Introduction to Applied Bayesian Analysis in Wildlife Ecology

Jeffrey W. Doser

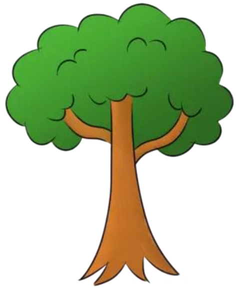
May 11, 2024

Course Website

https://doserjef.github.io/TWS24-Bayesian-Workshop/

A bit about me

Bayesian statistics



Wildlife and natural resources conservation

Software development

Course Overview

Model building

Frequentist vs. Bayesian

Bayesian basics

Bayesian linear models

Bayesian GLMs

Hierarchical Bayesian models

Course Learning Objectives

1. Understand foundational differences between frequentist and Bayesian approaches
2. Obtain a basic understanding of Bayesian analysis (and associated jargon) to impress your colleagues (and understand methods sections of papers)
3. Fit key statistical models such as linear models, generalized linear models, and mixed models in a Bayesian framework in R
4. Generate a solid Bayesian toolbox that you can build upon for your own work

Model building and modes of inference

What is a model?

An explanation of an observed pattern

Simplified version of nature

Statements that explain *why* the observations have occurred

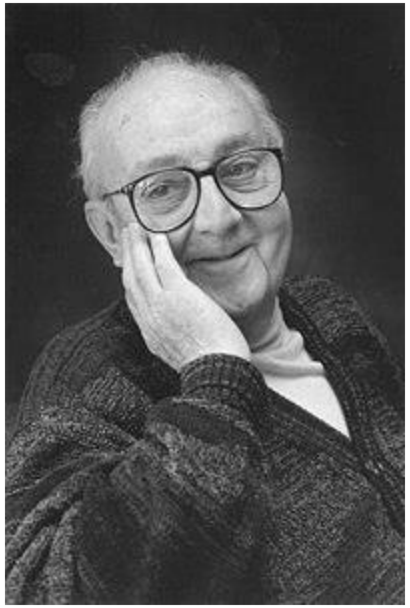
What is a model?

An explanation of an observed pattern

Simplified version of nature

Statements that explain *why* the observations have occurred

"All models are wrong, but some are useful"
- George Box



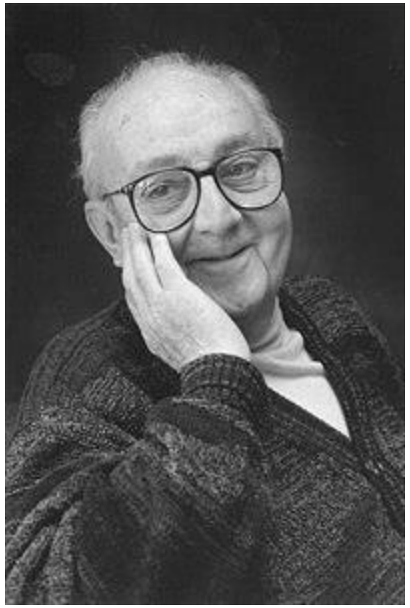
What is a model?

An explanation of an observed pattern

Simplified version of nature

Statements that explain *why* the observations have occurred

"All models are wrong, but some are useful"
- George Box



Nature is too complex. We need to simplify the complexity

Model objectives

The type of model we use depends on the objective of the model

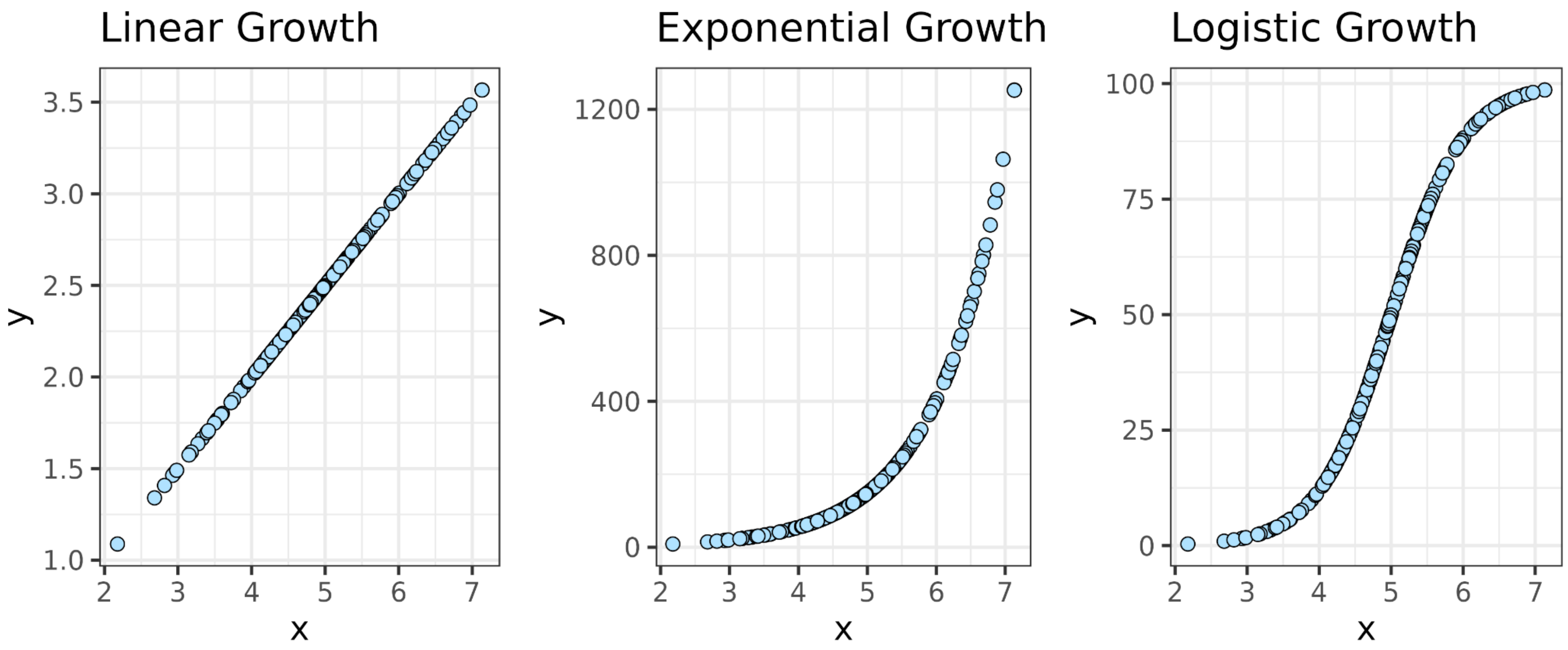
Model objectives

The type of model we use depends on the objective of the model

Inference:
understand
mechanisms

Prediction:
description

Mathematical models



$$y = \beta_0 + \beta_1 x$$

$$\frac{dN}{dt} = rN$$

$$\frac{dN}{dt} = rN\left(\frac{K-N}{K}\right)$$

Mathematical notation

Are there any differences between these equations?

$$y = \beta_0 + \beta_1 x$$

$$y = mx + b$$

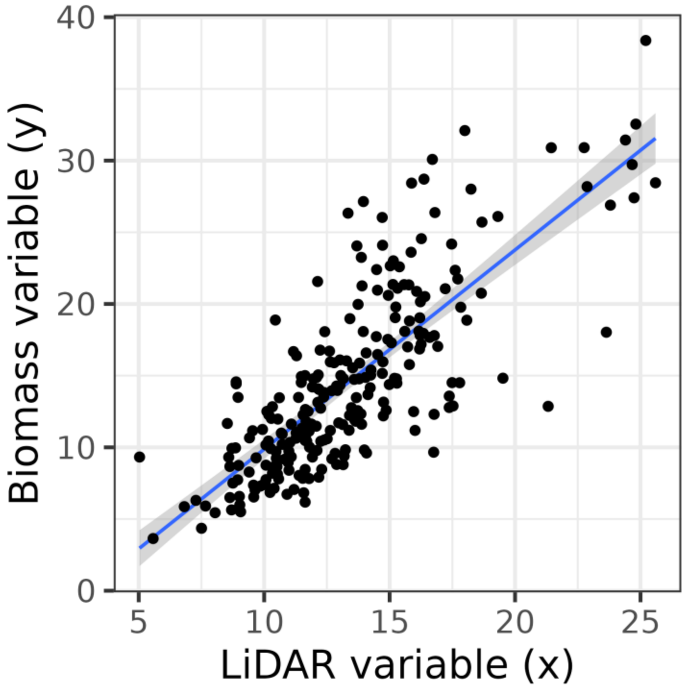
Statistical Models

- A description of a system composed of variables but where one or more **random variables** are related to other variables
- Explicitly acknowledge stochasticity (uncertainty/error) in systems
- Response = systematic component + random component
- Example: simple linear regression

Statistical Models

- A description of a system composed of variables but where one or more **random variables** are related to other variables
- Explicitly acknowledge stochasticity (uncertainty/error) in systems
- Response = systematic component + random component
- Example: simple linear regression

