Bayesian Data Analysis, class 1b

Andrew Gelman

Chapter 1: Probability and inference
“Bayesian inference” is too narrow; “Bayesian statistics” is too broad

“Bayes” is a good brand name; “Statistics using conditional probability” is confusing

Everyone uses Bayesian inference when it is appropriate. A Bayesian is a statistician who uses Bayesian inference even when it is inappropriate. I am a Bayesian.

First, second, and third editions
“Bayesian inference” is too narrow; “Bayesian statistics” is too broad

“Bayes” is a good brand name; “Statistics using conditional probability” is confusing

Everyone uses Bayesian inference when it is appropriate. A Bayesian is a statistician who uses Bayesian inference even when it is inappropriate. I am a Bayesian.

First, second, and third editions
“Bayesian inference” is too narrow; “Bayesian statistics” is too broad

“Bayes” is a good brand name; “Statistics using conditional probability” is confusing

Everyone uses Bayesian inference when it is appropriate. A Bayesian is a statistician who uses Bayesian inference even when it is inappropriate. I am a Bayesian.

First, second, and third editions
“Bayesian inference” is too narrow; “Bayesian statistics” is too broad

“Bayes” is a good brand name; “Statistics using conditional probability” is confusing

Everyone uses Bayesian inference when it is appropriate. A Bayesian is a statistician who uses Bayesian inference even when it is inappropriate. I am a Bayesian.

First, second, and third editions
“Bayesian inference” is too narrow; “Bayesian statistics” is too broad

“Bayes” is a good brand name; “Statistics using conditional probability” is confusing

Everyone uses Bayesian inference when it is appropriate. A Bayesian is a statistician who uses Bayesian inference even when it is inappropriate. I am a Bayesian.

First, second, and third editions
What is Bayes?

- Bayes is data + regularization
- Bayes is data + prior information
- Bayes is logical probabilistic reasoning
- Bayes is different things at different times
What is Bayes?

- Bayes is data + regularization
- Bayes is data + prior information
- Bayes is logical probabilistic reasoning
- Bayes is different things at different times
What is Bayes?

- Bayes is data + regularization
- Bayes is data + prior information
- Bayes is logical probabilistic reasoning
- Bayes is different things at different times
What is Bayes?

- Bayes is data + regularization
- Bayes is data + prior information
- Bayes is logical probabilistic reasoning
- Bayes is different things at different times
What is Bayes?

- Bayes is data + regularization
- Bayes is data + prior information
- Bayes is logical probabilistic reasoning
- Bayes is different things at different times
1. Probability and inference

- Different approaches to statistics:
  - Traditional likelihood
  - Pure nonparametric, robust
  - Full Bayes modeling

- "A chicken is an egg's way of making another egg"

Andrew Gelman  Bayesian Data Analysis, class 1b
1. Probability and inference

- Different approaches to statistics:
  - Traditional likelihood
  - Pure nonparametric, robust
  - Full Bayes modeling
  - “A chicken is an egg's way of making another egg”

Andrew Gelman
Bayesian Data Analysis, class 1b
1. Probability and inference

- Different approaches to statistics:
  - Traditional likelihood
    - Pure nonparametric, robust
    - Full Bayes modeling
  - “A chicken is an egg’s way of making another egg”
1. Probability and inference

- Different approaches to statistics:
  - Traditional likelihood
  - Pure nonparametric, robust
  - Full Bayes modeling

- “A chicken is an egg’s way of making another egg”

Andrew Gelman
Bayesian Data Analysis, class 1b
1. Probability and inference

- Different approaches to statistics:
  - Traditional likelihood
  - Pure nonparametric, robust
  - Full Bayes modeling

- “A chicken is an egg’s way of making another egg”
1. Probability and inference

- Different approaches to statistics:
  - Traditional likelihood
  - Pure nonparametric, robust
  - Full Bayes modeling

- “A chicken is an egg’s way of making another egg”
1.1. The three steps of Bayesian data analysis

Three steps:
1. Setting up a probability model
2. Inference
3. Model checking

Then go back and improve the model
1.1. The three steps of Bayesian data analysis

Three steps:
1. Setting up a probability model
2. Inference
3. Model checking

Then go back and improve the model
1.1. The three steps of Bayesian data analysis

