

Bayesian Basics:

A Gentle introduction

Bill Thompson

biltho@mpi.nl

What is Bayesian Statistics?



What is Bayesian Statistics?



What is Bayesian Statistics **Not**?
Frequentist Statistics



PHILOSOPHICAL
TRANSACTIONS:

**An Essay towards Solving a
Problem in the Doctrine of
Chances. By the Late Rev. Mr.
Bayes, F. R. S. Communicated by
Mr. Price, in a Letter to John
Canton, A. M. F. R. S.**

Mr. Bayes and Mr. Price



PHILOSOPHICAL
TRANSACTIONS:

**An Essay towards Solving a
Problem in the Doctrine of
Chances. By the Late Rev. Mr.
Bayes, F. R. S. Communicated by
Mr. Price, in a Letter to John
Canton, A. M. F. R. S.**

Mr. Bayes and Mr. Price

Modern Title:

*A Method of Calculating the Exact Probability
of All Conclusions founded on Induction*

A Crowd-Sourced Intuitive Explanation

A Popular Question on **stackexchange.com**:

The screenshot shows a Stack Exchange question page. At the top, the StackExchange logo and navigation links (sign up, log in, tour, help) are visible. Below the header, the 'Cross Validated' logo is on the left, and navigation tabs (QUESTIONS, TAGS, USERS, BADGES, UNANSWERED, ASK QUESTION) are on the right. The main content area features the question title 'Bayesian and frequentist reasoning in plain English'. Below the title, the question text is: 'How would you describe in plain English the characteristics that distinguish Bayesian from Frequentist reasoning?'. To the left of the question text, there are up and down vote arrows, the number '253', and a star icon with the number '183'. Below the question text, there are tags 'bayesian' and 'frequentist', and links 'share', 'cite', and 'improve this question'. To the right of the question text, there is a box showing the question was 'edited Oct 4 '11 at 7:05' by user 'chl' with a reputation of 41.2k. Another box shows the question was 'asked Jul 19 '10 at 19:25' by user 'Daniel Vassallo' with a reputation of 1,369. On the far right, there is a sidebar with statistics: 'asked 7 years, 2 months ago', 'viewed 119,085 times', and 'active 4 months ago'. At the bottom right, there is a 'BLOG' section with a link to 'Podcast #115 - Stack Overflow Reads Mean Tweets'.

<https://stats.stackexchange.com/questions/22/bayesian-and-frequentist-reasoning-in-plain-english>

A Crowd-Sourced Intuitive Explanation

A Popular Question on **stackexchange.com**:

“How would you describe in plain English the characteristics that distinguish Bayesian from Frequentist reasoning?”

<https://stats.stackexchange.com/questions/22/bayesian-and-frequentist-reasoning-in-plain-english>

A Crowd-Sourced Intuitive Explanation

The Most Popular **Answer:**

Problem:

I have misplaced my phone somewhere in the home. I can use the phone locator on the base of the instrument to locate the phone and when I press the phone locator the phone starts beeping.

Which area of my home should I search?

<https://stats.stackexchange.com/questions/22/bayesian-and-frequentist-reasoning-in-plain-english>

A Crowd-Sourced Intuitive Explanation

The Most Popular **Answer:**

Frequentist Reasoning

I can hear the phone beeping. I also have a mental model which helps me identify the area from which the sound is coming. Therefore, upon hearing the beep, I infer the area of my home I must search to locate the phone.

<https://stats.stackexchange.com/questions/22/bayesian-and-frequentist-reasoning-in-plain-english>

A Crowd-Sourced Intuitive Explanation

The Most Popular **Answer**:

Bayesian Reasoning

*I can hear the phone beeping. Now, apart from a mental model which helps me identify the area from which the sound is coming from, **I also know the locations where I have misplaced the phone in the past.** So, I combine my inferences using the beeps and my prior information about the locations I have misplaced the phone in the past to identify an area I must search to locate the phone.*

<https://stats.stackexchange.com/questions/22/bayesian-and-frequentist-reasoning-in-plain-english>

Example #2:

A Mystery Cough

Data: A Doctor's patient keeps coughing

Example #2:

A Mystery Cough

Data: A Doctor's patient keeps coughing

Hypothesis #1:
The patient has a cold

Example #2:

A Mystery Cough

Data: A Doctor's patient keeps coughing

Hypothesis #1:

The patient has a cold

Hypothesis #2:

The patient has chewing gum stuck inside the lungs

Example #2:

A Mystery Cough

Data: A Doctor's patient keeps coughing

Hypothesis #1:

The patient has a cold

Hypothesis #2:

The patient has chewing gum stuck inside the lungs

Inferential Question:

What is the cause of the cough (a cold or chewing gum?)

Example #2:

A Mystery Cough

Bayesian statistics divides inference into two main ingredients:

1. The **Prior** probability of a hypothesis
2. The **Likelihood** of data under a hypothesis

Bayes Theorem (see later slide) provides a **formula for combining these two source of information** when performing inference.

Example #2:

A Mystery Cough

Data: A Doctor's patient keeps coughing

Hypothesis #1:

The patient has a cold

Hypothesis #2:

The patient has chewing gum stuck inside the lungs

Example #2:

A Mystery Cough

Data: A Doctor's patient keeps coughing

Hypothesis #1: The patient has a cold

Example #2:

A Mystery Cough

Data: A Doctor's patient keeps coughing

Hypothesis #1: The patient has a cold

Likelihood:

The likelihood of observing the data if the hypothesis were true

Example #2:

A Mystery Cough

Data: A Doctor's patient keeps coughing

Hypothesis #1: The patient has a cold

Likelihood:

The likelihood of observing the data if the hypothesis were true

$P(\text{Data} \mid \text{Hypothesis})$

Example #2:

A Mystery Cough

Data: A Doctor's patient keeps coughing

Hypothesis #1: The patient has a cold

Likelihood:

$P(\text{Data} \mid \text{Hypothesis \# 1}) = \text{High or Low?}$

Example #2:

A Mystery Cough

Data: A Doctor's patient keeps coughing

Hypothesis #1: The patient has a cold

Likelihood:

$$P(\text{Data} \mid \text{Hypothesis \# 1}) = \textit{High}$$

Example #2:

A Mystery Cough

Data: A Doctor's patient keeps coughing

Hypothesis #1: The patient has a cold

Hypothesis #2: The patient has chewing gum in the lungs

Likelihood:

$P(\text{Data} \mid \text{Hypothesis \# 2}) = \text{High or Low?}$

Example #2:

A Mystery Cough

Data: A Doctor's patient keeps coughing

Hypothesis #1: The patient has a cold

Hypothesis #2: The patient has chewing gum in the lungs

Likelihood:

$$P(\text{Data} \mid \text{Hypothesis \# 2}) = \textit{High}$$

Example #2:

A Mystery Cough

Data: A Doctor's patient keeps coughing

Hypothesis #1:

The patient has a cold

Hypothesis #2:

The patient has chewing gum in the lungs

Likelihood:

$P(\text{Data} \mid \text{Hypothesis})$

High

High

Example #2:

A Mystery Cough

Data: A Doctor's patient keeps coughing

Hypothesis #1:

The patient has a cold

Hypothesis #2:

The patient has chewing gum in the lungs

Likelihood:

$P(\text{Data} \mid \text{Hypothesis})$

High

High

Prior:

$P(\text{Hypothesis})$

Example #2:

A Mystery Cough

Data: A Doctor's patient keeps coughing

Hypothesis #1:

The patient has a cold

Hypothesis #2:

The patient has chewing gum in the lungs

Likelihood:

$P(\text{Data} \mid \text{Hypothesis})$

High

High

Prior:

$P(\text{Hypothesis})$

High

Example #2:

A Mystery Cough

Data: A Doctor's patient keeps coughing

Hypothesis #1:

The patient has a cold

Hypothesis #2:

The patient has chewing gum in the lungs

Likelihood:

$P(\text{Data} \mid \text{Hypothesis})$

High

High

Prior:

$P(\text{Hypothesis})$

High

Low

Example #2:

A Mystery Cough

Data: A Doctor's patient keeps coughing

Hypothesis #1:

The patient has a cold

Hypothesis #2:

The patient has chewing gum in the lungs

Likelihood:

$P(\text{Data} \mid \text{Hypothesis})$

High

High

+

+

Prior:

$P(\text{Hypothesis})$

High

Low

=

=

High

Medium

Example #2:

A Mystery Cough

Data: A Doctor's patient keeps coughing

Hypothesis #1:

The patient has a cold

Hypothesis #2:

The patient has chewing gum in the lungs

Likelihood:

$P(\text{Data} \mid \text{Hypothesis})$

High

High

+

+

Prior:

$P(\text{Hypothesis})$

High

Low

=

=

Winner! →

High

Medium

Example #2:

A Mystery Cough

Two problems with the formulation above:

- 1:** We're just adding highs and lows — is there a better formula?
- 2:** We chose a single winner — is there a better selection technique?

Example #2: **A Mystery Cough**

Two problems with the formulation above:

1: We're just adding highs and lows — is there a better formula?

2: We chose a single winner — is there a better selection technique?

solution:

Bayes' Rule!

$$P(\text{Hypothesis} \mid \text{Data}) = P(\text{Data} \mid \text{Hypothesis}) \times P(\text{Hypothesis})$$

Example #2: **A Mystery Cough**

Two problems with the formulation above:

1: We're just adding highs and lows — is there a better formula?

2: We chose a single winner — is there a better selection technique?

solution:
Bayes' Rule!

$$P(\text{Hypothesis} \mid \text{Data}) = \mathbf{P(\text{Data} \mid \text{Hypothesis})} \times P(\text{Hypothesis})$$



likelihood

Example #2: **A Mystery Cough**

Two problems with the formulation above:

1: We're just adding highs and lows — is there a better formula?

2: We chose a single winner — is there a better selection technique?

solution:
Bayes' Rule!

$$P(\text{Hypothesis} \mid \text{Data}) = P(\text{Data} \mid \text{Hypothesis}) \times \mathbf{P(\text{Hypothesis})}$$

↑
prior

Example #2:

A Mystery Cough

Two problems with the formulation above:

1: We're just adding highs and lows — is there a better formula?

2: We chose a single winner — is there a better selection technique?

solution:

Sampling (repeated guessing)

Recap: In Bayesian Statistics:

- Every Hypothesis must be given a **Prior** probability: $P(H)$
- Every Hypothesis implies a **Likelihood** for the data: $P(D | H)$
- **Bayes Theorem** provides the recipe for combining **Prior** and **Likelihood** information optimally
- Bayesian statistics allows you to probabilistically **sample answers** to an inferential question, **rather than chose one single winner**
- **advanced exercises in the notebook —>**

Further Reading / Viewing / **Coding:**



How do you do a Bayesian Mixed Effects GLM in R?

Check out Bradley Rentz's Excellent tutorial on RSTANARM

SLIDES:

<https://github.com/rentzb/bayesian-pape>

R MARKDOWN:

[https://github.com/rentzb/bayesian-pape/blob/master/
presentation/intro_rstanarm.Rmd](https://github.com/rentzb/bayesian-pape/blob/master/presentation/intro_rstanarm.Rmd)

Further Reading / Viewing / **Coding:**



How do you do a Bayesian inference in Python using PYMC3?

Check out the notebook:

Bayesian Inference with PYMC3 -- Inferring the Size - Pitch Relationship.ipynb

Further Reading / **Viewing** / Coding:

Looking for more intuitive explanations?

Check out:

A 10-minute intuitive video explainer of Bayes by Julia Galef (writer and founder of the Centre for Applied Rationality in New York):

A Visual Guide to Bayesian Thinking

https://www.youtube.com/watch?v=BrK7X_XIGB8

Further **Reading** / Viewing / Coding:

How is Bayesian inference used in the cognitive sciences?

Check out:

Tom Griffiths's Bayesian reading list:

<http://cocosci.berkeley.edu/tom/bayes.html>