary distribution = posterior probability
(provided chain is irreducible etc.)

Markov chain Monte Carlo

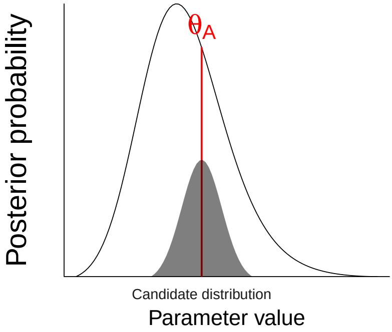


MCMC rule:

$$\text{if } \frac{P(\boldsymbol{\theta}_A)}{P(\boldsymbol{\theta}_B)} = \frac{J(\boldsymbol{\theta}_B \to \boldsymbol{\theta}_A)}{J(\boldsymbol{\theta}_A \to \boldsymbol{\theta}_B)}$$

then stationary distribution = posterior probability (provided chain is irreducible etc.)

Metropolis-(Hastings) updating



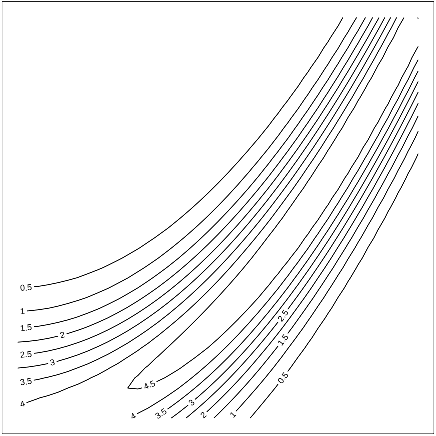
Metropolis rule:

accept B with probability

$$\min\left(1, \frac{P(\boldsymbol{\theta}_B)}{P(\boldsymbol{\theta}_A)}\right)$$

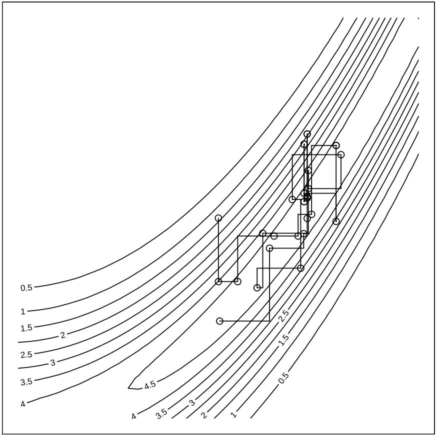
satisfies MCMC rule . . . Don't need to know denominator!

Gibbs sampling



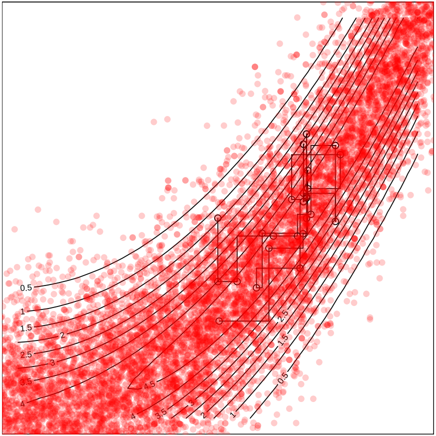
- **Joint** distribution of parameters often difficult
- Condition each element on "known" (i.e. imputed) values of all other elements: $P(a, b, c) = P(a|b, c)P(b, c)$
- **Gibbs sampling**: do this repeatedly for each element, or block of elements

Gibbs sampling



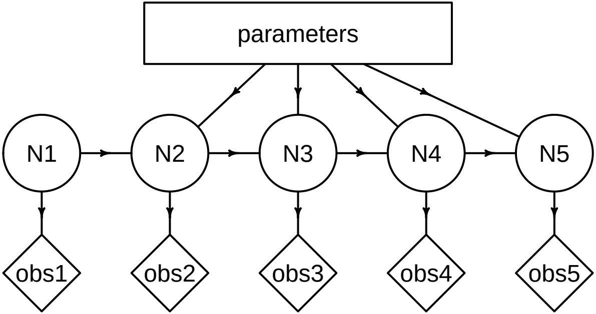
- **Joint** distribution of parameters often difficult
- Condition each element on "known" (i.e. imputed) values of all other elements: $P(a,b,c) = P(a|b,c)P(b,c)$
- **Gibbs sampling**: do this repeatedly for each element, or block of elements