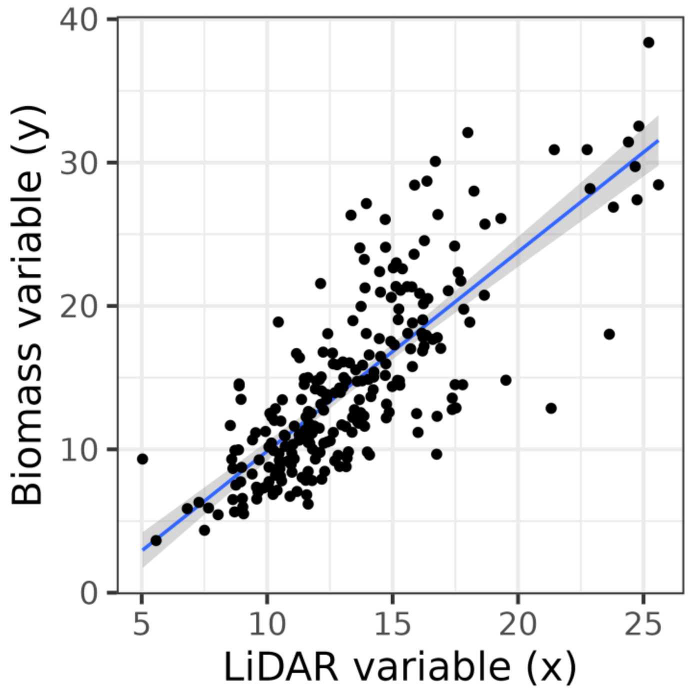
$$y = \beta_0 + \beta_1 x + \epsilon$$



Statistical Models

- A description of a system composed of variables but where one or more **random variables** are related to other variables
- Explicitly acknowledge stochasticity (uncertainty/error) in systems
- Response = systematic component + random component
- Example: simple linear regression

$$\underbrace{y = \beta_0 + \beta_1 x}_{\text{Systematic component}} + \epsilon$$

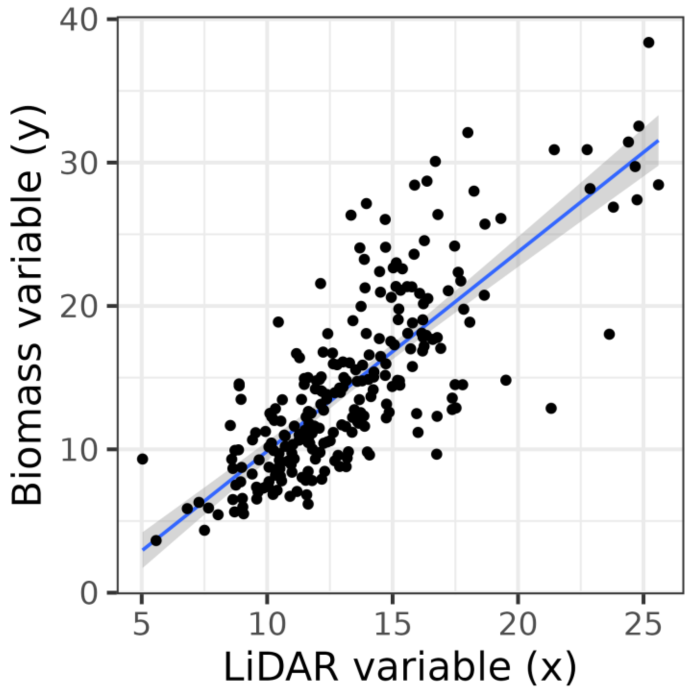


Statistical Models

- A description of a system composed of variables but where one or more **random variables** are related to other variables
- Explicitly acknowledge stochasticity (uncertainty/error) in systems
- Response = systematic component + random component
- Example: simple linear regression

$$y = \underbrace{\beta_0 + \beta_1 x}_{\text{Systematic component}} + \epsilon$$

Random error component



Why do we need statistical models in wildlife ecology?

Three essential kinds of random variability

Why do we need statistical models in wildlife ecology?

Three essential kinds of random variability

1. **Measurement error**: variability imposed by our imperfect observation of the world. It is often modeled by adding normally distribution variability around a mean value
   - **Sampling error**: a type of measurement error that results from us drawing conclusions about a *population* using a smaller *sample* of that population

Why do we need statistical models in wildlife ecology?

Three essential kinds of random variability

2. **Demographic Stochasticity**: the innate variability in outcomes due to random processes. For example, flipping a coin 20 times, you might get 10, 9, or 11 heads, even though you're flipping the coin the same way each time.

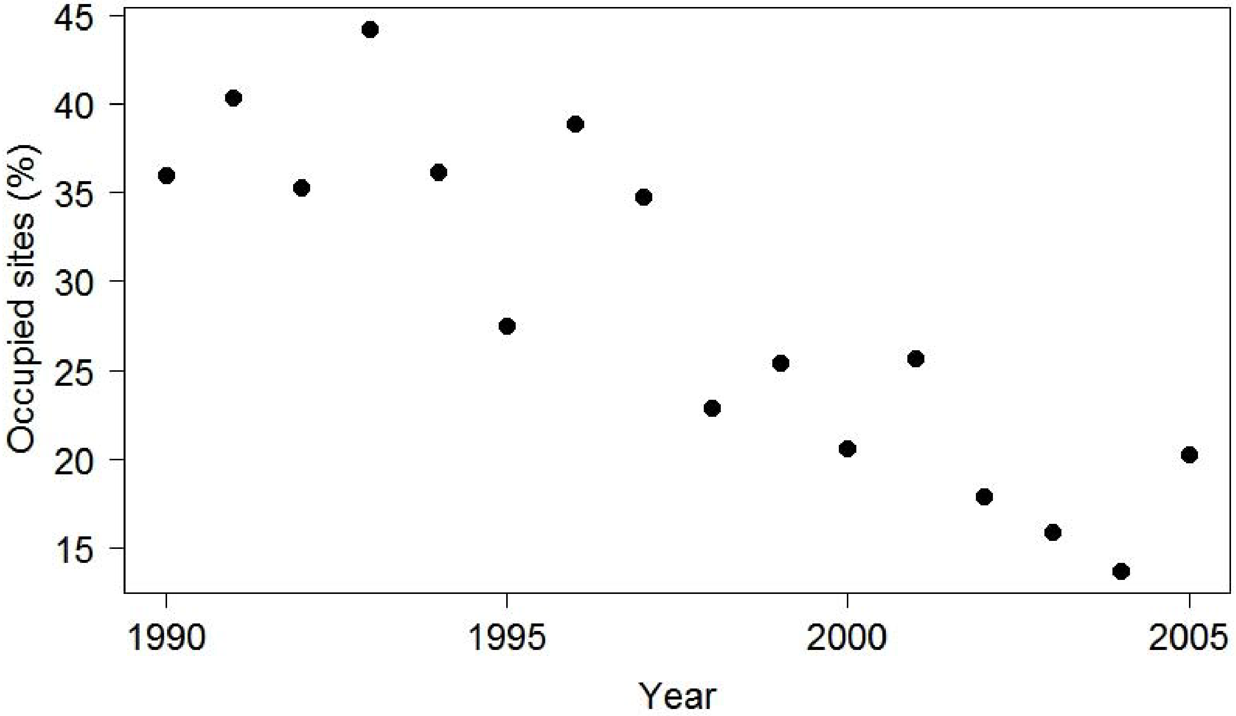
Why do we need statistical models in wildlife ecology?

Three essential kinds of random variability

3. Environmental Stochasticity: variability imposed from "outside" the ecological system, such as climatic, seasonal, or topographic variation

