

1 Fitting and assumptions

Fitting models to data

- What is the basic problem?
 - Given some observations, what are the “best” parameters for a model to fit it?
 - * For that matter, what is the best *model*?
- Why do we want to know this?
 - Want to predict the future
 - Want to evaluate the effects of various control strategies
 - Want to make inferences about individual-level mechanisms

Fitting approaches

- Least squares fit
- Likelihood fits
- Likelihood fits with process error
- Bayesian approaches

Least squares fit

- Find the model parameters that give the best fit to your data, measured by least squares
 - Optimization routine is usually required
- What is the justification for this approach?
 - The least-squares fit is the maximum likelihood fit, under the assumption that deviations are normal, with constant variation
- What do we think of this justification?
- What are some alternatives?
 - Weighted least squares is not much harder

Explicit likelihood

- Once we recognize where LS comes from, and recognize that we have computers, we can think of many different variations:
 - Choose numbers from a binomial distribution
 - Or simply use a normal with variance that matches the binomial
 - * Or variance that matches a beta-binomial
- We have a tremendous amount of flexibility to write down a likelihood, and have the computer optimize it for us, and also search the likelihood space so we can do stats
- Running a deterministic model and using it to calculate likelihood is a powerful, flexible approach
 - Do you have any concerns with this approach?

Process error

- In fact, our deterministic model is not usually a good representation of what we think is going on:
 - **Answer:** Individual-level random events
 - **Answer:** Large-scale random events
- A random event is anything that we cannot reasonably expect to account for in our model
 - Weather, economic changes, individual-scale coincidences

Estimating a random process

- We distinguish two main types of error: process error and observation error
- If we have no process error, we can use simple maximum likelihood methods (e.g., weighted least squares) to fit our model to the data
 - Need to estimate starting conditions
- If we have no observation error, we can use simple maximum likelihood with a *step-by-step* approach

The grim re world

- What if we have errors both in process and in observation?
 - We must make a model of the “hidden” variables, and see how well we can get it to fit the data
 - The likelihood for a set of parameters will be an integral over all possible values of the hidden variables

Likelihood and Bayesian approaches

- There are *philosophical* differences between likelihood and Bayesian approaches, but from a *practical* point of view, they are very similar
- In ML, the probability that a set of parameters is correct is assumed to be proportional (in some sense) to the likelihood of the observed data given these parameters
 - This is then interpreted in a frequentist sense
- In Bayesian inference, the probability that a set of parameters is correct is assumed to be proportional to the likelihood of the data they generate
 - ...
 - multiplied by a prior probability

Non-informative priors

- Typically, we want to start a Bayesian problem using a non-informative prior (we can add different types of information explicitly from there)
 - A uniform prior is often a good choice
 - It is often good to have the prior be uniform on the log-transformed scale
 - * The probability of being between 1 and 10 is the same as the probability of being between 10 and 100

MCMC sampling

- Bayesian methods are very flexible
- We can write down reasonable priors, and likelihoods, to cover a wide variety of assumptions and situations
- Unfortunately, we usually can't *integrate* – calculate the denominator of Bayes' formula
- Instead we use Markov chain Monte Carlo methods to sample randomly from the posterior distribution
 - Simple to do, but hard to know how long you have to simulate to get a good sample of the posterior

MCMC sampling

- Rules that assure that we will visit each point in parameter space in proportion to its likelihood . . . eventually
 - How do we know if we've been simulating long enough?
- Two rules of thumb for checking:
 - Check sensitivity of results to how long you simulate
 - Repeat the whole process with a different starting point (in parameter space)
- Packages are available: JAGS, BuGS (search for winbugs)

Sequential MC

- In some cases we need to break the problem down into parts in order to find good solutions.
- Sequential MC methods (related to particle filtering) allow solutions to evolve by progressively tightening criteria
- There is a “plug-and-play” package for R, called POMP, that is supposed to help analyze dynamical systems with both process error and observation error