Gibbs sampling



- **Joint** distribution of parameters often difficult
- Condition each element on "known" (i.e. imputed) values of all other elements: $P(a, b, c) = P(a|b, c)P(b, c)$
- **Gibbs sampling**: do this repeatedly for each element, or block of elements

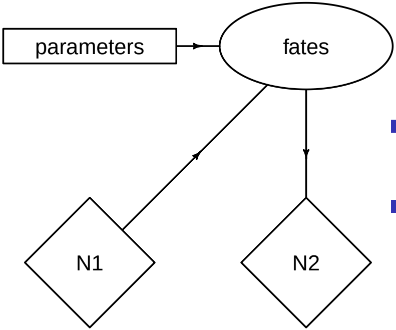
Back to the DAG



Outline

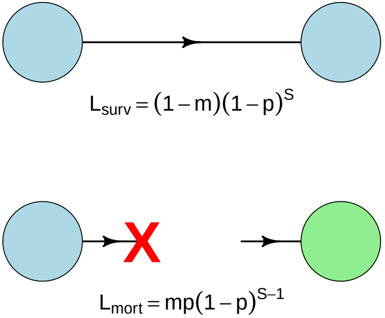
Outline

DAG



- Demographic parameters|fates: standard M-H
- Fates|parameters

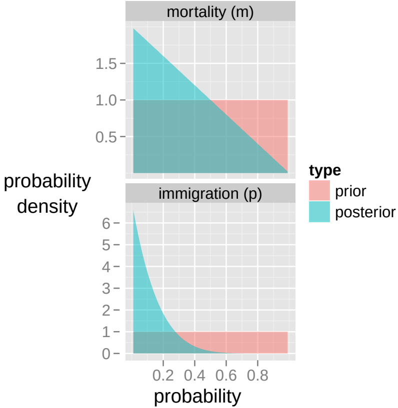
Scenario



One fish of the same size observed on two consecutive days . . .
was it the same individual?

$$P(\text{surv}|m,p) = \frac{(1-m)(1-p)}{1-m-p+2mp}$$

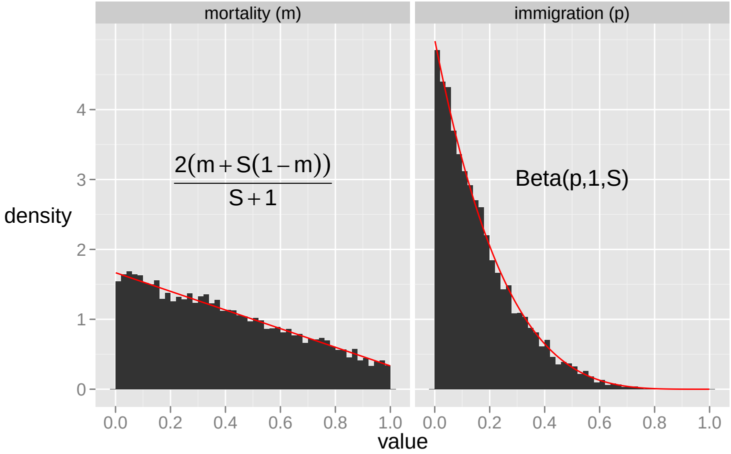
Updating probabilities for parameters



If the fish survived, then our posterior probabilities for the parameters are:

- $m \sim \text{Beta}(1, 2)$ (0 mortalities, 1 survival)
- $p \sim \text{Beta}(1, S + 1)$ (0 immigrations, S non-immigrations)