- Three steps:
  1. Setting up a probability model
  2. Inference
  3. Model checking
- Then go back and improve the model
1.1. The three steps of Bayesian data analysis

- Three steps:
  1. Setting up a probability model
  2. Inference
  3. Model checking

- Then go back and improve the model
1.1. The three steps of Bayesian data analysis

Three steps:
1. Setting up a probability model
2. Inference
3. Model checking

Then go back and improve the model
1.1. The three steps of Bayesian data analysis

- Three steps:
  1. Setting up a probability model
  2. Inference
  3. Model checking

- Then go back and improve the model
1.2. General notation for statistical inference

- x and y
- qoi’s
- Rubin philosophy: all statistics is inference about missing data
- In a world of predictions, what is the role of parameters?
- Rubin’s two questions:
1.2. General notation for statistical inference

- $x$ and $y$
- qoi’s
- Rubin philosophy: all statistics is inference about missing data
- In a world of predictions, what is the role of parameters?
- Rubin’s two questions:
1.2. General notation for statistical inference

- $x$ and $y$
- qoi’s

- Rubin philosophy: all statistics is inference about missing data
- In a world of predictions, what is the role of parameters?
- Rubin’s two questions:
1.2. General notation for statistical inference

- $x$ and $y$
- qoi’s
- Rubin philosophy: all statistics is inference about missing data
- In a world of predictions, what is the role of parameters?
- Rubin’s two questions:
  - What would you do if you had all the data?
  - What were you doing before you had any data?
1.2. General notation for statistical inference

- $x$ and $y$
- qoi's
- Rubin philosophy: all statistics is inference about missing data
- In a world of predictions, what is the role of parameters?
- Rubin's two questions:
  - What would you do if you had all the data?
  - What were you doing before you had any data?
1.2. General notation for statistical inference

- $x$ and $y$
- qoi’s
- Rubin philosophy: all statistics is inference about missing data
- In a world of predictions, what is the role of parameters?
- Rubin’s two questions:
  - What would you do if you had all the data?
  - What were you doing before you had any data?
1.2. General notation for statistical inference

- $x$ and $y$
- qoi’s
- Rubin philosophy: all statistics is inference about missing data
- In a world of predictions, what is the role of parameters?
- Rubin’s two questions:
  - What would you do if you had all the data?
  - What were you doing before you had any data?
1.2. General notation for statistical inference

- $x$ and $y$
- qoi’s
- Rubin philosophy: all statistics is inference about missing data
- In a world of predictions, what is the role of parameters?
- Rubin’s two questions:
  - What would you do if you had all the data?
  - What were you doing before you had any data?
1.3. Bayesian inference

- $p()$ and $\Pr()$
- $N(\theta|\mu, \sigma^2)$, etc., are precise mathematical expressions
- Details of distributions in Appendix A
- Check out our clean notation (compare to other books)
- Continuous random variables, conditioning on events of zero probability
- $p(\theta|y) \propto p(\theta)p(y|\theta)$ and the likelihood principle
1.3. Bayesian inference

- $p()$ and Pr()
- $N(\theta|\mu, \sigma^2)$, etc., are precise mathematical expressions
- Details of distributions in Appendix A
- Check out our clean notation (compare to other books)
- Continuous random variables, conditioning on events of zero probability
- $p(\theta|y) \propto p(\theta)p(y|\theta)$ and the likelihood principle
1.3. Bayesian inference

- $p()$ and $\text{Pr}()$
- $N(\theta | \mu, \sigma^2)$, etc., are precise mathematical expressions
- Details of distributions in Appendix A
- Check out our clean notation (compare to other books)
- Continuous random variables, conditioning on events of zero probability
- $p(\theta | y) \propto p(\theta)p(y | \theta)$ and the likelihood principle
1.3. Bayesian inference

- $p()$ and $Pr()$
- $N(\theta|\mu, \sigma^2)$, etc., are precise mathematical expressions
- Details of distributions in Appendix A
- Check out our clean notation (compare to other books)
- Continuous random variables, conditioning on events of zero probability
- $p(\theta|y) \propto p(\theta)p(y|\theta)$ and the likelihood principle
1.3. Bayesian inference

