



UPPSALA
UNIVERSITET

- Introduction
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

Bayesian Statistics and Data Analysis

Lecture 1

Måns Magnusson

Department of Statistics, Uppsala University
Thanks to Aki Vehtari, Aalto University



UPPSALA
UNIVERSITET

- **Introduction**
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

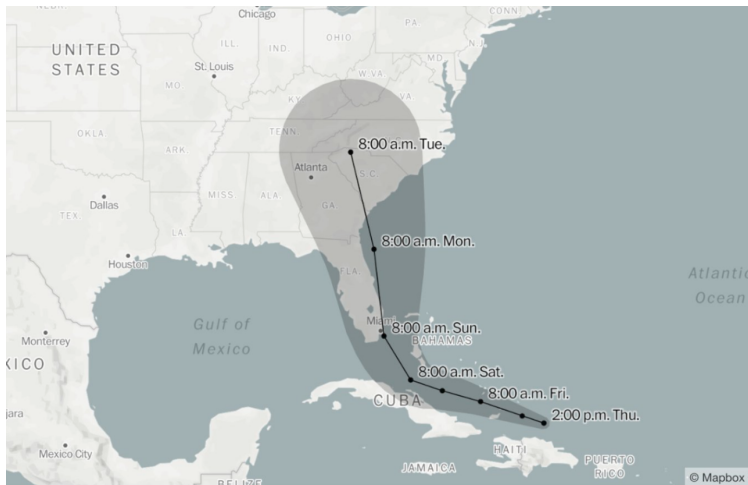
Section 1

Introduction



- Introduction
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

Decision making in case of uncertainties





UPPSALA
UNIVERSITET

Bayesian Analysis

- **Introduction**
 - Bayesian Reasoning
 - Bayesian Statistical inference
 - Probabilistic Modeling
 - Bayesian Computation
- Bayesian probability theory
 - uncertainty is presented with probabilities
 - probabilities are updated based on new information



UPPSALA
UNIVERSITET

Bayesian Analysis

- **Introduction**

- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

- Bayesian probability theory
 - uncertainty is presented with probabilities
 - probabilities are updated based on new information
- Thomas Bayes (170?–1761)
 - English nonconformist, Presbyterian minister, mathematician
 - considered the problem of *inverse probability*
 - significant part of the Bayesian theory



- **Introduction**

- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

- Bayesian probability theory
 - uncertainty is presented with probabilities
 - probabilities are updated based on new information
- Thomas Bayes (170?–1761)
 - English nonconformist, Presbyterian minister, mathematician
 - considered the problem of *inverse probability*
 - significant part of the Bayesian theory
- Bayes did not invent all, but was first to solve problem of inverse probability in special case
- Modern Bayesian theory with rigorous proofs developed in 20th century



Term Bayesian used first time in mid 20th century

- Earlier there was just "probability theory"
 - concept of the probability was not strictly defined, although it was close to modern Bayesian interpretation
 - in the end of 19th century there were increasing demand for more strict definition of probability (mathematical and philosophical problem)

- **Introduction**

- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation



Term Bayesian used first time in mid 20th century

- Earlier there was just "probability theory"
 - concept of the probability was not strictly defined, although it was close to modern Bayesian interpretation
 - in the end of 19th century there were increasing demand for more strict definition of probability (mathematical and philosophical problem)
- In the beginning of 20th century frequentist view gained popularity
 - accepts definition of probabilities only through frequencies
 - does not accept inverse probability or use of prior
 - gained popularity due to apparent objectivity and "cook book" like reference books



Term Bayesian used first time in mid 20th century

- Earlier there was just "probability theory"
 - concept of the probability was not strictly defined, although it was close to modern Bayesian interpretation
 - in the end of 19th century there were increasing demand for more strict definition of probability (mathematical and philosophical problem)
- In the beginning of 20th century frequentist view gained popularity
 - accepts definition of probabilities only through frequencies
 - does not accept inverse probability or use of prior
 - gained popularity due to apparent objectivity and "cook book" like reference books
- R. A. Fisher used in 1950 first time term "Bayesian" to emphasize the difference to general term "probability theory"
 - term became quickly popular, because alternative descriptions were longer