Trivial example: results



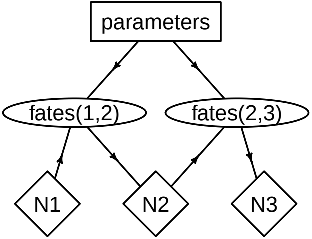
- $P(\text{mortality \& no immigration}) \approx 0.16 = 1/(S+1)$; $P(\text{survival \& immigration}) \approx 0.83 = S/(S+1)$;
- results are simple and make sense; closed form solution is possible but ugly ...
- combination of observations and **non-observation** of other individuals provides information (could fail with non-detection)

Outline

Non-trivial dynamics

- Multiple individuals, measured over multiple days
- Still simple: no density-dependent rates, immigration/emigration of large individuals
- **observations still error-free**

Parameter/fate sampling



- Parameter sampling: M-H updating with simple candidate distributions
- Fate sampling: ???
  - Enumerating possibilities is tedious

    . . .

M-H fate sampling

	6 mm	7 mm	dead
5 mm	$G(5) \cdot (1 - M(5))$	X	$M(5)$
6 mm	$(1 - G(6)) \cdot (1 - M(6))$	$G(6) \cdot (1 - M(6))$	$M(6)$
absent	$I(6)$	$I(7)$	X

M-H fate sampling

	6 mm	7 mm	dead
5 mm	$G(5) \cdot (1 - M(5))$	X	$M(5)$
6 mm	$(1 - G(6)) \cdot (1 - M(6))$	$G(6) \cdot (1 - M(6))$	$M(6)$
absent	$I(6)$	$I(7)$	X

M-H fate sampling

	6 mm	7 mm	dead
5 mm	$G(5) \cdot (1 - M(5))$	X	$M(5)$
6 mm	$(1 - G(6)) \cdot (1 - M(6))$	$G(6) \cdot (1 - M(6))$	$M(6)$
absent	$I(6)$	$I(7)$	X

M-H fate sampling

	6 mm	**7 mm**	dead
5 mm	$G(5) \cdot (1 - M(5))$	X	$M(5)$
6 mm	$(1 - G(6)) \cdot (1 - M(6))$	$G(6) \cdot (1 - M(6))$	$M(6)$
absent	$I(6)$	$I(7)$	X

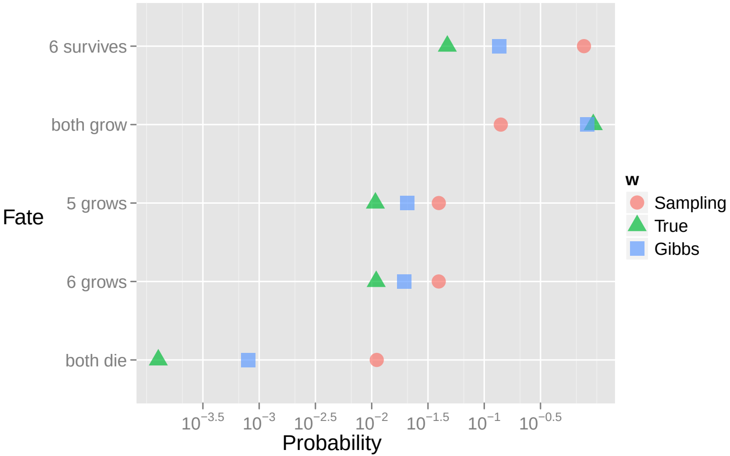
M-H fate sampling

	6 mm	**7 mm**	dead
5 mm	$G(5) \cdot (1 - M(5))$	X	$M(5)$
6 mm	$(1 - G(6)) \cdot (1 - M(6))$	$G(6) \cdot (1 - M(6))$	$M(6)$
absent	$I(6)$	$I(7)$	X

M-H fate sampling

	6 mm	7 mm	**dead**
5 mm	$G(5) \cdot (1 - M(5))$	X	$M(5)$
6 mm	$(1 - G(6)) \cdot (1 - M(6))$	$G(6) \cdot (1 - M(6))$	$M(6)$
absent	$I(6)$	$I(7)$	X