- $p()$ and $\text{Pr}()$
- $N(\theta|\mu, \sigma^2)$, etc., are precise mathematical expressions
- Details of distributions in Appendix A
- Check out our clean notation (compare to other books)
- Continuous random variables, conditioning on events of zero probability
- $p(\theta|y) \propto p(\theta)p(y|\theta)$ and the likelihood principle
1.3. Bayesian inference

- $p()$ and $\text{Pr}()$
- $\text{N}(\theta|\mu, \sigma^2)$, etc., are precise mathematical expressions
- Details of distributions in Appendix A
- Check out our clean notation (compare to other books)
- Continuous random variables, conditioning on events of zero probability
- $p(\theta|y) \propto p(\theta)p(y|\theta)$ and the likelihood principle
1.3. Bayesian inference

- $p()$ and $\Pr()$
- $\mathcal{N}(\theta|\mu, \sigma^2)$, etc., are precise mathematical expressions
- Details of distributions in Appendix A
- Check out our clean notation (compare to other books)
- Continuous random variables, conditioning on events of zero probability
- $p(\theta|y) \propto p(\theta)p(y|\theta)$ and the likelihood principle
1.4. Discrete probability examples: genetics and spell checking

- The typed word “Radom” is actually Random ($\theta = 1$), Radon ($\theta = 2$), or Radom ($\theta = 3$)

- Prior distribution:

<table>
<thead>
<tr>
<th>$\theta$</th>
<th>$p(\theta)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>random</td>
<td>$7.60 \times 10^{-5}$</td>
</tr>
<tr>
<td>radon</td>
<td>$6.05 \times 10^{-6}$</td>
</tr>
<tr>
<td>radom</td>
<td>$3.12 \times 10^{-7}$</td>
</tr>
</tbody>
</table>

- Likelihood:

| $\theta$ | $p(\text{“radom”} | \theta)$ |
|----------|----------------|
| random   | 0.00193         |
| radon    | 0.000143        |
| radom    | 0.975           |

Andrew Gelman  
Bayesian Data Analysis, class 1b
1.4. Discrete probability examples: genetics and spell checking

- The typed word “Radom” is actually Random ($\theta = 1$), Radon ($\theta = 2$), or Radom ($\theta = 3$)

- Prior distribution:

<table>
<thead>
<tr>
<th>$\theta$</th>
<th>$p(\theta)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>random</td>
<td>$7.60 \times 10^{-5}$</td>
</tr>
<tr>
<td>radon</td>
<td>$6.05 \times 10^{-6}$</td>
</tr>
<tr>
<td>radom</td>
<td>$3.12 \times 10^{-7}$</td>
</tr>
</tbody>
</table>

- Likelihood:

| $\theta$   | $p(\text{“radom”}|\theta)$ |
|------------|----------------------------|
| random     | 0.00193                    |
| radon      | 0.000143                   |
| radom      | 0.975                      |
1.4. Discrete probability examples: genetics and spell checking

- The typed word “Radom” is actually Random ($\theta = 1$), Radon ($\theta = 2$), or Radom ($\theta = 3$)

- Prior distribution:

<table>
<thead>
<tr>
<th>$\theta$</th>
<th>$p(\theta)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>random</td>
<td>$7.60 \times 10^{-5}$</td>
</tr>
<tr>
<td>radon</td>
<td>$6.05 \times 10^{-6}$</td>
</tr>
<tr>
<td>radom</td>
<td>$3.12 \times 10^{-7}$</td>
</tr>
</tbody>
</table>

- Likelihood:

<table>
<thead>
<tr>
<th>$\theta$</th>
<th>$p(\text{“radom”} \mid \theta)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>random</td>
<td>0.00193</td>
</tr>
<tr>
<td>radon</td>
<td>0.000143</td>
</tr>
<tr>
<td>radom</td>
<td>0.975</td>
</tr>
</tbody>
</table>
1.4. Discrete probability examples: genetics and spell checking

- The typed word “Radom” is actually Random ($\theta = 1$), Radon ($\theta = 2$), or Radom ($\theta = 3$)

- Prior distribution:

<table>
<thead>
<tr>
<th>$\theta$</th>
<th>$p(\theta)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>random</td>
<td>$7.60 \times 10^{-5}$</td>
</tr>
<tr>
<td>radon</td>
<td>$6.05 \times 10^{-6}$</td>
</tr>
<tr>
<td>radom</td>
<td>$3.12 \times 10^{-7}$</td>
</tr>
</tbody>
</table>

- Likelihood:

| $\theta$   | $p(\text{“radom”} | \theta)$ |
|------------|-------------|
| random     | 0.00193     |
| radon      | 0.000143    |
| radom      | 0.975       |
Prior, likelihood, posterior distributions:

| θ   | p(θ)   | p(y|θ) | p(θ)p(y|θ) | p(θ|y) |
|-----|--------|--------|------------|--------|
| random | 7.60 × 10^{-5} | 0.00193 | 1.47 × 10^{-7} | 0.325 |
| radon  | 6.05 × 10^{-6} | 0.000143 | 8.65 × 10^{-10} | 0.002 |
| radom  | 3.12 × 10^{-7} | 0.975   | 3.04 × 10^{-7}  | 0.673 |

Decision making

Model checking

Model improvement
Prior, likelihood, posterior distributions:

| θ     | p(θ)     | p(y|θ)   | p(θ)p(y|θ)   | p(θ|y)  |
|-------|----------|---------|--------------|---------|
| random| 7.60 × 10^{-5} | 0.00193 | 1.47 × 10^{-7} | 0.325   |
| radon | 6.05 × 10^{-6} | 0.000143 | 8.65 × 10^{-10} | 0.002   |
| radom | 3.12 × 10^{-7} | 0.975   | 3.04 × 10^{-7}  | 0.673   |

Decision making

Model checking

Model improvement
Prior, likelihood, posterior distributions:

| $\theta$  | $p(\theta)$ | $p(y|\theta)$ | $p(\theta)p(y|\theta)$ | $p(\theta|y)$ |
|-----------|-------------|----------------|------------------------|--------------|
| random    | $7.60 \times 10^{-5}$ | 0.00193         | $1.47 \times 10^{-7}$  | 0.325        |
| radon     | $6.05 \times 10^{-6}$ | 0.000143        | $8.65 \times 10^{-10}$ | 0.002        |
| radom     | $3.12 \times 10^{-7}$ | 0.975           | $3.04 \times 10^{-7}$  | 0.673        |

Decision making

- Model checking
- Model improvement
Prior, likelihood, posterior distributions:

| θ     | p(θ)    | p(y|θ)  | p(θ)p(y|θ) | p(θ|y)  |
|-------|---------|--------|------------|---------|
| random| $7.60 \times 10^{-5}$ | 0.00193 | $1.47 \times 10^{-7}$ | 0.325 |
| radon | $6.05 \times 10^{-6}$ | 0.000143 | $8.65 \times 10^{-10}$ | 0.002 |
| radom | $3.12 \times 10^{-7}$ | 0.975   | $3.04 \times 10^{-7}$ | 0.673 |

Decision making

Model checking

Model improvement
Prior, likelihood, posterior distributions:

| $\theta$     | $p(\theta)$  | $p(y|\theta)$ | $p(\theta)p(y|\theta)$ | $p(\theta|y)$ |
|--------------|---------------|----------------|--------------------------|---------------|
| random       | $7.60 \times 10^{-5}$ | 0.00193        | $1.47 \times 10^{-7}$    | 0.325         |
| radon        | $6.05 \times 10^{-6}$ | 0.000143       | $8.65 \times 10^{-10}$   | 0.002         |
| radom        | $3.12 \times 10^{-7}$ | 0.975          | $3.04 \times 10^{-7}$    | 0.673         |

Decision making

Model checking

Model improvement
1.5. Probability as a measure of uncertainty

- Foundations of probability
- Equally likely events
- Calibration on events defined by physical symmetry
- “Suppose a coin having probability 0.7 of coming up heads is tossed”
- I’m unsatisfied by axiomatic or betting rationales for Bayes
- Frequency reference sets = Bayes probability
1.5. Probability as a measure of uncertainty