A simple example of a statistical model

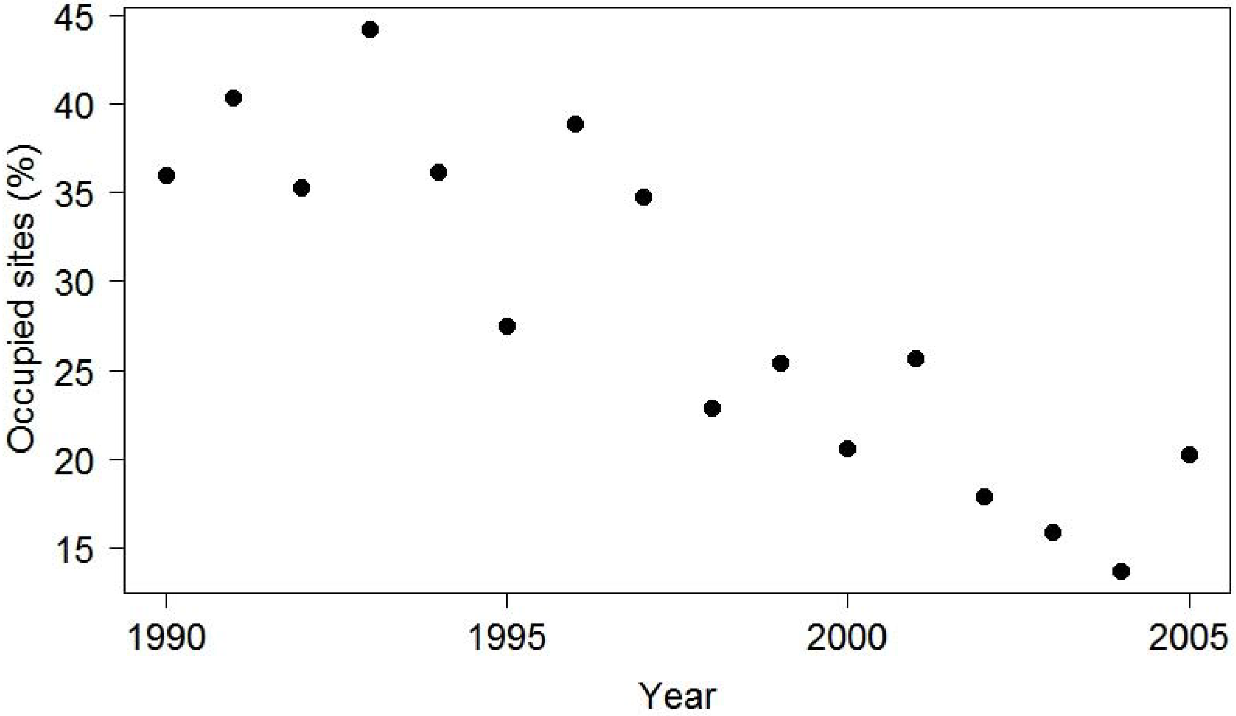
Spotted Salamander



A simple example of a statistical model

Spotted Salamander

$$y = \beta_0 + \beta_1 x + \epsilon$$

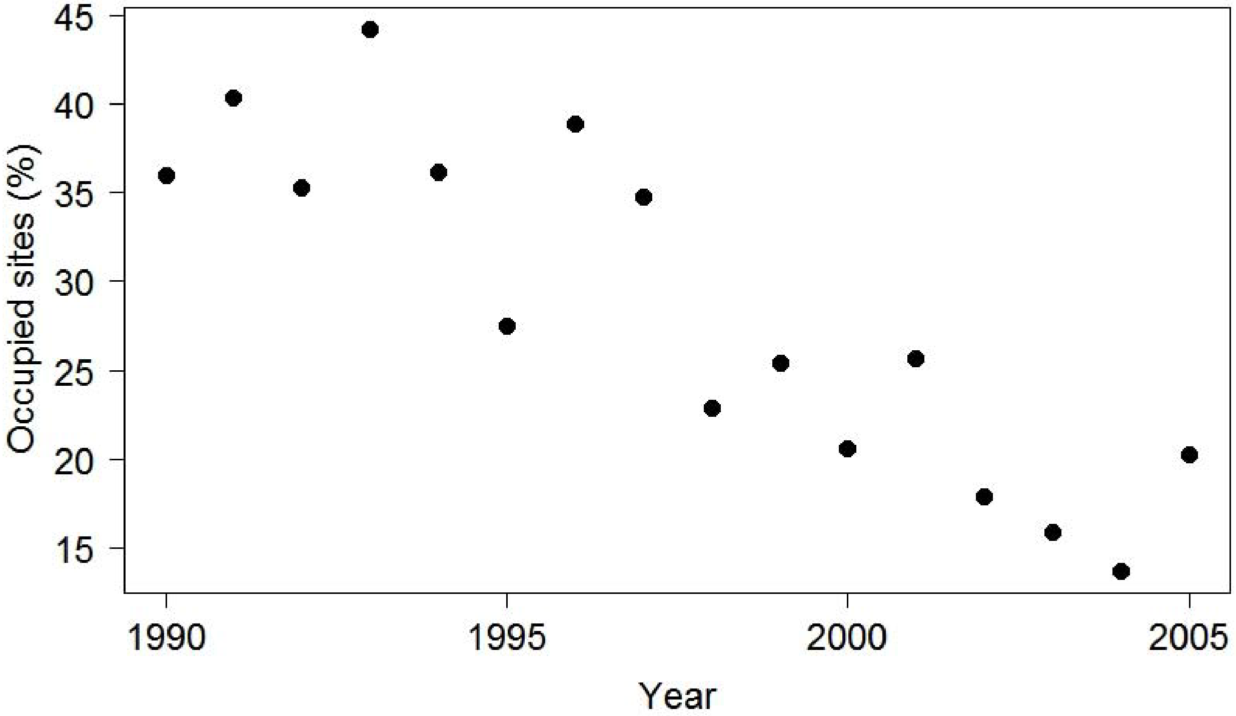


A simple example of a statistical model

Spotted Salamander

$$y = \beta_0 + \beta_1 x + \epsilon$$

$$\epsilon \sim \text{Normal}(0, \sigma^2)$$



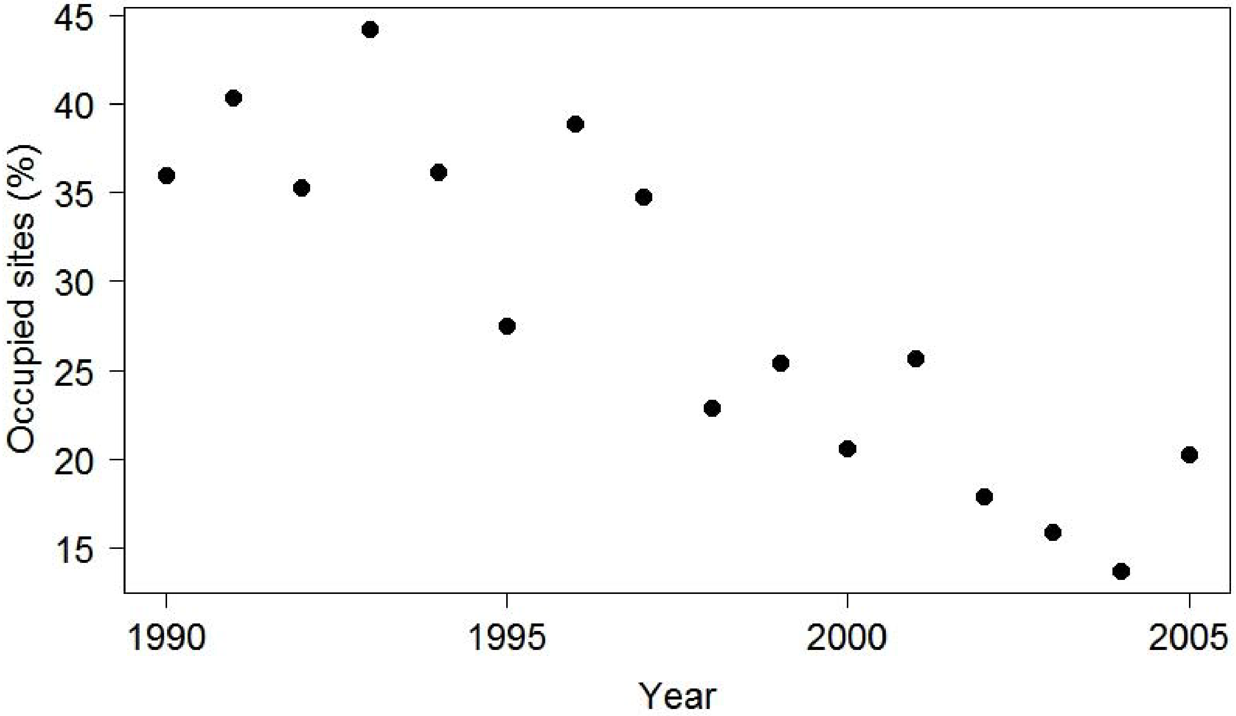
A simple example of a statistical model

Spotted Salamander

$$y = \beta_0 + \beta_1 x + \epsilon$$

$$\epsilon \sim \text{Normal}(0, \sigma^2)$$

"is distributed as"



A simple example of a statistical model

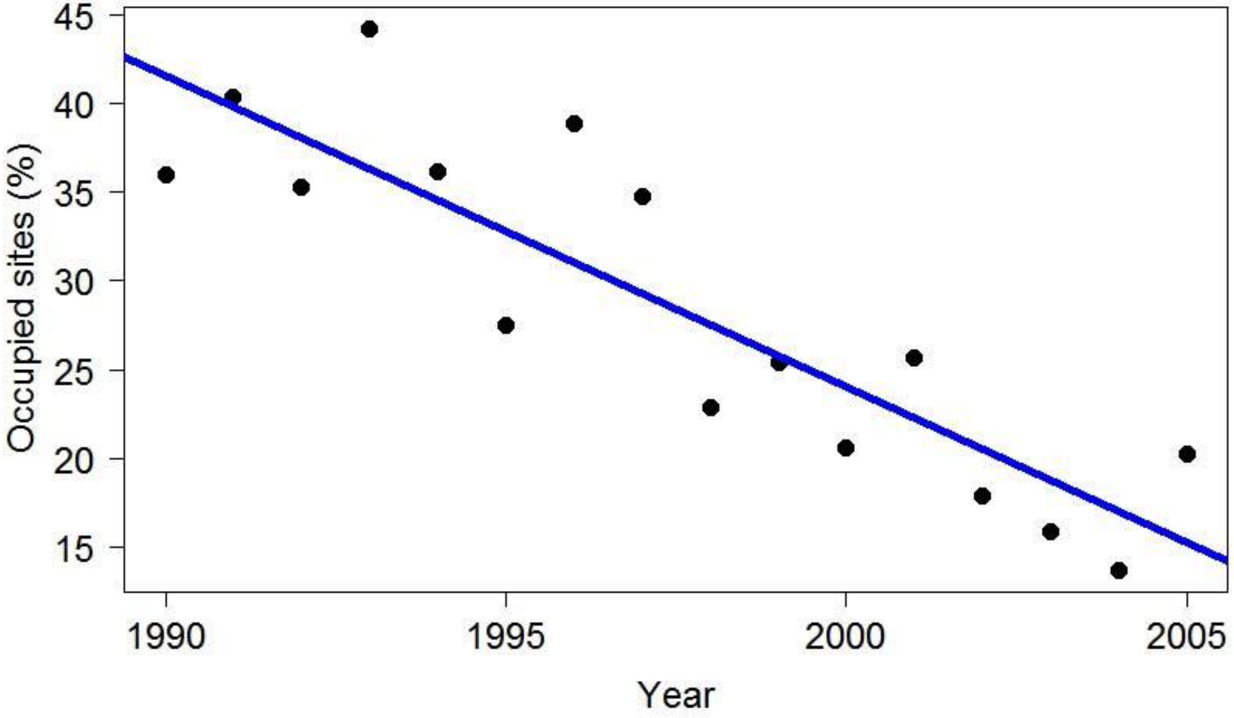
Spotted Salamander

$$y = \beta_0 + \beta_1 x + \epsilon$$

$$\epsilon \sim \text{Normal}(0, \sigma^2)$$

"is distributed as"

Trend Estimate



A simple example of a statistical model

Spotted Salamander

$$y = \beta_0 + \beta_1 x + \epsilon$$

$$\epsilon \sim \text{Normal}(0, \sigma^2)$$

An alternative, but equivalent model description

$$y \sim \text{Normal}(\beta_0 + \beta_1 x, \sigma^2)$$

Statistical models

- **Parametric Statistical Models**: draw statistical conclusions from the data on the ecological process of interest using probability distributions thought to have produced the data

Statistical models

- **Parametric Statistical Models**: draw statistical conclusions from the data on the ecological process of interest using probability distributions thought to have produced the data
- **Non-parametric Statistical Models**: do not assume the data belong to a particular distribution (but make other assumptions)

Statistical models

- **Parametric Statistical Models**: draw statistical conclusions from the data on the ecological process of interest using probability distributions thought to have produced the data
- **Non-parametric Statistical Models**: do not assume the data belong to a particular distribution (but make other assumptions)
- **Generalized Linear Models (GLMs)**: quintessential parametric statistical model
    - A majority of the statistical models used in wildlife ecology can be viewed as an extension of a GLM

Statistical models

Two frequently used GLMs in wildlife ecology

- **Normal response**
  - Random component $y \sim \text{Normal}(\text{mean} = \mu, \text{variance} = \sigma^2)$
  - Systematic component $\mu = \beta_0 + \beta_1 x$

Statistical models

Two frequently used GLMs in wildlife ecology

- **Normal response**
  - Random component $y \sim \text{Normal}(\text{mean} = \mu, \text{variance} = \sigma^2)$
  - Systematic component $\mu = \beta_0 + \beta_1 x$
- **Binomial response**
  - Random component: $y \sim \text{Binomial}(\text{number of trials} = N, \text{success probability} = p)$
  - Systematic component: $\text{logit}(p) = \beta_0 + \beta_1 x$

What is a parameter?