UPPSALA
UNIVERSITET

Bayesian core: Uncertainty and probabilistic modeling

- **Introduction**

- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

- Representing uncertainty with probabilities
- Updating uncertainty (with empirical observations)



UPPSALA
UNIVERSITET

Two types of uncertainty

- Introduction

- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

- Aleatoric uncertainty due to randomness
 - we are not able to obtain observations which could reduce this uncertainty
- Epistemic uncertainty due to lack of knowledge



UPPSALA
UNIVERSITET

Two types of uncertainty

- Introduction

- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

- Aleatoric uncertainty due to randomness
 - we are not able to obtain observations which could reduce this uncertainty
- Epistemic uncertainty due to lack of knowledge
 - we are able to obtain observations which can reduce this uncertainty



Two types of uncertainty

- Introduction

- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

- Aleatoric uncertainty due to randomness
 - we are not able to obtain observations which could reduce this uncertainty
- Epistemic uncertainty due to lack of knowledge
 - we are able to obtain observations which can reduce this uncertainty
 - two observers may have different epistemic uncertainty



UPPSALA
UNIVERSITET

- Introduction
- **Bayesian Reasoning**
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

Section 2

Bayesian Reasoning



UPPSALA
UNIVERSITET

Probability in Bayesian Reasoning

- Uncertainty and knowledge is expressed as probability distributions, why?

- Introduction
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation



- Introduction
- **Bayesian Reasoning**
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

Probability in Bayesian Reasoning

- Uncertainty and knowledge is expressed as probability distributions, why?
- Analogy to physical randomness
- Knowledge as coherence of bets (the Dutch book)
Example: If you say: $P(\text{rain}) = 0.7$ and $P(\text{no rain}) = 0.4$, those add up to 1.1 (too high).
A bookmaker could then take both bets from you and be sure to profit, because you've made your odds **incoherent**.



- Introduction
- **Bayesian Reasoning**
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

Probability in Bayesian Reasoning

- Uncertainty and knowledge is expressed as probability distributions, why?
- Analogy to physical randomness
- Knowledge as coherence of bets (the Dutch book)
Example: If you say: $P(\text{rain}) = 0.7$ and $P(\text{no rain}) = 0.4$, those add up to 1.1 (too high).
A bookmaker could then take both bets from you and be sure to profit, because you've made your odds **incoherent**.
- Everyone can have their own 'subjective' uncertainty, e.g. $P(\text{rain tomorrow})$, $P(\text{Magdalena Andersson will be the next primeminister})$
- Frequency arguments can be difficult in some situations: $P(\text{other life in the universe})$



UPPSALA
UNIVERSITET

Bayesian epistemology

- Introduction
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation



UPPSALA
UNIVERSITET

Bayesian epistemology

- Introduction
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

- What is Bayesian epistemology?



UPPSALA
UNIVERSITET

Bayesian epistemology

- Introduction
 - Bayesian Reasoning
 - Bayesian Statistical inference
 - Probabilistic Modeling
 - Bayesian Computation
- What is **epistemology**?
"the theory of knowledge" (Wikipedia)



UPPSALA
UNIVERSITET

Bayesian epistemology

- Introduction
 - Bayesian Reasoning
 - Bayesian Statistical inference
 - Probabilistic Modeling
 - Bayesian Computation
- What is **epistemology**?
"the theory of knowledge" (Wikipedia)
 - Bayesian epistemology:



UPPSALA
UNIVERSITET

Bayesian epistemology

- Introduction
 - Bayesian Reasoning
 - Bayesian Statistical inference
 - Probabilistic Modeling
 - Bayesian Computation
- What is **epistemology**?
"the theory of knowledge" (Wikipedia)
 - Bayesian epistemology:
 - State of knowledge is a probability distribution,
 - e.g. unlike Popperian approaches, we can talk about $P(\text{black swan})$
 - (New) Research in Philosophy of Science



UPPSALA
UNIVERSITET

- Introduction
- Bayesian Reasoning
- **Bayesian Statistical inference**
- Probabilistic Modeling
- Bayesian Computation

Section 3

Bayesian Statistical inference



UPPSALA
UNIVERSITET

Statistical inference

- Introduction
 - Bayesian Reasoning
 - Bayesian Statistical inference
 - Probabilistic Modeling
 - Bayesian Computation
- Draw conclusions of **unobserved** entities, based on **data**



- Introduction
 - Bayesian Reasoning
 - **Bayesian Statistical inference**
 - Probabilistic Modeling
 - Bayesian Computation
- Draw conclusions of **unobserved** entities, based on **data**
 - Different types of unobserved entities
 - **potentially observed**: future observations, treatment effects
 - **parameters**: data-generating process (e.g. regression coefficients)



- Introduction
- Bayesian Reasoning
- **Bayesian Statistical inference**
- Probabilistic Modeling
- Bayesian Computation

- $p(\cdot|\cdot)$: conditional pdf/pmf
- $p(\cdot)$: marginal pdf/pmf
- $P(\cdot)$ or $Pr(\cdot)$: probability, e.g. $P(\theta > 0) = \int_0^\infty p(\theta)d\theta$
- Random variable: $\theta \sim N(\mu, \sigma)$
- pdf/pmf: $p(\theta) = N(\theta|\mu, \sigma)$



- Introduction
- Bayesian Reasoning
- **Bayesian Statistical inference**
- Probabilistic Modeling
- Bayesian Computation

- Model parameters: $\theta = (\theta_1, \dots, \theta_p)$
- Observed data: $y = (y_1, \dots, y_n)$
 - y_i can be a vector and is assumed to come from a **data generating process** (probabilistic model)
- Potentially observed data: $\tilde{y} = (\tilde{y}_1, \dots, \tilde{y}_m)$
- Observed (known) covariates: x



- Introduction
- Bayesian Reasoning
- **Bayesian Statistical inference**
- Probabilistic Modeling
- Bayesian Computation

- Model parameters: $\theta = (\theta_1, \dots, \theta_p)$
- Observed data: $y = (y_1, \dots, y_n)$
 - y_i can be a vector and is assumed to come from a **data generating process** (probabilistic model)
- Potentially observed data: $\tilde{y} = (\tilde{y}_1, \dots, \tilde{y}_m)$
- Observed (known) covariates: x
- I'm sloppy: e.g. often implicitly condition on x , i.e. $p(\theta|y) = p(\theta|y, x)$



UPPSALA
UNIVERSITET

The basic steps of Bayesian inference

- Introduction
- Bayesian Reasoning
- **Bayesian Statistical inference**
- Probabilistic Modeling
- Bayesian Computation

1. Setting up a full probability model $p(y|\theta) \cdot p(\theta) = p(y, \theta)$



The basic steps of Bayesian inference

- Introduction
- Bayesian Reasoning
- **Bayesian Statistical inference**
- Probabilistic Modeling
- Bayesian Computation

1. Setting up a full probability model $p(y|\theta) \cdot p(\theta) = p(y, \theta)$
2. Conditioning on observed data y to calculate the posterior distribution $p(\theta|y)$



The basic steps of Bayesian inference

- Introduction
- Bayesian Reasoning
- **Bayesian Statistical inference**
- Probabilistic Modeling
- Bayesian Computation

1. Setting up a full probability model $p(y|\theta) \cdot p(\theta) = p(y, \theta)$
2. Conditioning on observed data y to calculate the posterior distribution $p(\theta|y)$
3. Evaluate the model. If not satisfied, go back to 1. Else use the model.