- Foundations of probability
  - Equally likely events
  - Calibration on events defined by physical symmetry
  - “Suppose a coin having probability 0.7 of coming up heads is tossed”
  - I’m unsatisfied by axiomatic or betting rationales for Bayes
  - Frequency reference sets = Bayes probability
1.5. Probability as a measure of uncertainty

- Foundations of probability
- Equally likely events
- Calibration on events defined by physical symmetry
- “Suppose a coin having probability 0.7 of coming up heads is tossed”
- I’m unsatisfied by axiomatic or betting rationales for Bayes
- Frequency reference sets = Bayes probability
1.5. Probability as a measure of uncertainty

- Foundations of probability
- Equally likely events
- Calibration on events defined by physical symmetry
  - “Suppose a coin having probability 0.7 of coming up heads is tossed”
- I’m unsatisfied by axiomatic or betting rationales for Bayes
- Frequency reference sets = Bayes probability
1.5. Probability as a measure of uncertainty

- Foundations of probability
- Equally likely events
- Calibration on events defined by physical symmetry
  - “Suppose a coin having probability 0.7 of coming up heads is tossed”
- I’m unsatisfied by axiomatic or betting rationales for Bayes
- Frequency reference sets = Bayes probability
1.5. Probability as a measure of uncertainty

- Foundations of probability
- Equally likely events
- Calibration on events defined by physical symmetry
- “Suppose a coin having probability 0.7 of coming up heads is tossed”
- I’m unsatisfied by axiomatic or betting rationales for Bayes
- Frequency reference sets = Bayes probability
1.5. Probability as a measure of uncertainty

- Foundations of probability
- Equally likely events
- Calibration on events defined by physical symmetry
- “Suppose a coin having probability 0.7 of coming up heads is tossed”
- I’m unsatisfied by axiomatic or betting rationales for Bayes
- Frequency reference sets = Bayes probability
1.6. Example of probability assignment: football point spreads
Raw data plus assumptions
Estimating the probability a vote is decisive

States where your vote is most likely to matter
The numbers

Probability that your state is pivotal and that it is tied

Pr (your vote matters): 1 in 10 million
Pr (your vote matters): 1 in a billion
Pr (your vote matters): 1 in 100 billion
Raw data plus assumptions

- Estimating the probability a vote is decisive:
  - Pure data
  - Pure model
  - Combining data and model

- There is never a pure “pure data” estimate; we still need a reference set (that is, exchangeability)
Raw data plus assumptions

- Estimating the probability a vote is decisive:
  - Pure data
  - Pure model
  - Combining data and model

- There is never a pure “pure data” estimate; we still need a reference set (that is, exchangeability)
Raw data plus assumptions

- Estimating the probability a vote is decisive:
  - Pure data
  - Pure model
  - Combining data and model

- There is never a pure “pure data” estimate; we still need a reference set (that is, exchangeability)
Estimating the probability a vote is decisive:

- Pure data
- Pure model
- Combining data and model

There is never a pure “pure data” estimate; we still need a reference set (that is, exchangeability)
Raw data plus assumptions

- Estimating the probability a vote is decisive:
  - Pure data
  - Pure model
  - Combining data and model

- There is never a pure “pure data” estimate; we still need a reference set (that is, exchangeability)
Raw data plus assumptions

- Estimating the probability a vote is decisive:
  - Pure data
  - Pure model
  - Combining data and model

- There is never a pure “pure data” estimate; we still need a reference set (that is, exchangeability)
1.7. Example: estimating the accuracy of record linkage

Another example of empirical probability assignment
Another example of empirical probability assignment
1.8. Some useful results from probability theory

- The math you need: derivatives, integrals, multivariable calculus
- Being able to read an expression and separate constants from variables:

\[
\prod_{i=1}^{n} N(y_i | \mu_i, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp \left( -\frac{1}{2\sigma^2} (y_i - \mu_i)^2 \right)
\]

- Late-twentieth-century probability modeling
1.8. Some useful results from probability theory

- The math you need: derivatives, integrals, multivariable calculus
- Being able to read an expression and separate constants from variables:

\[ \prod_{i=1}^{n} \mathcal{N}(y_i|\mu_i, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp \left( -\frac{1}{2\sigma^2}(y_i - \mu_i)^2 \right) \]

