Bayesian Statistics

Introduction and Bayes' Rule

Alexandra Sarafoglou

June 1, 2026

Teaching Team: Lectures, Quizzes, and Exam

Alexandra Sarafoglou

Johnny van Doorn

Mathilde ter Veen

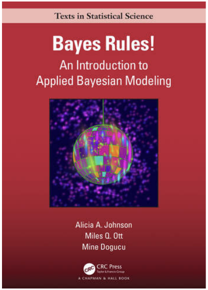
Teaching Team: Group Projects

Vipasha Goyal

Florian Metwaly

Nikola Sekulovski

Course Book



https://www.bayesrulesbook.com

This course covers the first chapters of the book. The remaining chapters will be covered in the Research Master Psychology Course *Computational Statistics*.

Intended Learning Outcomes

At the end of the course...

- you are able to define the likelihood, the prior, and the posterior distributions for the models covered in the course.
- you are able to conceptually explain the various components of Bayesian analysis covered in this course.
- you are able to perform a Bayesian analysis, using the models and methodologies covered in the course, to determine if there is an effect (Bayesian hypothesis testing), what the size of that effect is (posterior estimation), and to express how certain they are about their conclusions.
- you are able to report a Bayesian analysis in an empirical report.

Intended Learning Outcomes

In addition, we want to introduce you to the Bayesian research we do in our department:

- Alexandra: Meta-science/Epistemic uncertainty/Computational modeling
- Johnny: Methodologies to compute Bayes factors/ Accessible statistics in JASP
- Maarten: Network Psychometrics and Sampling Methods

Contact us via email, if you would like to know about current thesis and internship projects.

Schedule

Week 1

- Introduction & Bayes' rule
- The Beta-Binomial model & sequential testing
- Quiz and Q&A

Week 2

- Conjugacy, choosing suitable priors, & prior prediction
- Obtaining posterior, posterior inference & hypothesis testing
- Q&A

Week 3

- Posterior prediction & hypothesis testing
- Reporting guidelines & Exam Prep/Q&A
- **Exam**

Week 4

- Group project (introduction)
- Group project
- Group presentation

Attendance

Attendance during the lectures is **not** mandatory. Attendance during the group project in week 4 is **not** mandatory, but desired.

Support

We try to support you in your learning as best as we can:

- Ungraded weekly homework assignments that are representative of the exam questions
- Ungraded quiz "Kahoot" in two lectures with questions that are representative for the quiz
- Office hours on Wednesdays to discuss questions about the homework, exam/quiz, or any other issues
- Weekly Q&A's about the course materials
- Week 4: Supervised group work
- Week 4: Ungraded group presentations as feedback moments for the group report

Office hours

Office hours (with Alexandra) take place from 13:00 to 14:00 in REC G0.33, on

- Wednesday, June 03
- Wednesday, June 10
- Wednesday, June 17
- Wednesday, June 24

Examination

- Week 1: Onboarding Quiz
  - Friday, June 5, 12:00 - 13:00 in REC M1.02.
  - Relevant: Materials of the first two lectures
  - Pen-and-paper format
  - No cheat sheet
  - You do *not* need to pass the quiz to pass the course. No resit.
- Week 3: Exam
  - Friday, June 19, 09:00 - 11:00 in IWO 4.04C (Blauw)
  - Relevant: All lecture materials
  - ANS exam including R. No internet access; but book will be made available as a PDF.
  - 1 A4 cheat sheet (written/printed on both sides)
  - You need to pass the exam to pass the course. Resit July 7.

Examination

- Week 1 - 2: Submission of a Q&A question
  - Submit one question about the lecture or homework either in week 1 or week 2
  - Assignment is graded pass/fail
- Week 4: Group report
  - You can choose from 3 projects. When assigned, you will work in a group of 5 students
  - Monday and Wednesday: introduction to the group project topic & supervised group work
  - Friday: Ungraded group presentations and feedback
  - Friday, June 26, 21:00: Deadline for the group report
  - Group report will feature a paragraph on contributorship
  - Coordinate with your group if you miss seminars in week 4

Examination

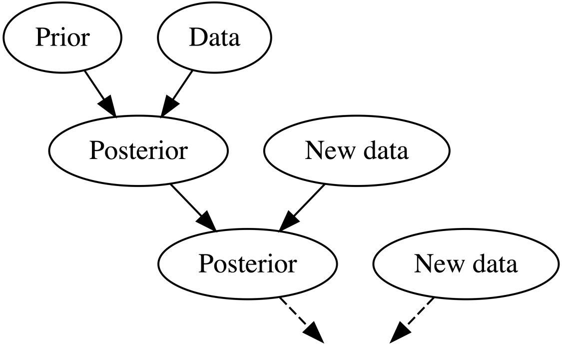
The final grade will be calculated as the average of the following components: the quiz (5%), the submitted questions (5%), the exam (50%), and the group assignment (40%). To pass the course, the grade of the exam should be 5.5 or above.