UPPSALA
UNIVERSITET

Bayesian Inference I

- Introduction
 - Bayesian Reasoning
 - Bayesian Statistical inference
 - Probabilistic Modeling
 - Bayesian Computation
- Bayesian conclusions about θ or \tilde{y} are made using probabilities, conditional on data y



UPPSALA
UNIVERSITET

Bayesian Inference I

- Introduction
 - Bayesian Reasoning
 - **Bayesian Statistical inference**
 - Probabilistic Modeling
 - Bayesian Computation
- Bayesian conclusions about θ or \tilde{y} are made using probabilities, **conditional on data y**
 - We state our uncertainty about θ or \tilde{y} as **distributions**



- Introduction
 - Bayesian Reasoning
 - **Bayesian Statistical inference**
 - Probabilistic Modeling
 - Bayesian Computation
- Bayesian conclusions about θ or \tilde{y} are made using probabilities, **conditional on data** y
 - We state our uncertainty about θ or \tilde{y} as **distributions**
 - potentially observed: $p(\tilde{y}|y)$
 - parameters: $p(\theta|y)$



Bayesian Inference: Setting up the model

- Introduction
- Bayesian Reasoning
- **Bayesian Statistical inference**
- Probabilistic Modeling
- Bayesian Computation

- Start out with a **model**: a *joint distribution* for data and parameters:

$$p(y, \theta) = p(y|\theta)p(\theta)$$



- Introduction
- Bayesian Reasoning
- **Bayesian Statistical inference**
- Probabilistic Modeling
- Bayesian Computation

- Start out with a **model**: a *joint distribution* for data and parameters:

$$p(y, \theta) = p(y|\theta)p(\theta)$$

- $p(y|\theta)$ is our data model, and when conditioned on y , the *likelihood*



- Introduction
- Bayesian Reasoning
- **Bayesian Statistical inference**
- Probabilistic Modeling
- Bayesian Computation

- Start out with a **model**: a *joint distribution* for data and parameters:

$$p(y, \theta) = p(y|\theta)p(\theta)$$

- $p(y|\theta)$ is our data model, and when conditioned on y , the *likelihood*
- $p(\theta)$ is our prior distribution for our parameters



Bayesian Inference: Computing the posterior

- Conditioning on data y , using *Bayes theorem*, we can compute the *posterior distribution*

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)}$$

where

$$p(y) = \int p(y|\theta)p(\theta)d\theta$$

- Introduction
- Bayesian Reasoning
- **Bayesian Statistical inference**
- Probabilistic Modeling
- Bayesian Computation



Bayesian Inference: Computing the posterior

- Introduction
- Bayesian Reasoning
- **Bayesian Statistical inference**
- Probabilistic Modeling
- Bayesian Computation

- Conditioning on data y , using *Bayes theorem*, we can compute the *posterior distribution*

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)}$$

where

$$p(y) = \int p(y|\theta)p(\theta)d\theta$$

- $p(y)$ is the *marginal likelihood*



Bayesian Inference: Computing the posterior

- Conditioning on data y , using *Bayes theorem*, we can compute the *posterior distribution*

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)}$$

where

$$p(y) = \int p(y|\theta)p(\theta)d\theta$$

- $p(y)$ is the *marginal likelihood*
- $p(\theta|y)$ summarize our knowledge about θ



Bayesian Inference: Computing the posterior

- Conditioning on data y , using *Bayes theorem*, we can compute the *posterior distribution*

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)}$$

where

$$p(y) = \int p(y|\theta)p(\theta)d\theta$$

- $p(y)$ is the *marginal likelihood*
- $p(\theta|y)$ summarize our knowledge about θ
- Bayesian statistics obey the *likelihood principle*: data only affects $p(\theta|y)$ through the likelihood $p(y|\theta)$