- Late-twentieth-century probability modeling
1.8. Some useful results from probability theory

- The math you need: derivatives, integrals, multivariable calculus
- Being able to read an expression and separate constants from variables:

\[
\prod_{i=1}^{n} N(y_i | \mu_i, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(y_i - \mu_i)^2\right)
\]

- Late-twentieth-century probability modeling
1.8. Some useful results from probability theory

- The math you need: derivatives, integrals, multivariable calculus
- Being able to read an expression and separate constants from variables:
  \[
  \prod_{i=1}^{n} N(y_i | \mu_i, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp \left( -\frac{1}{2\sigma^2} (y_i - \mu_i)^2 \right)
  \]
- Late-twentieth-century probability modeling
Figure 1. Histogram of Democratic Share of the Two-Party Vote in Congressional Elections in 1988. Only districts that were contested by both major parties are shown here.
Figure 2. Histogram of Democratic Share of the Two-Party Vote in Congressional Elections in 1988, in Districts With (a) Republican Incumbents, (b) Democratic Incumbents, and (c) Open Seats. Combined, the three distributions yield the bimodal distribution in Figure 1.
Models of congressional elections

The early-twentieth-century method of modeling: find a distributional family

The late-twentieth-century approach: modeling using conditional distributions

The twenty-first-century approach: hierarchical nonparametric modeling?
Models of congressional elections

The early-twentieth-century method of modeling: find a distributional family

The late-twentieth-century approach: modeling using conditional distributions

The twenty-first-century approach: hierarchical nonparametric modeling?
Models of congressional elections

The early-twentieth-century method of modeling: find a distributional family

The late-twentieth-century approach: modeling using conditional distributions

The twenty-first-century approach: hierarchical nonparametric modeling?
Models of congressional elections

The early-twentieth-century method of modeling: find a distributional family

The late-twentieth-century approach: modeling using conditional distributions

The twenty-first-century approach: hierarchical nonparametric modeling?
Models of congressional elections

The early-twentieth-century method of modeling: find a distributional family

The late-twentieth-century approach: modeling using conditional distributions

The twenty-first-century approach: hierarchical nonparametric modeling?
“If you wanted to do foundational research in statistics in the mid-twentieth century, you had to be a bit of a mathematician, whether you wanted to or not. If you want to do statistical research at the turn of the twenty-first century, you have to be a computer programmer.”

- Programming
- Graphics
- Your working environment
- Problem-solving skills
“If you wanted to do foundational research in statistics in the mid-twentieth century, you had to be bit of a mathematician, whether you wanted to or not . . . If you want to do statistical research at the turn of the twenty-first century, you have to be a computer programmer.”

- Programming
- Graphics
- Your working environment
- Problem-solving skills
“If you wanted to do foundational research in statistics in the mid-twentieth century, you had to be a bit of a mathematician, whether you wanted to or not. . . If you want to do statistical research at the turn of the twenty-first century, you have to be a computer programmer.”

- Programming
  - Graphics
  - Your working environment
  - Problem-solving skills
“If you wanted to do foundational research in statistics in the mid-twentieth century, you had to be a bit of a mathematician, whether you wanted to or not . . . If you want to do statistical research at the turn of the twenty-first century, you have to be a computer programmer.”

- Programming
- Graphics
- Your working environment
- Problem-solving skills
“If you wanted to do foundational research in statistics in the mid-twentieth century, you had to be a bit of a mathematician, whether you wanted to or not . . . If you want to do statistical research at the turn of the twenty-first century, you have to be a computer programmer.”

- Programming
- Graphics
- Your working environment
- Problem-solving skills
“If you wanted to do foundational research in statistics in the mid-twentieth century, you had to be bit of a mathematician, whether you wanted to or not. . . If you want to do statistical research at the turn of the twenty-first century, you have to be a computer programmer.”