- Unknown effects that we want to estimate
- More generally, a summary measure of the population characteristic of interest

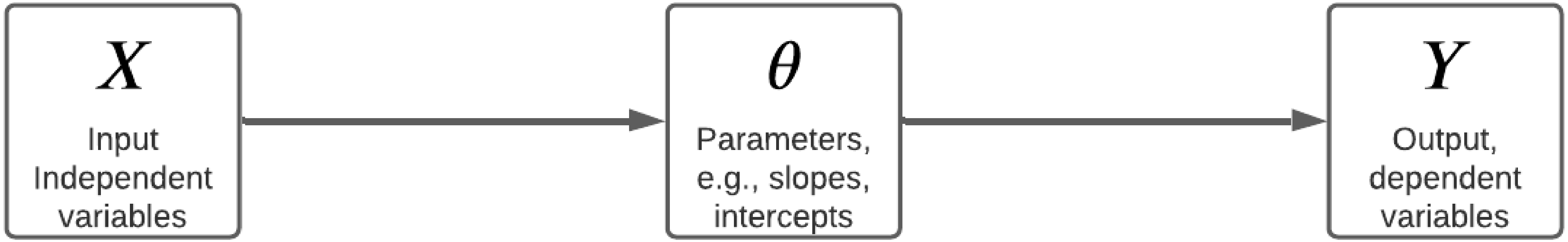
What is a parameter?

- Unknown effects that we want to estimate
- More generally, a summary measure of the population characteristic of interest
- **Statistical Inference**: collect data from a sample of the population to estimate parameters that describe the population

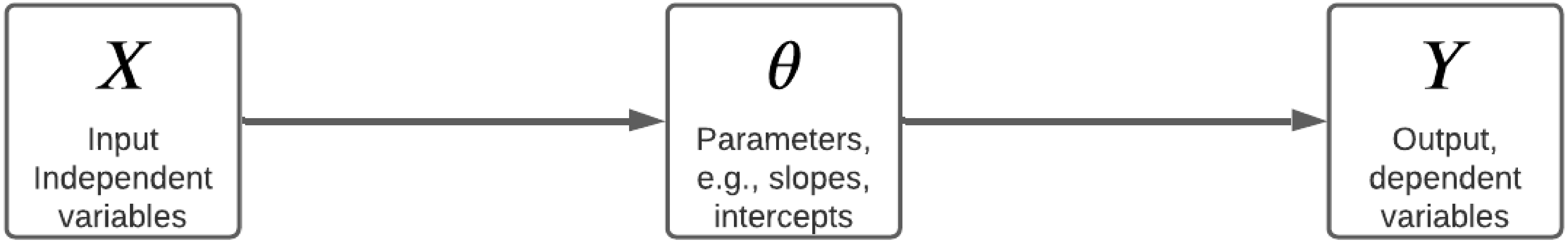
What is a parameter?

- Unknown effects that we want to estimate
- More generally, a summary measure of the population characteristic of interest
- **Statistical Inference**: collect data from a sample of the population to estimate parameters that describe the population
- Examples
    - Average biomass of trees in a forest (mean)
    - Total number of mammals in a management unit (total)
    - Relationship between number of spring growing degree days and monarch abundance (slope)
    - Survival probability of breeding birds over the non-breeding season (probability)

Analysis of a statistical model

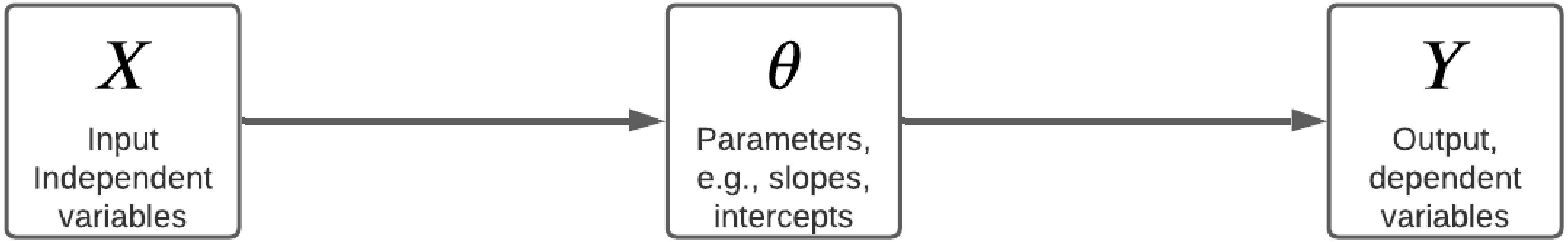


Analysis of a statistical model



- Data (Y) viewed as a result of random processes. Independent variables (X) typically viewed as known.
- How can we guess at values of θ?

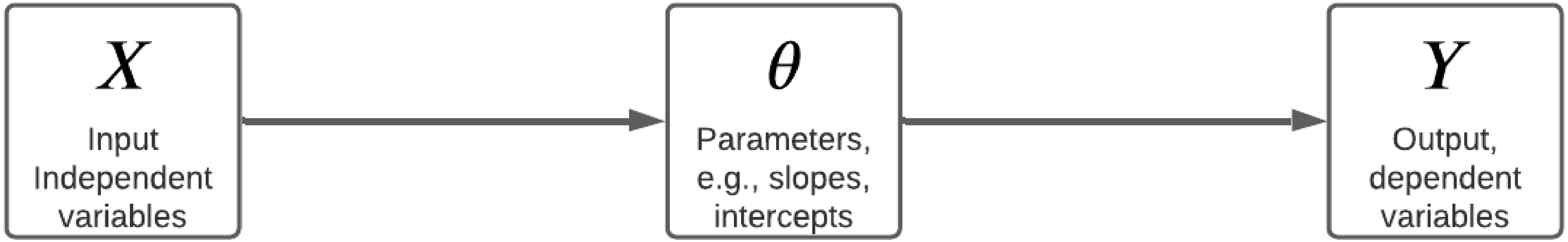
Analysis of a statistical model



- Data (Y) viewed as a result of random processes. Independent variables (X) typically viewed as known.
- How can we guess at values of θ?
- Frameworks of statistical inference:
  - Method of moments
  - Ordinary-least squares (OLS)
  - Maximum likelihood
  - Bayesian analysis

Frequentist analysis

Analysis of a statistical model



- Data (Y) viewed as a result of random processes. Independent variables (X) typically viewed as known.
- How can we guess at values of θ?
- Frameworks of statistical inference:
  - Method of moments
  - Ordinary-least squares (OLS)
  - Maximum likelihood
  - Bayesian analysis

Frequentist analysis

A simple example of a statistical model

Spotted Salamander

$$y = \beta_0 + \beta_1 x + \epsilon$$

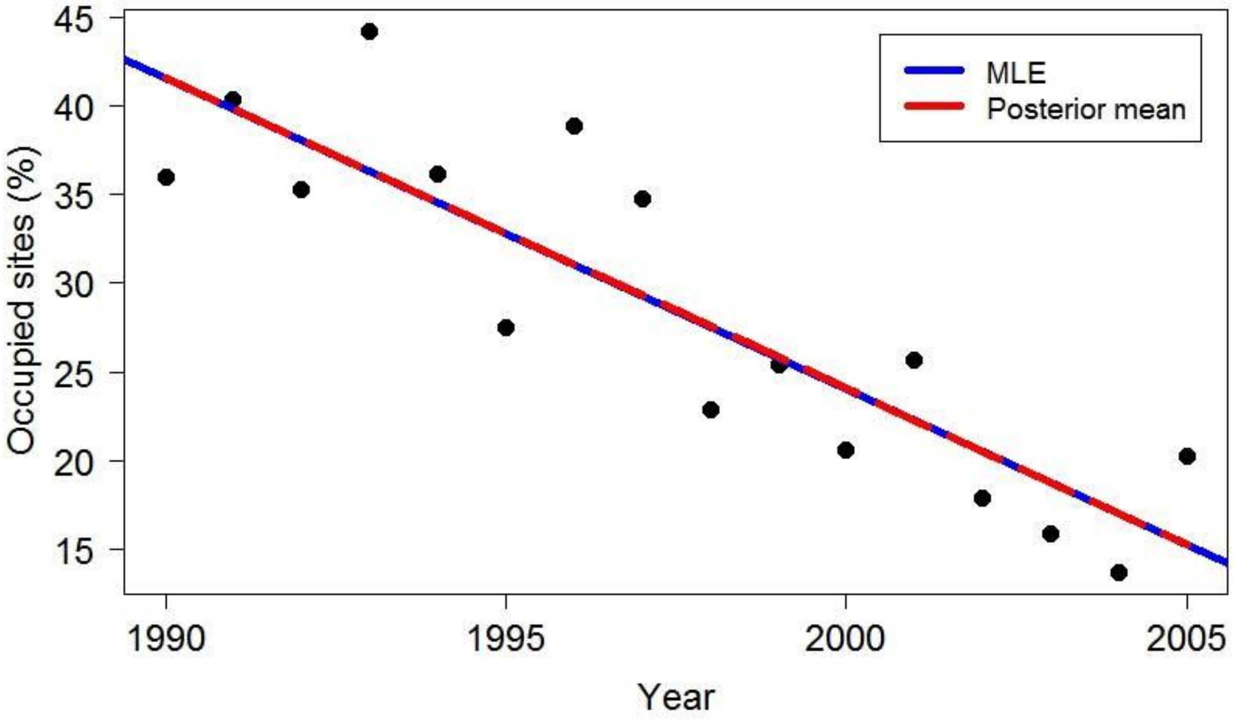
$$\epsilon \sim \text{Normal}(0, \sigma^2)$$