Sampling test



Testbed

- Simulate for 80 days, with a settlement rate (`tot.settle`) of 5% day:
- 29 distinct individuals, size range 5–30, total of 618 fish-days (observations).
- Sample fates only (1000 MCMC steps), fixed demographic parameters

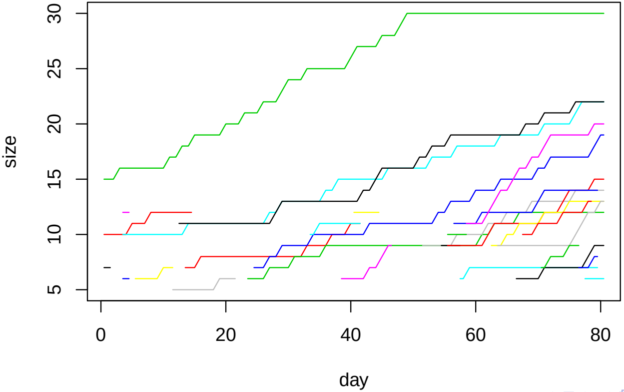
Sampled 1000 unique fates (but only 895 unique likelihoods)

Testbed

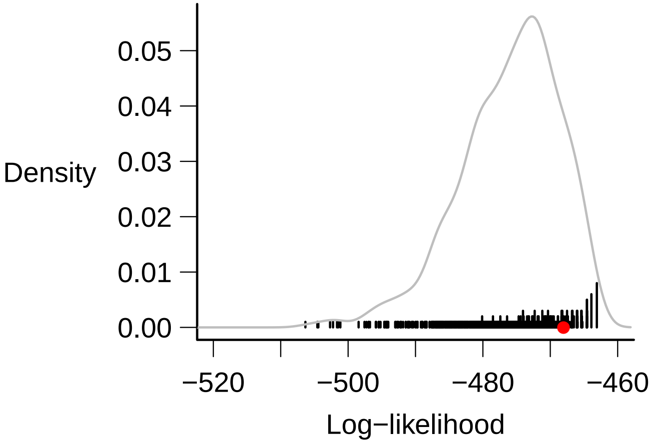
- Simulate for 80 days, with a settlement rate (`tot.settle`) of 5% day:
- 29 distinct individuals, size range 5–30, total of 618 fish-days (observations).
- Sample fates only (1000 MCMC steps), fixed demographic parameters

Sampled 1000 unique fates (but only 895 unique likelihoods)