UPPSALA
UNIVERSITET

Predictions

- Introduction
 - Bayesian Reasoning
 - **Bayesian Statistical inference**
 - Probabilistic Modeling
 - Bayesian Computation
- How to do inference on an unknown (potential) observable \tilde{y} ? E.g. a future observation



- Introduction
- Bayesian Reasoning
- **Bayesian Statistical inference**
- Probabilistic Modeling
- Bayesian Computation

- How to do inference on an unknown (potential) observable \tilde{y} ? E.g. a future observation
- We use our *data* model and our posterior and *marginalize* over the uncertainty

$$p(\tilde{y}|y) = \int p(\tilde{y}|\theta)p(\theta|y)d\theta$$

- The *posterior predictive distribution*



UPPSALA
UNIVERSITET

Example: Updating uncertainty

- Introduction
- Bayesian Reasoning
- **Bayesian Statistical inference**
- Probabilistic Modeling
- Bayesian Computation



UPPSALA
UNIVERSITET

Example: Updating uncertainty

- Probability of red $\frac{\#red}{\#red + \#yellow} = \theta$

- Introduction
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation



- Introduction
- Bayesian Reasoning
- **Bayesian Statistical inference**
- Probabilistic Modeling
- Bayesian Computation

Example: Updating uncertainty

- Probability of red $\frac{\#red}{\#red + \#yellow} = \theta$
- $p(y = \text{red}|\theta) = \theta$ aleatoric uncertainty



- Introduction
- Bayesian Reasoning
- **Bayesian Statistical inference**
- Probabilistic Modeling
- Bayesian Computation

Example: Updating uncertainty

- Probability of red $\frac{\#red}{\#red + \#yellow} = \theta$
- $p(y = \text{red}|\theta) = \theta$ aleatoric uncertainty
- $p(\theta)$ epistemic uncertainty



- Introduction
- Bayesian Reasoning
- **Bayesian Statistical inference**
- Probabilistic Modeling
- Bayesian Computation

Example: Updating uncertainty

- Probability of red $\frac{\#red}{\#red + \#yellow} = \theta$
- $p(y = \text{red}|\theta) = \theta$ aleatoric uncertainty
- $p(\theta)$ epistemic uncertainty
- Picking many chips updates our uncertainty about the proportion
- $p(\theta|y = \text{red, yellow, red, red, } \dots) = ?$



- Introduction
- Bayesian Reasoning
- **Bayesian Statistical inference**
- Probabilistic Modeling
- Bayesian Computation

Example: Updating uncertainty

- Probability of red $\frac{\#red}{\#red + \#yellow} = \theta$
- $p(y = \text{red}|\theta) = \theta$ aleatoric uncertainty
- $p(\theta)$ epistemic uncertainty
- Picking many chips updates our uncertainty about the proportion
- $p(\theta|y = \text{red, yellow, red, red, } \dots) = ?$
- Bayes rule $p(\theta|y) = \frac{p(y|\theta)p(\theta)}{\int p(y|\theta)p(\theta)d\theta}$



- Introduction
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

Model vs. Likelihood

- Bayes rule

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)} \propto p(y|\theta)p(\theta)$$

- Data model: $p(y|\theta)$ as a function of y given fixed θ describes the **aleatoric** uncertainty
- Likelihood: $p(y|\theta) = L(\theta|y)$ as a function of θ given fixed y provides information about epistemic uncertainty. but is not a probability distribution.