"is distributed as"

Trend Estimate

$\beta_1 = -1.754$

$\beta_1 = -1.754$



Maximum Likelihood Estimation

Frequentist analysis

Example: estimate the detection probability of spring peepers

- Originally released n = 50 in an artificial pond
- Recaptured y = 20

Maximum likelihood analysis

- Based upon the concept of the *Likelihood*
    - Bayesian inference also relies on the likelihood
- Likelihood is probability in reverse

Maximum likelihood analysis

- Based upon the concept of the *Likelihood*
  - Bayesian inference also relies on the likelihood
- Likelihood is probability in reverse

Probability: parameters given and data unknown

- Given that Pr(frog detected) = 0.4, how many of the 50 frogs will probably be detected?

Maximum likelihood analysis

- Based upon the concept of the *Likelihood*
  - Bayesian inference also relies on the likelihood
- Likelihood is probability in reverse

Probability: parameters given and data unknown

- Given that Pr(frog detected) = 0.4, how many of the 50 frogs will probably be detected?
- 20 will probably be detected

Maximum likelihood analysis

- Based upon the concept of the *Likelihood*
  - Bayesian inference also relies on the likelihood
- Likelihood is probability in reverse

Probability: parameters given and data unknown

- Given that Pr(frog detected) = 0.4, how many of the 50 frogs will probably be detected?
- 20 will probably be detected

Likelihood: data given and parameters unknown

- Given that 20 of my 50 frogs were recaptured, what is the most likely tadpole detection probability?

Maximum likelihood analysis

- Based upon the concept of the *Likelihood*
  - Bayesian inference also relies on the likelihood
- Likelihood is probability in reverse

Probability: parameters given and data unknown

- Given that Pr(frog detected) = 0.4, how many of the 50 frogs will probably be detected?
- 20 will probably be detected

Likelihood: data given and parameters unknown

- Given that 20 of my 50 frogs were recaptured, what is the most likely tadpole detection probability?
- 0.4 is the most likely detection probability

Maximum likelihood analysis

- One way to estimate θ: **Maximum Likelihood Estimation**
- **Sampling distribution**: p(y | θ)
  - Read as "probability of observing data y, given a fixed parameter value θ
- Note the probability statement is about the data, not about θ
- Probability is defined as the long-run frequency in hypothetical replicate data sets
- Example: binomial sampling distribution (y ~ Binomial(θ, N))

$$p(y \mid \theta) = \frac{N!}{y!(N-y)!}\theta^y(1-\theta)^{N-y}$$

Maximum likelihood analysis

- One way to estimate θ: **Maximum Likelihood Estimation**
- **Sampling distribution**: p(y | θ)
  - Read as "probability of observing data y, given a fixed parameter value θ
- Note the probability statement is about the data, not about θ
- Probability is defined as the long-run frequency in hypothetical replicate data sets
- Example: binomial sampling distribution (y ~ Binomial(θ, N))

"Given" or "conditional on"

$$p(y \mid \theta) = \frac{N!}{y!(N-y)!}\theta^y(1-\theta)^{N-y}$$

Maximum likelihood analysis

- One way to estimate θ: **Maximum Likelihood Estimat**ion
- **Sampling distribution**: p(y | θ)
  - Read as "probability of observing data y, given a fixed parameter value θ
- Note the probability statement is about the dat
- Probability is defined as the long-run frequency
  replicate data sets
- Example: binomial sampling distribution (y ~ Bin

"Given" or "conditional on"

$$p(y \mid \theta) = \frac{N!}{y!(N-y)!}\theta^y(1-\theta)^{N-y}$$

Maximum likelihood: fundamental idea

- For a given statistical model, maximum likelihood finds a set of parameters that makes the observed data *most likely* to occur
- A good choice of θ is that which maximizes the function value of the sampling distribution of the data set
- Answer the question: "what parameter value is most likely to have generated the data I observed?"

Maximum likelihood: fundamental idea

- Probability describes a function of the outcome given a fixed parameter value
  - If a coin is flipped 10 times and it is a fair coin, what is the probability of it landing heads every time?
- Likelihood is used when describing a function of a parameter given an outcome
  - If a coin is flipped 10 times and it has landed heads 10 times, what is the likelihood that the coin is fair?

Maximum likelihood: fundamental idea

- **Likelihood Function**: read the sampling distribution "in reverse" as a function of θ.

$$p(y \mid \theta) = L(\theta \mid y)$$

Maximum likelihood: fundamental idea

- **Likelihood Function**: read the sampling distribution "in reverse" as a function of θ.

$$p(y \mid \theta) = L(\theta \mid y)$$

$$L(\theta \mid y) = \frac{N!}{y!(N-y)!}\theta^y(1-\theta)^{N-y}$$

Maximum likelihood: fundamental idea

- **Likelihood Function**: read the sampling distribution "in reverse" as a function of θ.

$$p(y \mid \theta) = L(\theta \mid y)$$

$$L(\theta \mid y) = \frac{N!}{y!(N-y)!}\theta^y(1-\theta)^{N-y}$$

- Call the value of θ that maximizes L(θ | y) the **Maximum Likelihood Estimate (MLE)**

Maximum likelihood: how to find the MLE?

Maximum likelihood: how to find the MLE?

Analytically

- Take the derivative of the likelihood and find maximum values
- Use calculus to determine the value of θ that maximizes the likelihood function
- Not commonly used. Often only possible for simple models.

Maximum likelihood: how to find the MLE?

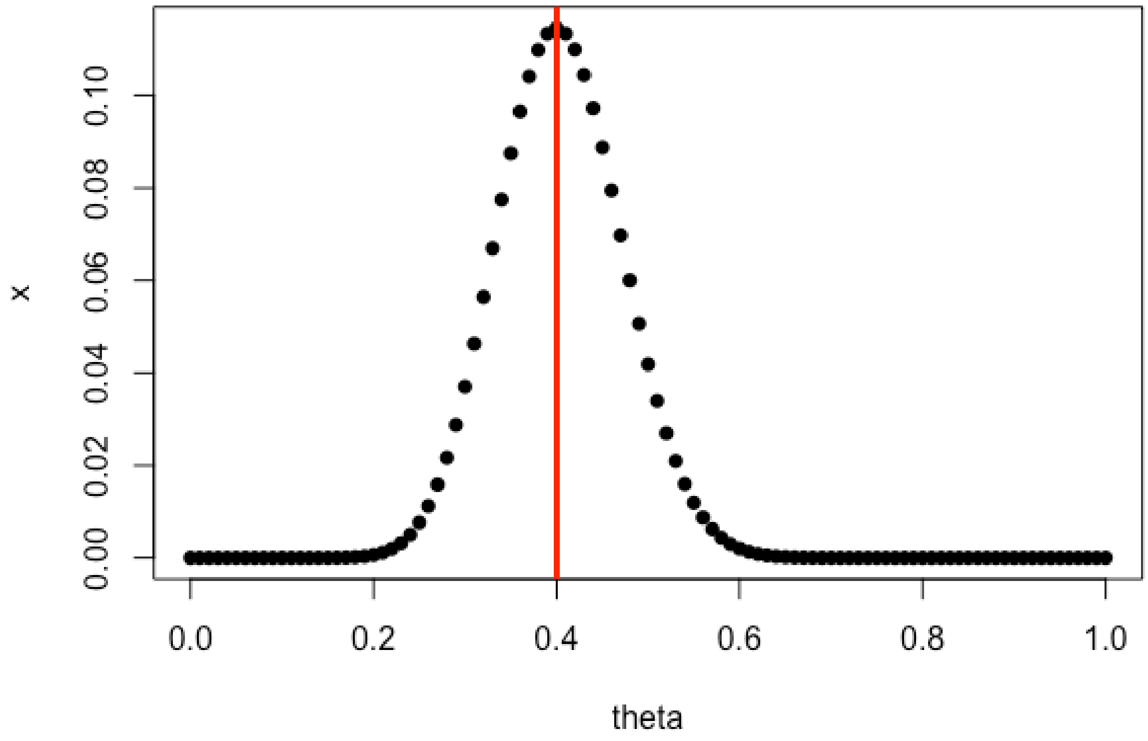
Analytically

- Take the derivative of the likelihood and find maximum values
- Use calculus to determine the value of θ that maximizes the likelihood function
- Not commonly used. Often only possible for simple models.