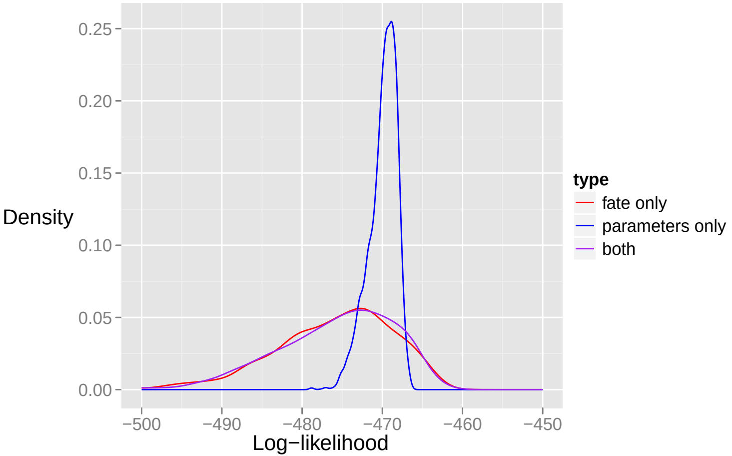
True demography



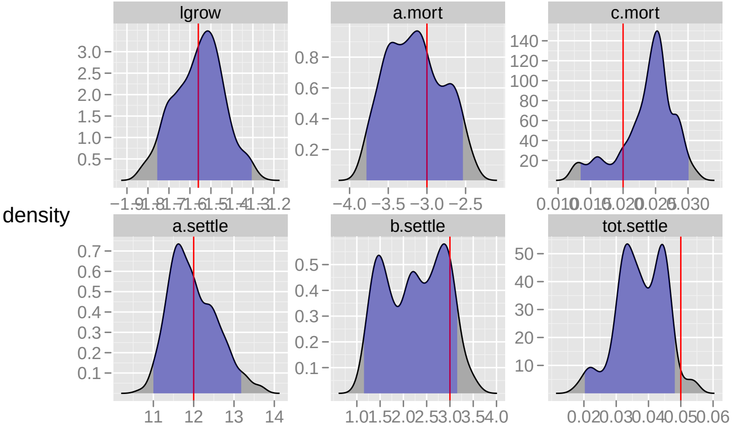
Fate-only results



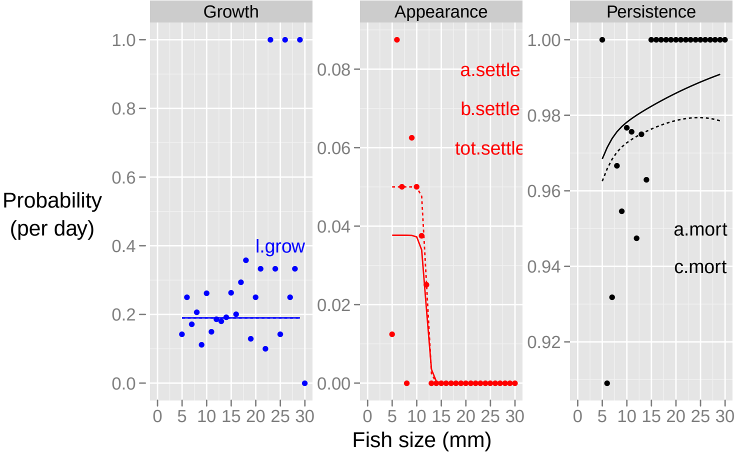
Sampling fates, parameters, both . . .



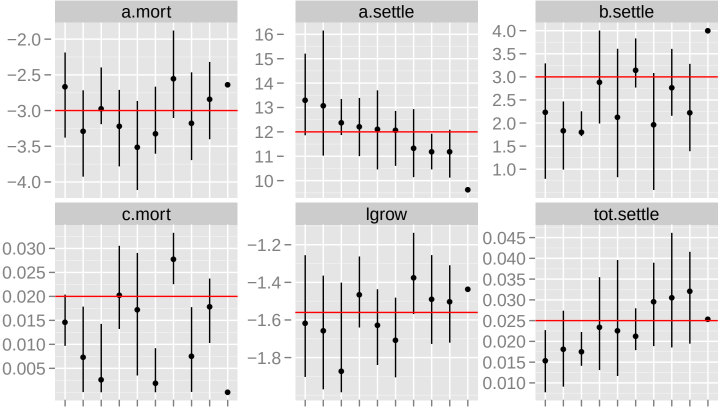
Parameter estimates 1



Parameter estimates 2

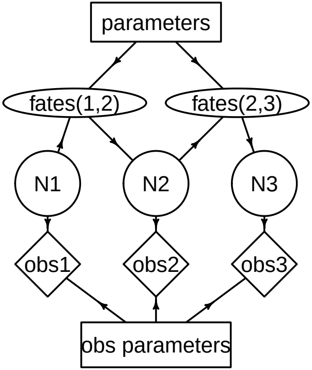


Parameter estimates 3



Outline

Adding errors



Back to the more typical graph: incorporate information from $N(t-1)$, $N(t+1)$, and current observation

New algorithm

For each individual:

- Pick true state at $N(t)$, from **all** possible states, conditioning on $N(t-1)$
- Pick identity at time $N(t+1)$ from feasible set: update probabilities
- Pick corresponding observation from observed(t): update probabilities
- Calculate likelihood, M-H update

Open questions, caveats

- Will it work ??
- Add complexities:
  - Multiple populations
  - Spatial variation, correlation among parameters
- Simplify/generalize code
- Speed?

Conclusions & musings

- MCMC as powerful technology
- . . . can it be democratized?
- Data limitations shift over time:
  hierarchical models are great for “modern” (high-volume, low-quality) data