- Introduction
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

Model vs. Likelihood

- Bayes rule

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)} \propto p(y|\theta)p(\theta)$$

- Data model: $p(y|\theta)$ as a function of y given fixed θ describes the **aleatoric** uncertainty
- Likelihood: $p(y|\theta) = L(\theta|y)$ as a function of θ given fixed y provides information about epistemic uncertainty, but is not a probability distribution.
- Bayes rule combines the likelihood with prior uncertainty $p(\theta)$ and transforms them to updated posterior uncertainty $p(\theta|y)$



UPPSALA
UNIVERSITET

Example application: Effects on roaches

- Question: Effect of treatments on captured roaches

- Introduction
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation



UPPSALA
UNIVERSITET

Example application: Effects on roaches

- Question: Effect of treatments on captured roaches
- Outcome (y): The number of roaches
- Coefficients (x): Treatment, Senior home, Pre-treatment number of roaches

- Introduction
- Bayesian Reasoning
- **Bayesian Statistical inference**
- Probabilistic Modeling
- Bayesian Computation



- Introduction
- Bayesian Reasoning
- **Bayesian Statistical inference**
- Probabilistic Modeling
- Bayesian Computation

Example application: Effects on roaches

- Question: Effect of treatments on captured roaches
- Outcome (y): The number of roaches
- Coefficients (x): Treatment, Senior home, Pre-treatment number of roaches
- Data model:

$$p(y|\theta) = Po(\lambda) = \frac{1}{y!} \lambda^y \exp(-\lambda)$$

where

$$\lambda = \exp(\alpha + \beta X)$$

- Prior:

$$p(\theta) = p(\alpha) = p(\beta_k) = N(\theta|\mu, \sigma)$$



- Introduction
- Bayesian Reasoning
- **Bayesian Statistical inference**
- Probabilistic Modeling
- Bayesian Computation

Example application: Effects on roaches

- Question: Effect of treatments on captured roaches
- Outcome (y): The number of roaches
- Coefficients (x): Treatment, Senior home, Pre-treatment number of roaches
- Data model:

$$p(y|\theta) = Po(\lambda) = \frac{1}{y!} \lambda^y \exp(-\lambda)$$

where

$$\lambda = \exp(\alpha + \beta X)$$

- Prior:

$$p(\theta) = p(\alpha) = p(\beta_k) = N(\theta|\mu, \sigma)$$

- Posterior:

$$p(\alpha, \beta|y) \propto N(\alpha|\mu, \sigma) \prod_{k=1}^K N(\beta_k|\mu, \sigma) \prod_{i=1}^N Po(\exp(\alpha + \beta x_i))$$



- Introduction
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

Example application: Effects on roaches

- Question: Effect of treatments on captured roaches
- Outcome (y): The number of roaches
- Coefficients (x): Treatment, Senior home, Pre-treatment number of roaches
- Data model:

$$p(y|\theta) = Po(\lambda) = \frac{1}{y!} \lambda^y \exp(-\lambda)$$

where

$$\lambda = \exp(\alpha + \beta X)$$

- Prior:

$$p(\theta) = p(\alpha) = p(\beta_k) = N(\theta|\mu, \sigma)$$

- Posterior:

$$p(\alpha, \beta|y) \propto N(\alpha|\mu, \sigma) \prod_{k=1}^K N(\beta_k|\mu, \sigma) \prod_{i=1}^N Po(\exp(\alpha + \beta x_i))$$

- Predictive distribution:

$$p(\tilde{y}|y, \tilde{x}) = \int Po(\exp(\alpha + \beta \tilde{x})) p(\alpha, \beta|y) d\alpha d\beta$$



UPPSALA
UNIVERSITET

- Introduction
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

Section 4

Probabilistic Modeling



UPPSALA
UNIVERSITET

- Introduction
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

The art of probabilistic modeling

- Subjectivity: we need to specify both $p(\theta)$ and $p(y|\theta)$
- The art of probabilistic modeling is to describe in a mathematical form (model and prior distributions) what we already know and what we don't know



UPPSALA
UNIVERSITET

- Introduction
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

The art of probabilistic modeling

- Subjectivity: we need to specify both $p(\theta)$ and $p(y|\theta)$
- The art of probabilistic modeling is to describe in a mathematical form (model and prior distributions) what we already know and what we don't know
- “Easy” part is to use Bayes rule to update the uncertainties
 - computational challenges