Numerically

- Estimation by brute force
- Try out and plot a large number of values for θ
- Find the value that maximizes the log likelihood (or minimizes the negative log likelihood)

Maximum likelihood: brute force numerical estimation



$$L(\theta \mid 20) = \frac{50!}{20!(50-20)!}\theta^{20}(1-\theta)^{50-20}$$

Maximum likelihood: brute force numerical estimation

- Numerical estimation by function minimization
  - `optim()` function in R (broad)
  - Specialized functions in R (e.g., `nlm()`, `glm()`)
- Basic approach
  - Specify the likelihood function
  - Take the derivative through numerical approximation
  - The MLE is the value that minimizes the negative log-likelihood
- Minimizing the negative log-likelihood is equivalent to maximizing the likelihood, but more computationally efficient

Benefits of the frequentist approach

- "Gold standard" in wildlife ecology

Benefits of the frequentist approach

- "Gold standard" in wildlife ecology
- Well-developed, consistent, and reliable. Works with a large set of problems.

Benefits of the frequentist approach

- "Gold standard" in wildlife ecology
- Well-developed, consistent, and reliable. Works with a large set of problems.
- Desirable mathematical properties
  - Unbiased at large sample sizes and consistent
  - Have approximate normal distributions and sample variances, which is great for generating confidence intervals

Benefits of the frequentist approach

- "Gold standard" in wildlife ecology
- Well-developed, consistent, and reliable. Works with a large set of problems.
- Desirable mathematical properties
  - Unbiased at large sample sizes and consistent
  - Have approximate normal distributions and sample variances, which is great for generating confidence intervals
- Lots of (easy to use) software packages

Drawbacks of the frequentist approach

- Working out likelihoods can be difficult, especially when model structure is complicated

Drawbacks of the frequentist approach

- Working out likelihoods can be difficult, especially when model structure is complicated
- Biased at small sample sizes (and difficult to determine what sample size is big enough)

Drawbacks of the frequentist approach

- Working out likelihoods can be difficult, especially when model structure is complicated
- Biased at small sample sizes (and difficult to determine what sample size is big enough)
- Tricky interpretation
    - **Confidence interval**: if this procedure was repeated on multiple samples, the 95% confidence interval (which is different for each sample) would encompass the true population parameter 95% of the time

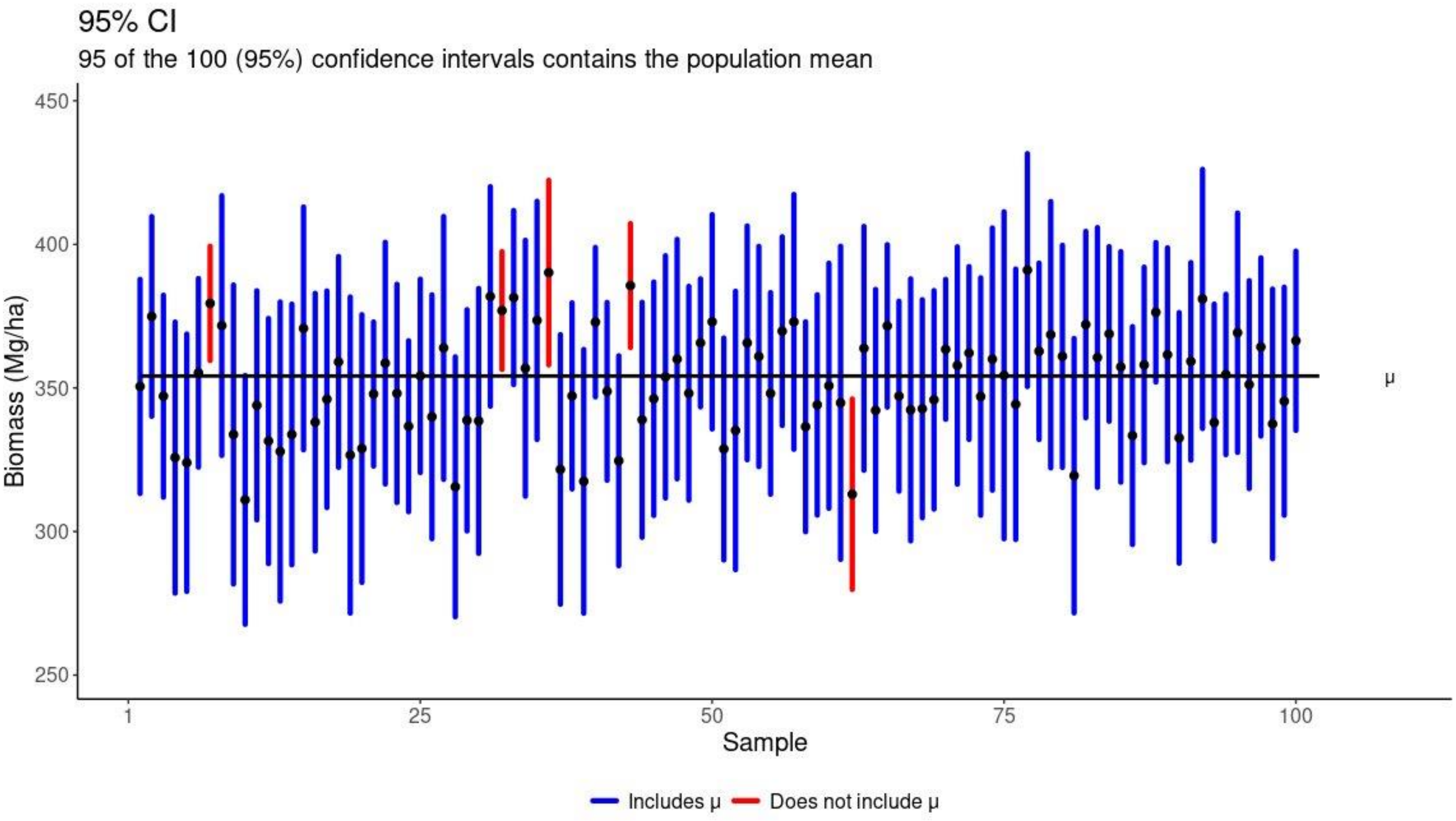
Drawbacks of the frequentist approach

- Working out likelihoods can be difficult, especially when model structure is complicated
- Biased at small sample sizes (and difficult to determine what sample size is big enough)
- Tricky interpretation
    - **Confidence interval**: if this procedure was repeated on multiple samples, the 95% confidence interval (which is different for each sample) would encompass the true population parameter 95% of the time
    - **P-value**: probability of obtaining a result equal to or "more extreme" than what was observed, assuming that the null hypothesis under consideration is true

Drawbacks of the frequentist approach

- Working out likelihoods can be difficult, especially when model structure is complicated
- Biased at small sample sizes (and difficult to determine what sample size is big enough)
- Tricky interpretation
  - **Confidence interval**: if this procedure was repeated on multiple samples, the 95% confidence interval (which is different for each sample) would encompass the true population parameter 95% of the time
  - **P-value**: probability of obtaining a result equal to or "more extreme" than what was observed, assuming that the null hypothesis under consideration is true
  - Frequentist statistics cannot assign a probability to a hypothesis (e.g., no probability associated with the value of a parameter)

Drawbacks: confidence interval interpretation



Questions?

Bayesian Analysis

Recall: fundamental idea of maximum likelihood

- A good choice of θ is that which maximizes the likelihood function
- Answers the question: "What parameter value is most likely to have generated the data I observed?"
- Probability statement is about the data, **not the parameters**

Bayesian analysis

Our overall question is: How do we estimate values of θ?

Bayesian analysis

Our overall question is: How do we estimate values of θ?

Fundamental difference between frequentist and Bayesian analysis:

Bayesian analysis

Our overall question is: How do we estimate values of θ?

Fundamental difference between frequentist and Bayesian analysis:

- **Frequentist**: parameters are unknown, but fixed values

Bayesian analysis

Our overall question is: How do we estimate values of θ?

Fundamental difference between frequentist and Bayesian analysis:

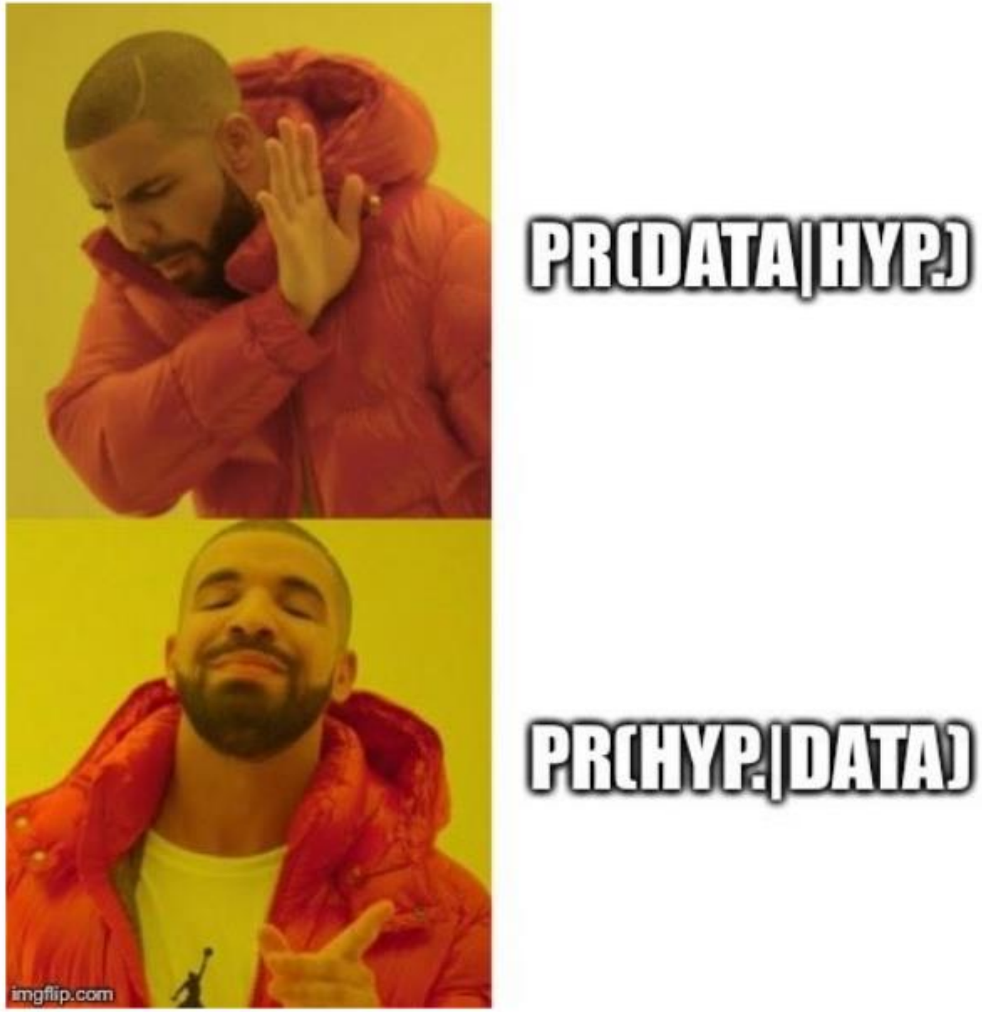
- **Frequentist**: parameters are unknown, but fixed values
- **Bayesian**: parameters are unknown, random quantities