Bayesian versus frequentist knowledge building

Bayesian statistics is about...

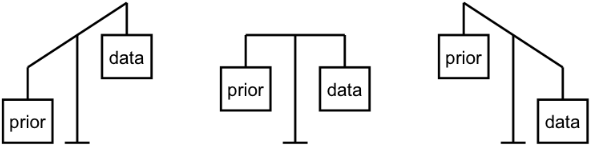
- using our existing knowledge, our *prior beliefs*, to make sense of the world
- continuously update our beliefs with data
- learning from past experiences and data to form our *posterior beliefs*

Bayesian versus frequentist knowledge building



Bayesian versus frequentist knowledge building

We chose the prior based on theoretical considerations.



- Guessing parameter in a task *cannot* be lower than chance level
- Combining an original study with replication data to establish an effect
- Older adults have lower values for their source-memory parameter than younger adults

Frequentist knowledge building

Frequentist analysis gives no voice to our prior knowledge:

Summary: Bayesian versus frequentist knowledge building

Concept	**Frequentist interpretation**	**Bayesian interpretation**
Probability	The long-run relative *frequency* of a repeatable event (hence "frequentist")	A measure of the relative plausibility of an event
Data	Data alone should drive our analysis	Data should be weighed against our prior beliefs
Research questions	If the hypothesis is not correct, what are the chances I would have observed these data or more extreme ones?	In light of these data, what is the probability that the hypothesis is correct?

Summary: Bayesian knowledge building

- We construct our posterior knowledge by balancing information from our data with our prior knowledge
- As more data come in, we continue to refine this knowledge and (most of the time) the influence of our original prior fades into the background

Building a Bayesian model

- Explore properties of probability models
- Explore the marginal, conditional, and joint probability
- Conduct your first formal Bayesian analysis using Bayes' Rule

Fake News

Article Title:

NYC Terrorist Ahmad Rahami Sued Police Department for 'Religious Persecution' in 2011!

- What is the probability that *this* article is fake?

Fake News

Article Title:

NYC Terrorist Ahmad Rahami Sued Police Department for 'Religious Persecution' in 2011!

How can we use the Bayesian philosophy to make a distinction between real and fake news? We have the following information:

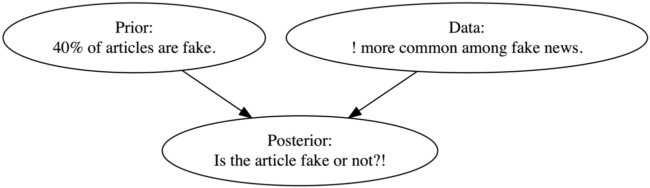
- Database consists of 150 articles
- 60% (90/150) of articles are real
- 2.22% (2/90) of real news titles use an exclamation point
- 40% (60/150) of articles are fake
- 26.67% (16/60) of fake news titles use an exclamation point

Fake News

Our goal is to fill in the following table:

	B	B^c	**Total**
A	?	?	?
A^c	?	?	?
Total	0.4	0.6	1.0

Ingredients for Bayesian inference



Variables

Two variables:

- An article is fake vs real
- An article title contains an exclamation point

We can represent the randomness in these variables using probability models:

- A **prior probability model** for our prior understanding of whether an article is fake
- A model for interpreting the exclamation point **data**
- A **posterior probability model** which summarizes the posterior plausibility that the article is fake

Ingredients for Bayesian inference

Probability model:

A valid probability model must:

1. Account for all possible events
2. Assign probabilities to each event
3. These probabilities sum to one

Building a Bayesian model

Step 1: Formalize our prior beliefs as a probability model.
Let B denote the event that an article is fake and B^c denote the event that it's not fake, we have:

$$P(B) = 0.40 \text{ and } P(B^c) = 0.60.$$

Building a Bayesian model

Step 2: We collected a new article and we will now summarize our insights about it.

- *If* an article is fake, *then* there is a roughly 26.67% chance it uses exclamation points
- *If* an article is real, *then* there is only a roughly 2.22% chance it uses exclamation points

The occurrence of the exclamation point is *conditioned* upon whether the article is fake. We specify this dependence by conditional probabilities.