- Introduction
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

The art of probabilistic modeling

- Subjectivity: we need to specify both $p(\theta)$ and $p(y|\theta)$
- The art of probabilistic modeling is to describe in a mathematical form (model and prior distributions) what we already know and what we don't know
- “Easy” part is to use Bayes rule to update the uncertainties
 - computational challenges
- Other parts of the art of probabilistic modeling are, for example,
 - model checking/assessment: is data in conflict with our prior knowledge?
 - model choice: which model should we use?
 - presentation: presenting the model and the results to the application experts



- Introduction
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

Example applications

- Galaxy clusters for cosmology
- Coagulation of blood
- Gene regulation
- Pharmacokinetics and -dynamics
- Decision support
- Effects of nutrition for diabetes
- Evolutionary anthropology
- Clinical trial designs
- Daily demand for gas
- Brain structure trees
- School enrollment
- Sports
- Product demand
- Cocoa bean fermentation
- Marine propulsion power
- Alcohol consumption trends
- Flood probability
- Instantaneous heart rate distributions
- Drug dosing regimens in pediatrics
- Human T stem cell memory cells
- Fairness in university admission policies
- Destruction of bacteria and bacterial spores under heat



- Benefits of Bayesian approach
 - uncertainty is the foundation
 - easier interpretation of uncertainty intervals
 - integrate over uncertainties to focus to interesting parts
 - straight-forward predictive distributions
 - use relevant prior information
 - straight-forward hierarchical models
 - model checking and evaluation
 - easier to build and evaluate very complex models

- Introduction
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation



- Introduction
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

- Benefits of Bayesian approach
 - uncertainty is the foundation
 - easier interpretation of uncertainty intervals
 - integrate over uncertainties to focus to interesting parts
 - straight-forward predictive distributions
 - use relevant prior information
 - straight-forward hierarchical models
 - model checking and evaluation
 - easier to build and evaluate very complex models
- Complications of Bayesian approach
 - most models does not have nice analytical posteriors
 - we need to *approximate* our posterior
 - can be computationally costly/take long time to estimate complex models



Example: Swedish vote intention estimation

- Introduction
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

- Estimation of Swedish vote intention

<https://www.bottenada.se/>

- approximately 29 000 parameters
(every party, every week, since the 90ies)
- dynamic house bias
- autoregressive industry bias
- elections are known
- composite data



Example: Swedish vote intention estimation

- Introduction
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

- Estimation of Swedish vote intention
<https://www.bottenada.se/>
 - approximately 29 000 parameters
(every party, every week, since the 90ies)
 - dynamic house bias
 - autoregressive industry bias
 - elections are known
 - composite data
- But it takes 12 h to estimate the model



UPPSALA
UNIVERSITET

- Introduction
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- **Bayesian Computation**

Section 5

Bayesian Computation



We want to compute expectations with respect to a posterior distribution $p(\theta|y)$

$$\mathbb{E}_{\theta|y} [g(\theta)] = \int p(\theta|y)g(\theta)d\theta$$

- Introduction
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

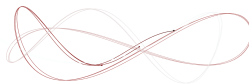
- Analytic
 - only for very simple models
- Monte Carlo, Markov chain Monte Carlo
 - generic, high-quality
 - computationally expensive
- Distributional approximations
 - e.g. Laplace, variational inference
 - less generic, but can be much faster with sufficient accuracy



UPPSALA
UNIVERSITET

- Introduction
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- **Bayesian Computation**

Probabilistic programming



Enables agile workflow for developing probabilistic models

language – automated inference – diagnostics



mc-stan.org



- Introduction
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- **Bayesian Computation**

Binomial model for treatment/control comparison

- Two groups of patients: treatment and control
 - Binary outcome (1 treatment, 0 no treatment)
 - Is the treatment useful?
- Model:
Likelihood

$$y_0 \sim \text{Bin}(\theta_0, N_0)$$

$$y_1 \sim \text{Bin}(\theta_1, N_1)$$



- Introduction
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

Binomial model for treatment/control comparison

- Two groups of patients: treatment and control
 - Binary outcome (1 treatment, 0 no treatment)
 - Is the treatment useful?