Bayesian analysis

- Bayesian approach: in the face of uncertainty about θ, use conditional probability, $p(\theta \mid y)$
  - "Probability of the parameter, given the observed data"
- Make probability statements about hypotheses (parameters)

Bayesian analysis

- Bayesian approach: in the face of uncertainty about θ, use conditional probability, $p(\theta \mid y)$
  - "Probability of the parameter, given the observed data"
- Make probability statements about hypotheses (parameters)



Basic recipe of a statistical analysis

Basic recipe of a statistical analysis

1. What information do we have?
   - The data ($y = 20$, $N = 50$)

Basic recipe of a statistical analysis

1. What information do we have?
   - The data ($y = 20$, $N = 50$)
2. What do we want to learn?
   - The parameters (θ) we're interested in (e.g., detection probability)

Basic recipe of a statistical analysis

1. What information do we have?
   - The data ($y = 20$, $N = 50$)
2. What do we want to learn?
   - The parameters (θ) we're interested in (e.g., detection probability)
3. What to do?
   - Frequentist: Use the likelihood $L(\theta \mid y)$ to calculate the MLE
     - Probability statement about the data, not parameters

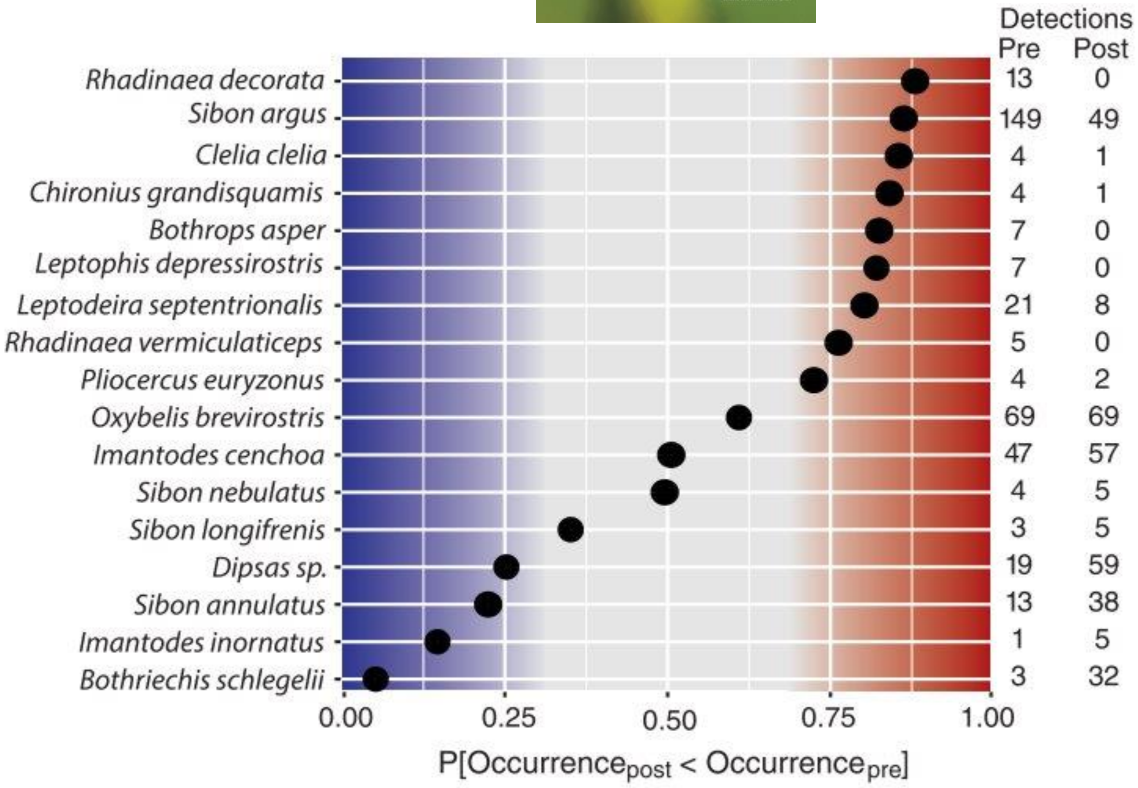
Basic recipe of a statistical analysis

1. What information do we have?
   - The data ($y = 20$, $N = 50$)
2. What do we want to learn?
   - The parameters (θ) we're interested in (e.g., detection probability)
3. What to do?
   - Frequentist: Use the likelihood $L(\theta \mid y)$ to calculate the MLE
     - Probability statement about the data, not parameters
   - Bayesian: Calculate the posterior distribution $p(\theta \mid y)$
     - Probability statement about the parameters, not data

"Probability statement about the parameters"

In a Bayesian framework, you can say things like:

- There is a 97.3% probability the treatment had a positive effect on the outcome
- There is an 88.5% probability the population is declining



Foundation of Bayesian stats: Bayes' Theorem

$$p(A \mid B) = \frac{p(B|A)p(A)}{p(B)} = \frac{p(A,B)}{p(B)}$$

- Mathematical fact of probability
- Thomas Bayes (1702-1761)
- Applied the rule to parameters for parameter estimation
- But why did Bayesian inference only recently become popular?

Bayes' Theorem

$$p(A \mid B) = \frac{p(B|A)p(A)}{p(B)} = \frac{p(A,B)}{p(B)}$$

Bayes' Theorem

"Given" or "conditional on"

$$p(A \mid B) = \frac{p(B|A)p(A)}{p(B)} = \frac{p(A,B)}{p(B)}$$

Bayes' Theorem

"Given" or "conditional on"

$$p(A \mid B) = \frac{p(B|A)p(A)}{p(B)} = \frac{p(A,B)}{p(B)}$$

Probability event
A occurs "given"
event B occurs

Bayes' Theorem

Probability of A and B

"Given" or "conditional on"

$$p(A \mid B) = \frac{p(B|A)p(A)}{p(B)} = \frac{p(A,B)}{p(B)}$$

Probability event
A occurs "given"
event B occurs

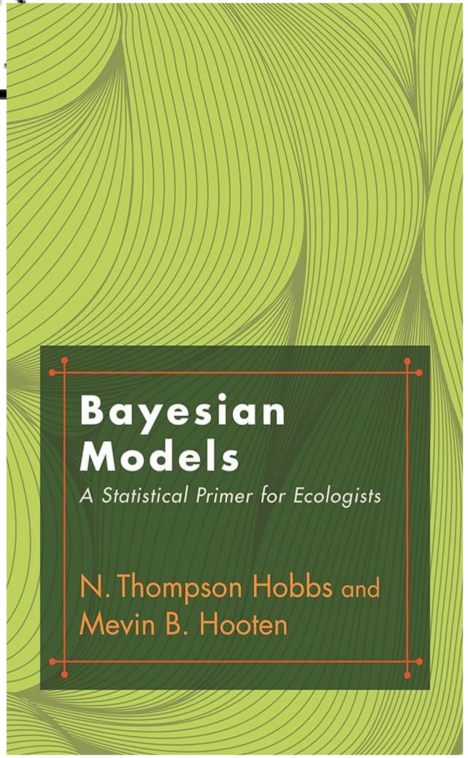
Bayes' Theorem

"Given" or "conditional on"

Probability of A and B

$$p(A \mid B) = \frac{p(B|A)p(A)}{p(B)} = \frac{p(A,B)}{p(B)}$$

Probability event A occurs "given" event B occurs



Bayes' Theorem

$$p(A \mid B) = \frac{p(B|A)p(A)}{p(B)} = \frac{p(A,B)}{p(B)}$$

- Basic tool of Bayesian analysis

Bayes' Theorem

$$p(A \mid B) = \frac{p(B|A)p(A)}{p(B)} = \frac{p(A,B)}{p(B)}$$

- Basic tool of Bayesian analysis
- Allows us to assess the probability of an event (A) in the light of new information (B)

Bayes' Theorem

$$p(A \mid B) = \frac{p(B|A)p(A)}{p(B)} = \frac{p(A,B)}{p(B)}$$

- Basic tool of Bayesian analysis
- Allows us to assess the probability of an event (A) in the light of new information (B)
- The modification of **prior** information to **posterior** information based on observed data (**likelihood**)

Bayes' Theorem

$$p(A \mid B) = \frac{p(B|A)p(A)}{p(B)} = \frac{p(A,B)}{p(B)}$$

- Basic tool of Bayesian analysis
- Allows us to assess the probability of an event (A) in the light of new information (B)
- The modification of **prior** information to **posterior** information based on observed data (**likelihood**)
- Used for estimating parameters and testing hypotheses about those parameters