- Programming
- Graphics
- Your working environment
- Problem-solving skills
### Summarizing inferences by simulation

<table>
<thead>
<tr>
<th>Simulation draw</th>
<th>Parameters</th>
<th>Predictive quantities</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\theta_1$</td>
<td>$\ldots$</td>
</tr>
<tr>
<td>1</td>
<td>$\theta_1^1$</td>
<td>$\ldots$</td>
</tr>
<tr>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\ddots$</td>
</tr>
<tr>
<td>L</td>
<td>$\theta_1^S$</td>
<td>$\ldots$</td>
</tr>
</tbody>
</table>
Example: a regression model for forecasting elections

<table>
<thead>
<tr>
<th>sim</th>
<th>σ</th>
<th>β₀</th>
<th>β₁</th>
<th>β₂</th>
<th>ŷ₁</th>
<th>ŷ₂</th>
<th>...</th>
<th>ŷ₅₅</th>
<th>...</th>
<th>ŷ₄₃₅</th>
<th>∑ᵢ I(ŷᵢ &gt; 0.5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.065</td>
<td>.19</td>
<td>.62</td>
<td>.067</td>
<td>.69</td>
<td>.57</td>
<td>...</td>
<td>NA</td>
<td>...</td>
<td>.79</td>
<td>251</td>
</tr>
<tr>
<td>2</td>
<td>.069</td>
<td>.25</td>
<td>.50</td>
<td>.097</td>
<td>.75</td>
<td>.63</td>
<td>...</td>
<td>NA</td>
<td>...</td>
<td>.76</td>
<td>254</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td>.067</td>
<td>.23</td>
<td>.51</td>
<td>.089</td>
<td>.73</td>
<td>.57</td>
<td>...</td>
<td>NA</td>
<td>...</td>
<td>.69</td>
<td>251</td>
</tr>
<tr>
<td>median</td>
<td>.068</td>
<td>.20</td>
<td>.58</td>
<td>.077</td>
<td>.73</td>
<td>.65</td>
<td>...</td>
<td>NA</td>
<td>...</td>
<td>.72</td>
<td>253</td>
</tr>
<tr>
<td>mean</td>
<td>.067</td>
<td>.20</td>
<td>.58</td>
<td>.078</td>
<td>.73</td>
<td>.65</td>
<td>...</td>
<td>NA</td>
<td>...</td>
<td>.72</td>
<td>252.4</td>
</tr>
<tr>
<td>sd</td>
<td>.003</td>
<td>.02</td>
<td>.04</td>
<td>.007</td>
<td>.07</td>
<td>.07</td>
<td>...</td>
<td>NA</td>
<td>...</td>
<td>.07</td>
<td>3.1</td>
</tr>
</tbody>
</table>
Computing languages and software

- Bayesian
  - Stan
  - Bugs, Jags
  - Others
- Statistical/mathematical
- General
Computing languages and software

- Bayesian
  - Stan
  - Bugs, Jags
  - Others
- Statistical/mathematical

- General
Computing languages and software

- Bayesian
  - Stan
    - Bugs, Jags
    - Others
  - Statistical/mathematical

- General
Computing languages and software

- Bayesian
  - Stan
  - Bugs, Jags
  - Others

- Statistical/mathematical
  - R (see Chapters 7 and 8 of ARM)
  - Matlab
  - Stata
  - Others

- General
Computing languages and software

- Bayesian
  - Stan
  - Bugs, Jags
  - Others

- Statistical/mathematical
  - R (see Chapters 7 and 8 of ARM)
  - Matlab
  - Stata
  - Others

- General
Computing languages and software

- Bayesian
  - Stan
  - Bugs, Jags
  - Others
- Statistical/mathematical
  - R (see Chapters 7 and 8 of ARM)
  - Matlab
  - Stata
  - Others
- General
Computing languages and software

- Bayesian
  - Stan
  - Bugs, Jags
  - Others

- Statistical/mathematical
  - R (see Chapters 7 and 8 of ARM)
  - Matlab
  - Stata
  - Others

- General
Computing languages and software

- Bayesian
  - Stan
  - Bugs, Jags
  - Others
- Statistical/mathematical
  - R (see Chapters 7 and 8 of ARM)
  - Matlab
  - Stata
  - Others
- General
Computing languages and software

- Bayesian
  - Stan
  - Bugs, Jags
  - Others
- Statistical/mathematical
  - R (see Chapters 7 and 8 of ARM)
  - Matlab
  - Stata
  - Others
- General
  - C
  - Linpack, Eigen, etc
Computing languages and software