- Model:
Likelihood

$$y_0 \sim \text{Bin}(\theta_0, N_0)$$

$$y_1 \sim \text{Bin}(\theta_1, N_1)$$

Prior

$$\theta_0 \sim \text{Beta}(1, 1)$$

$$\theta_1 \sim \text{Beta}(1, 1)$$



- Introduction
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

Binomial model for treatment/control comparison

- Two groups of patients: treatment and control
 - Binary outcome (1 treatment, 0 no treatment)
 - Is the treatment useful?

- Model:
Likelihood

$$y_0 \sim \text{Bin}(\theta_0, N_0)$$

$$y_1 \sim \text{Bin}(\theta_1, N_1)$$

Prior

$$\theta_0 \sim \text{Beta}(1, 1)$$

$$\theta_1 \sim \text{Beta}(1, 1)$$

Goal (odds ratio), e.g. $P(OR > 1)$

$$OR = \frac{\frac{\theta_1}{1-\theta_1}}{\frac{\theta_0}{1-\theta_0}}$$



- Introduction
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

Binomial model for treatment/control comparison

```
data {  
  int<lower=0> N1;  
  int<lower=0> y1;  
  int<lower=0> N0;  
  int<lower=0> y0;  
}  
parameters {  
  real<lower=0, upper=1> theta1;  
  real<lower=0, upper=1> theta0;  
}  
model {  
  theta1 ~ beta(1,1);  
  theta0 ~ beta(1,1);  
  y1 ~ binomial(N1, theta1);  
  y0 ~ binomial(N0, theta0);  
}  
generated quantities {  
  real oddsratio;  
  oddsratio = (theta1/(1-theta1))/(theta0/(1-theta0));  
}
```




Binomial model for treatment/control comparison

- Introduction
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- **Bayesian Computation**

RStanARM

```
fit_bin2 <- stan_glm(y/N ~ grp2, family = binomial(),  
                    data = d_bin2, weights = N)
```



UPPSALA
UNIVERSITET

Modeling nature

- Introduction
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

- Drop a ball from different heights and measure time



UPPSALA
UNIVERSITET

Modeling nature

- Introduction
 - Bayesian Reasoning
 - Bayesian Statistical inference
 - Probabilistic Modeling
 - Bayesian Computation
- Drop a ball from different heights and measure time
 - Newton
 - air resistance, air pressure, shape and surface structure of the ball
 - relativity



UPPSALA
UNIVERSITET

Modeling nature

- Introduction
 - Bayesian Reasoning
 - Bayesian Statistical inference
 - Probabilistic Modeling
 - Bayesian Computation
- Drop a ball from different heights and measure time
 - Newton
 - air resistance, air pressure, shape and surface structure of the ball
 - relativity
 - Taking into account the accuracy of the measurements, how accurate model is needed?



- Introduction
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

- Drop a ball from different heights and measure time
 - Newton
 - air resistance, air pressure, shape and surface structure of the ball
 - relativity
- Taking into account the accuracy of the measurements, how accurate model is needed?
 - often simple models are adequate and useful
 - *All models are wrong, but some of them are useful,* George P. Box



Recap: Uncertainty and probabilistic modeling

- Introduction
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

- Two types of uncertainty: aleatoric and epistemic
- Representing uncertainty with probabilities
- Updating uncertainty using data



UPPSALA
UNIVERSITET

Rest of the course

- Introduction
- Bayesian Reasoning
- Bayesian Statistical inference
- Probabilistic Modeling
- Bayesian Computation

- Basic models which can be used as building blocks for complex ones
- Basic computation of posterior and predictive distributions
- Typical simple scientific data analysis cases
 - e.g. comparison of treatments
- Presentation of the results