Building a Bayesian model

Let A denote the event that an article uses an exclamation point. We can formulate the conditional probability as follows:

$$P(A \mid B) = 0.2667 \text{ and } P(A \mid B^c) = 0.0222.$$

With conditional probabilities, we learn the extent to which information about event B (the article is fake) informs our understanding of event A (the article uses an exclamation point).

Building a Bayesian model

Conditional versus unconditional probability:

Let A and B be two events. The **unconditional probability** A, $P(A)$, measures the probability of observing A, without any knowledge of B. The **conditional probability** of A given B, $P(A \mid B)$, measures the probability of observing A in light of the information that B occurred.

Building a Bayesian model

Important: The order of conditioning is important: $P(A \mid B) \neq P(B \mid A)$. For instance, roughly 100% of puppies are adorable. Thus, if the next object you pass on the street is a puppy, $P(\text{adorable} \mid \text{puppy}) = 1$. However, the reverse is not true. Not every adorable object is a puppy, thus $P(\text{puppy} \mid \text{adorable}) < 1$.

Building a Bayesian model

Important: Information about event B does not always affect our understanding of event A. For instance, suppose we also have information about whether a given article includes a graph or a picture. Before viewing the article, the probability that it is fake is $P(B) = 0.40$. However, the presence of graphs or pictures is not a reliable indicator of an article's credibility. Even if we know the article contains a graph or picture, the probability that it is fake remains 40%: $P(B \mid \text{picture}) = 0.40$.

Building a Bayesian model

Independent events:

Two events A and B are independent if and only if the occurrence of B does not tell us anything about the occurrence of A:

$$P(A \mid B) = P(A).$$

Difference between conditional probability and likelihood

- Normally, $P(A \mid B)$ means we *know* about B and ask about A.
- In our fake news example, however, we *observe* A (use of exclamation points) and we ask whether the article is fake (B).
- Importantly, we do not want to refer to a posterior probability $P(B \mid A)$ but to the likelihood of observing A given B

Difference between conditional probability and likelihood

To differentiate these two cases, we introduce the **likelihood function** notation $L(B \mid A)$.

- The numerical values of $L(B \mid A)$ and $P(A \mid B)$ are the same, but they serve different purposes.
- $P(A \mid B)$: Used to describe the probability of observing A after observing that B is true.
- $L(B \mid A)$: Used to describe the probability of the observed data A, as a function of different possible values of B. Or formally: used to compare how well different values of B explain the observed data A.
- $L(B \mid A) = P(A \mid B)$ and $L(B^c \mid A) = P(A \mid B^c)$.

Difference between conditional probability and likelihood

Probability vs likelihood:

When B is known, the conditional probability function $P(\cdot \mid B)$ allows us to compare the probabilities of an unknown event, A or A^c, occurring with B:

$$P(A \mid B) \text{ vs } P(A^c \mid B).$$

When A is known, the likelihood function $L(\cdot \mid A) = P(A \mid \cdot)$ allows us to evaluate the relative compatibility of data A with events, B or B^c:

$$L(B \mid A) \text{ vs } L(B^c \mid A).$$

Building a Bayesian model

Event	B	B^c	**Total**
Prior probability	0.4	0.6	1.0
Likelihood $L(\cdot \mid A)$	0.2667	0.0222	0.2889

- The prior probabilities add up to 1.
- The likelihoods do not add up to 1 $\rightarrow$ the likelihood function is *not* a valid probability model.

Building a Bayesian model

Next, we look at the marginal probability of observing exclamation points across *all* news articles, $P(A)$. To do this, we need our prior model and the likelihood function to fill in the table below:

	B	B^c	**Total**
A	?	?	?
A^c	?	?	?
Total	0.4	0.6	1.0

First, let us focus on the B (fake article) column.

Building a Bayesian model

	B	B^c	**Total**
A	?	?	?
A^c	?	?	?
Total	0.4	0.6	1.0

First, let us focus on the B (fake article) column. We have two groups:

- proportion of all articles that are *fake* and use *exclamation points*, $A \cap B$
- proportion of all articles that are *fake* and don't use *exclamation points*, $A^c \cap B$

Building a Bayesian model

We know:

- 40% of articles are fake; $P(B) = 0.40$
- 26.67% of fake articles use exclamation points; $P(A \mid B) = 0.2667$
- Thus, considering all articles, 26.67% of 40% of articles are fake *and* have exclamation points
- The **joint probability** of observing both A and B is:

$$P(A \cap B) = P(A \mid B)P(B) = 0.2667 \times 0.4 = 0.1067.$$

Building a Bayesian model

Calculating joint probabilities:

For events A and B, the joint probability $A \cap B$ is calculated by weighting the conditional probability of A given B by the marginal probability of B:

$$P(A \cap B) = P(A \mid B)P(B).$$

When A and B are independent,

$$P(A \cap B) = P(A)P(B).$$

Building a Bayesian model

	B	B^c	**Total**
A	0.1067	?	?
A^c	?	?	?
Total	0.4	0.6	1.0

$P(A^c \cap B) = P(B) - P(A \cap B) = 0.2933$. But we can also get this number using conditional probabilities.