Bayes' Theorem

$$p(A \mid B) = \frac{p(B|A)p(A)}{p(B)} = \frac{p(A,B)}{p(B)}$$

- Basic tool of Bayesian analysis
- Allows us to assess the probability of an event (A) in the light of new information (B)
- The modification of **prior** information to **posterior** information based on observed data (**likelihood**)
- Used for estimating parameters and testing hypotheses about those parameters
- Use probability to express imperfect knowledge

Bayes' Theorem for statistical inference

$$p(\theta \mid y) = \frac{p(y \mid \theta)p(\theta)}{p(y)} = \frac{p(\theta, y)}{p(y)}$$

Bayes' Theorem for statistical inference

Likelihood

$$p(\theta \mid y) = \frac{p(y \mid \theta)p(\theta)}{p(y)} = \frac{p(\theta, y)}{p(y)}$$

Bayes' Theorem for statistical inference

Likelihood

Prior distribution

$$p(\theta \mid y) = \frac{p(y \mid \theta)p(\theta)}{p(y)} = \frac{p(\theta, y)}{p(y)}$$

Bayes' Theorem for statistical inference

Likelihood

Prior distribution

$$p(\theta \mid y) = \frac{p(y \mid \theta)p(\theta)}{p(y)} = \frac{p(\theta, y)}{p(y)}$$

Marginal distribution of the data

Bayes' Theorem for statistical inference

Likelihood

Prior distribution

$$p(\theta \mid y) = \frac{p(y \mid \theta)p(\theta)}{p(y)} = \frac{p(\theta, y)}{p(y)}$$

Posterior distribution

Marginal distribution of the data

Bayes' Theorem for statistical inference

Likelihood

Prior distribution

$$p(\theta \mid y) = \frac{p(y \mid \theta)p(\theta)}{p(y)} = \frac{p(\theta, y)}{p(y)}$$

Posterior distribution

Margin
the dat

Some numeric constant that we don't need to think about

Bayes' Theorem for statistical inference

Likelihood

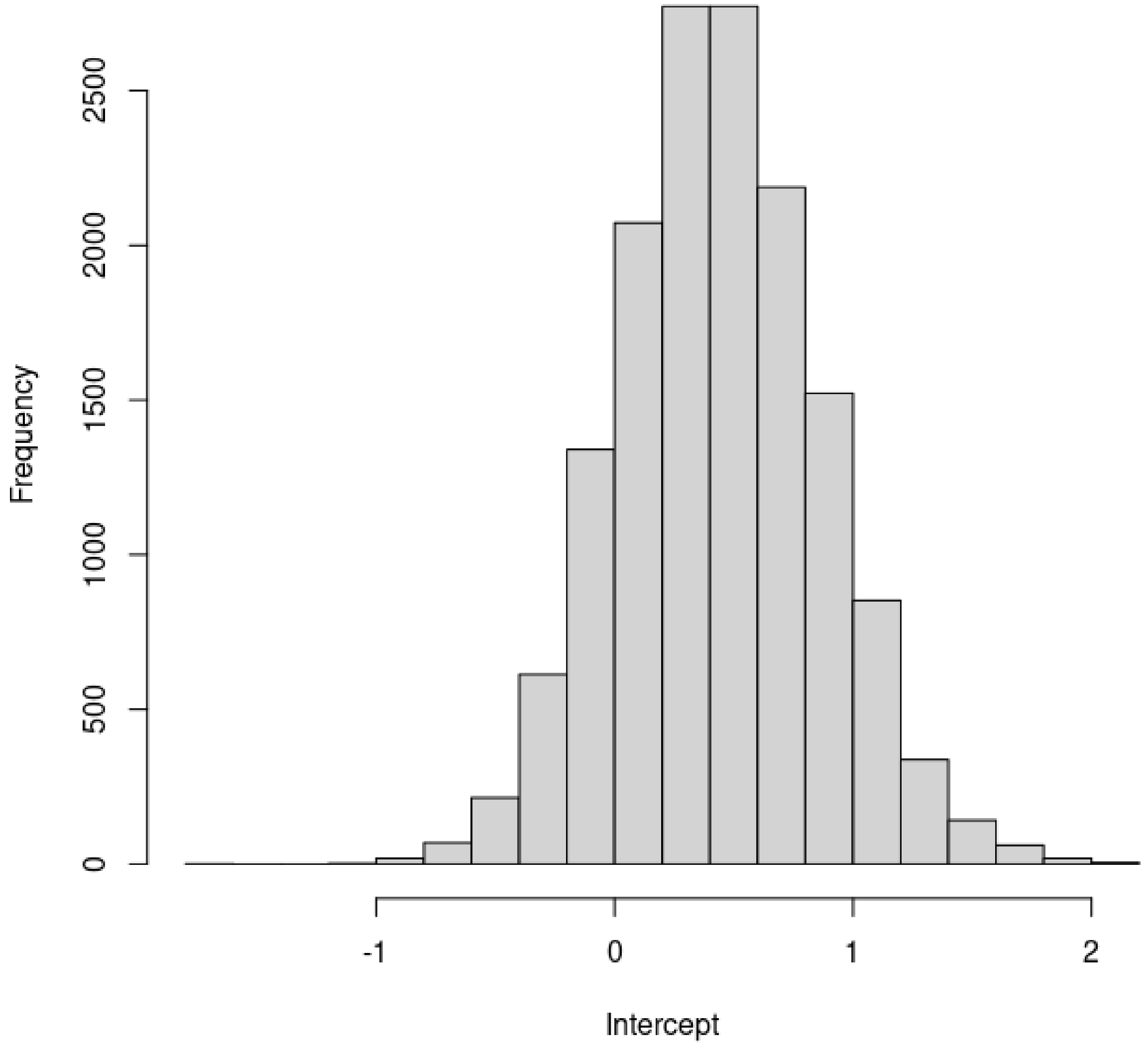
Prior distribution

$$p(\theta \mid y) \propto p(y \mid \theta) p(\theta)$$

Posterior distribution

The posterior distribution

- The posterior distribution is what we use to summarize results from a Bayesian analysis
- For each parameter, we get a distribution of possible values.
- We summarize that distribution in different ways (e.g., mean, median, mode)



Bayesian analysis: Summary of major points

- Probability as a measure of uncertainty for *all* unknown quantities

Bayesian analysis: Summary of major points

- Probability as a measure of uncertainty for *all* unknown quantities
- All statistical inference is a simple probability calculation

Bayesian analysis: Summary of major points

- Probability as a measure of uncertainty for *all* unknown quantities
- All statistical inference is a simple probability calculation
- Express prior information via $p(\theta)$

Bayesian analysis: Summary of major points

- Probability as a measure of uncertainty for *all* unknown quantities
- All statistical inference is a simple probability calculation
- Express prior information via $p(\theta)$
- Update your prior information using the likelihood, $p(y \mid \theta)$, to form the posterior $p(\theta \mid y)$

Bayesian analysis: Summary of major points

- Probability as a measure of uncertainty for *all* unknown quantities
- All statistical inference is a simple probability calculation
- Express prior information via $p(\theta)$
- Update your prior information using the likelihood, $p(y \mid \theta)$, to form the posterior $p(\theta \mid y)$
- **Final result is a distribution of possible values for the unknown parameters, unlike maximum likelihood estimation, where our result is a single value**

Benefits of the Bayesian approach

1. Natural concept of probability
   - What is the probability my detection probability is less than 0.5?

Benefits of the Bayesian approach

1. Natural concept of probability
   - What is the probability my detection probability is less than 0.5?
2. Clear link with how we (humans) learn
   - Our final result is the combination of experience (prior) and new information (likelihood)
   - New information changes, or updates, my previous state of knowledge to my current state of knowledge

Benefits of the Bayesian approach

1. Natural concept of probability
    - What is the probability my detection probability is less than 0.5?
2. Clear link with how we (humans) learn
    - Our final result is the combination of experience (prior) and new information (likelihood)
    - New information changes, or updates, my previous state of knowledge to my current state of knowledge
3. Much more flexible than frequentist approaches to account for complex data types
    - Multiple species
    - Multiple data sources
    - Lots of different random effect structures
    - Complex interactions across parameters

Drawbacks of the Bayesian approach

1. Results are always dependent on the prior distribution
   - Can specify weakly informative (vague) priors
   - Be transparent
   - Justify your choice

Drawbacks of the Bayesian approach

1. Results are always dependent on the prior distribution
   - Can specify weakly informative (vague) priors
   - Be transparent
   - Justify your choice
2. Typically more computationally intensive than frequentist approaches

Drawbacks of the Bayesian approach

1. Results are always dependent on the prior distribution
    - Can specify weakly informative (vague) priors
    - Be transparent
    - Justify your choice
2. Typically more computationally intensive than frequentist approaches
3. Often more coding involved.

Drawbacks of the Bayesian approach

1. Results are always dependent on the prior distribution
    - Can specify weakly informative (vague) priors
    - Be transparent
    - Justify your choice
2. Typically more computationally intensive than frequentist approaches
3. Often more coding involved. But...
    - Is this really a bad thing? `JAGS`, `NIMBLE`, `Stan` foster learning

Drawbacks of the Bayesian approach

1. Results are always dependent on the prior distribution
   - Can specify weakly informative (vague) priors
   - Be transparent
   - Justify your choice
2. Typically more computationally intensive than frequentist approaches
3. Often more coding involved. But…
   - Is this really a bad thing? `JAGS`, `NIMBLE`, `Stan` foster learning
   - LOTS of canned software in active development
     - Ex: `brms`, `spOccupancy`, `oSCR`, `marked`

Frequentist

Bayesian

Frequentist

Bayesian

$

Frequentist

Bayesian

$

$

Frequentist

Bayesian

$

$

$

Frequentist

$

$

$

$

Bayesian

Frequentist

$

$

$

$

Bayesian

$$$

Frequentist

$

$

$

$

Total = $$$$

Bayesian

$$$

Total = $$$

Questions?