- Bayesian
  - Stan
  - Bugs, Jags
  - Others

- Statistical/mathematical
  - R (see Chapters 7 and 8 of ARM)
  - Matlab
  - Stata
  - Others

- General
  - C
  - Linpack, Eigen, etc
Computing languages and software

- Bayesian
  - Stan
  - Bugs, Jags
  - Others
- Statistical/mathematical
  - R (see Chapters 7 and 8 of ARM)
  - Matlab
  - Stata
  - Others
- General
  - C
  - Linpack, Eigen, etc
Computing languages and software

- **Bayesian**
  - Stan
  - Bugs, Jags
  - Others

- **Statistical/mathematical**
  - R (see Chapters 7 and 8 of ARM)
  - Matlab
  - Stata
  - Others

- **General**
  - C
    - Linpack, Eigen, etc
Computing languages and software

- Bayesian
  - Stan
  - Bugs, Jags
  - Others

- Statistical/mathematical
  - R (see Chapters 7 and 8 of ARM)
  - Matlab
  - Stata
  - Others

- General
  - C
  - Linpack, Eigen, etc
Simple models that can be fit successfully

Complex models that cannot be fit, or that give nonsensical results
1.10. Bayesian inference in applied statistics

- Flexibility
- Combine multiple sources of information
- Uncertainty and variation
1.10. Bayesian inference in applied statistics

- Flexibility
  - Combine multiple sources of information
  - Uncertainty and variation
1.10. Bayesian inference in applied statistics

- Flexibility
- Combine multiple sources of information
- Uncertainty and variation
1.10. Bayesian inference in applied statistics

- Flexibility
- Combine multiple sources of information
- Uncertainty and variation
Summary of Chapter 1

- 3 steps of Bayesian data analysis
- Bayesian inference for simple discrete probabilities
- Assigning probabilities from data
- Simulation and software
Summary of Chapter 1

- 3 steps of Bayesian data analysis
  - Bayesian inference for simple discrete probabilities
  - Assigning probabilities from data
  - Simulation and software
Summary of Chapter 1

- 3 steps of Bayesian data analysis
- Bayesian inference for simple discrete probabilities
  - Assigning probabilities from data
  - Simulation and software
Summary of Chapter 1

- 3 steps of Bayesian data analysis
- Bayesian inference for simple discrete probabilities
- Assigning probabilities from data
- Simulation and software
Summary of Chapter 1

- 3 steps of Bayesian data analysis
- Bayesian inference for simple discrete probabilities
- Assigning probabilities from data
- Simulation and software
Homework due beginning of class 2b

All assignments are at http://www.stat.columbia.edu/~gelman/bda.course/homeworks.pdf

- Theory problem: prior distributions for the exponential distribution
- Computing problem: arranging national poll data by state, Bayesian inference estimating prior distribution using method of moments
- Applied problem: noninformative, subjective, and weakly informative prior distributions
Homework due beginning of class 2b

- All assignments are at http://www.stat.columbia.edu/~gelman/bda.course/homeworks.pdf
  - Theory problem: prior distributions for the exponential distribution
  - Computing problem: arranging national poll data by state, Bayesian inference estimating prior distribution using method of moments
  - Applied problem: noninformative, subjective, and weakly informative prior distributions
All assignments are at http://www.stat.columbia.edu/~gelman/bda.course/homeworks.pdf

- Theory problem: prior distributions for the exponential distribution
- Computing problem: arranging national poll data by state, Bayesian inference estimating prior distribution using method of moments
- Applied problem: noninformative, subjective, and weakly informative prior distributions
Homework due beginning of class 2b

- All assignments are at http://www.stat.columbia.edu/~gelman/bda.course/homeworks.pdf
  - Theory problem: prior distributions for the exponential distribution
  - Computing problem: arranging national poll data by state, Bayesian inference estimating prior distribution using method of moments
  - Applied problem: noninformative, subjective, and weakly informative prior distributions
All assignments are at http://www.stat.columbia.edu/~gelman/bda.course/homeworks.pdf

- Theory problem: prior distributions for the exponential distribution
- Computing problem: arranging national poll data by state, Bayesian inference estimating prior distribution using method of moments
- Applied problem: noninformative, subjective, and weakly informative prior distributions