Building a Bayesian model

Since 26.67% of fake articles use exclamation points, 73.33% do not. Therefore:

- The conditional probability that an article does *not* use exclamation points (A^c) given it is fake (B) is: $P(A^c \mid B) = 1 - P(A \mid B) = 1 - 0.2667 = 0.7333$.
- Thus, considering all articles, 73.33% of 40% of articles are fake *and* do not use exclamation points
- The **joint probability** of observing both A^c and B is:

$$P(A^c \cap B) = P(A^c \mid B)P(B) = 0.7333 \times 0.4 = 0.2933.$$

Building a Bayesian model

	B	B^c	**Total**
A	0.1067	?	?
A^c	0.2933	?	?
Total	0.4	0.6	1.0

The **marginal probability** of observing a fake article is the sum of the joint events:
$P(B) = P(A \cap B) + P(A^c \cap B) = 0.1067 + 0.2933 = 0.40.$

Building a Bayesian model

We can take the same approach to the real articles:

	B	B^c	**Total**
A	0.1067	?	?
A^c	0.2933	?	?
Total	0.4	0.6	1.0

- 60% of articles are real; $P(B^c) = 0.60$
- 2.22% of real articles use exclamation points; $P(A \mid B^c) = 0.0222$

Building a Bayesian model

We can take the same approach to the real articles:

- $P(A \cap B^c) = P(A \mid B^c)P(B^c) = 0.0222 \times 0.6 = 0.0133.$

	B	B^c	**Total**
A	0.1067	0.0133	?
A^c	0.2933	?	?
Total	0.4	0.6	1.0

Building a Bayesian model

We can take the same approach to the real articles:

- $P(A^c \mid B^c) = 1 - P(A \mid B^c) = 1 - 0.0222 = 0.9778.$
- $P(A^c \cap B^c) = P(A^c \mid B^c)P(B^c) = 0.9778 \times 0.6 = 0.5867.$

Building a Bayesian model

	B	B^c	**Total**
A	0.1067	0.0133	0.12
A^c	0.2933	0.5867	0.88
Total	0.4	0.6	1.0

The **marginal probability** of observing a real article is:
$P(B^c) = P(A \cap B^c) + P(A^c \cap B^c) = 0.0133 + 0.5867 = 0.60.$

Building a Bayesian model

Law of total probability

Let A and B be events with $P(B) > 0$ and $P(B^c) > 0$. The Law of Total Probability states:

$$P(A) = P(B) \times P(A \mid B) + P(B^c) \times P(A \mid B^c).$$

Building a Bayesian model

The ultimate question: What is the probability that this article is fake?

Article Title:

NYC Terrorist Ahmad Rahami Sued Police Department for 'Religious Persecution' in 2011!

Building a Bayesian model

Formally speaking, we want to calculate the **posterior probability** that the article is fake given that is uses exclamation points $P(B \mid A)$:

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}.$$

where **total probability** of $P(A)$ is:

$$P(A) = P(B)P(A \mid B) + P(B^c)P(A \mid B^c).$$

Building a Bayesian model

Written in terms of a prior, a likelihood, and a marginal probability the equation becomes:

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{P(B)L(B \mid A)}{P(A)},$$

and the **total probability** of $P(A)$ is:

$$P(A) = P(B)L(B \mid A) + P(B^c)L(B^c \mid A).$$

This gives us Bayes' rule:

$$\text{posterior} = \frac{\text{prior} \times \text{likelihood}}{\text{normalizing constant}}.$$

Building a Bayesian model

The ultimate question: What is the probability that this article is fake?

Article Title:

NYC Terrorist Ahmad Rahami Sued Police Department for 'Religious Persecution' in 2011!

$$P(B \mid A) = \frac{P(B)L(B \mid A)}{P(A)} = \frac{0.4 \times 0.2667}{0.12} = 0.889.$$

Building a Bayesian model

Calculating joint probabilities:

The conditional probability of A given B is calculated by:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}.$$

Summary

- Explore properties of probability models
- Explore the marginal, conditional, and joint probability
- Conduct your first formal Bayesian analysis using Bayes' Rule

Ungraded Homework Assignment (Exam Prep)

In the book:

- Exercise 1.8
- Exercise 2.4, 2.5., 2.9
- Exercise 2.17, 2.20, 2.21