# Session 1.1: Bayesian thinking

Bayesian modelling for Spatial and Spatio-temporal data, Imperial College

# Learning objectives

After this lecture you should be able to

- Introduce Bayesian way of thinking
- Compare Bayesian and Frequentist approaches in terms of parameters and probability definition

The topics treated in this lecture are presented in Chapter 3 of Blangiardo and Cameletti (2015) and in Chapter 1 of Johnson, Ott, and Dogucu (2022).

# Main resources

For the first part of this module you will be able to follow these two books

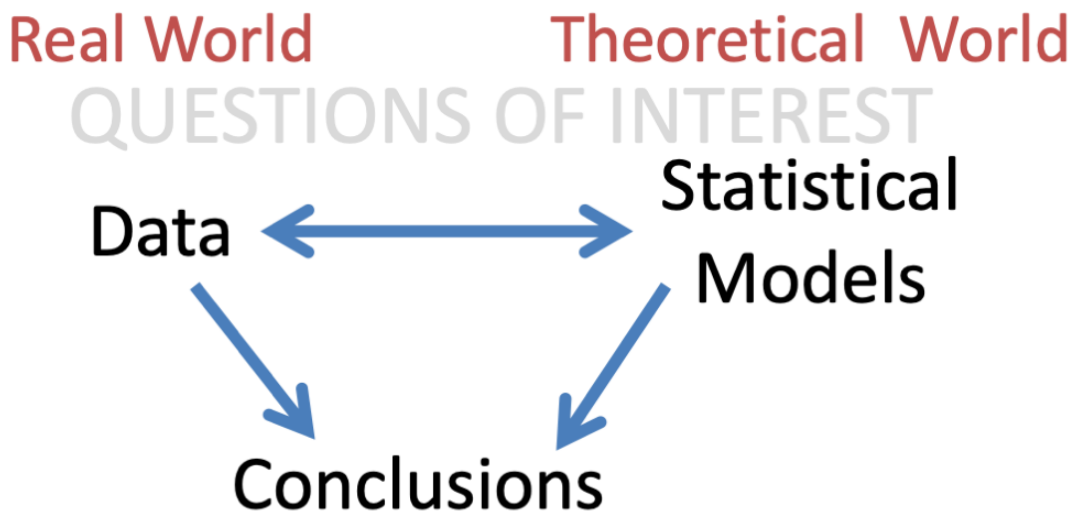
# Main resources

# Outline

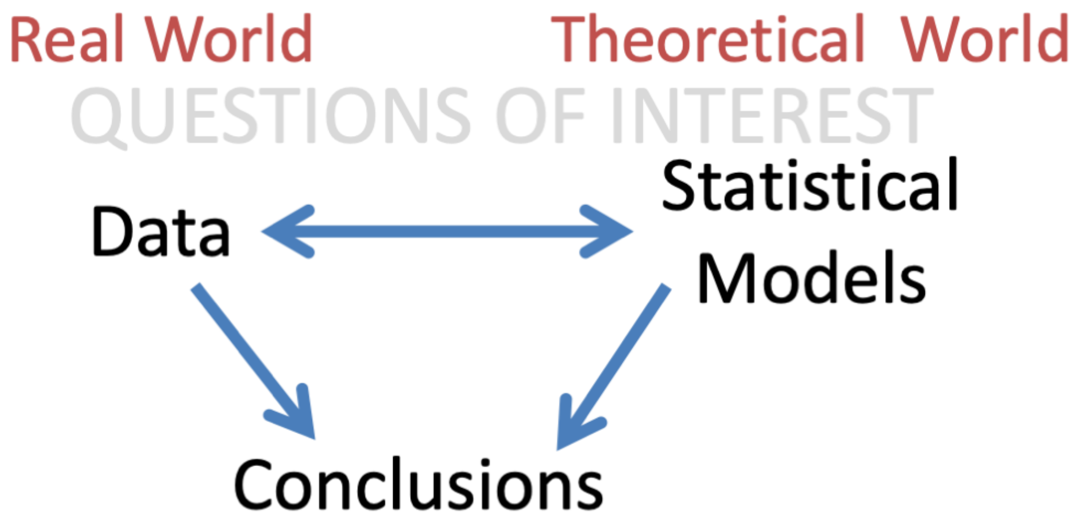
# Why Bayesian

## QUESTIONS OF INTEREST

# Statistics - The 'Big Picture'



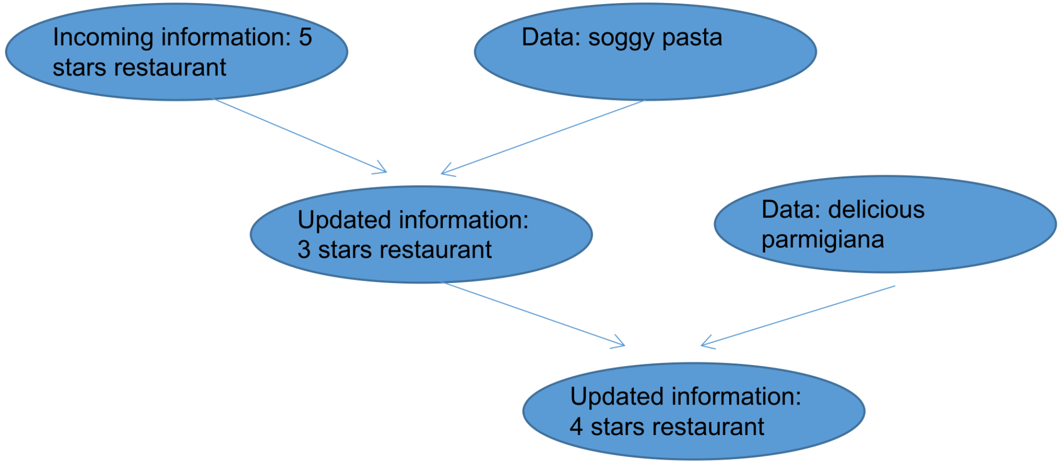
# Statistics - The 'Big Picture'




- Several different ways of formulating statistical models and computing inferences from these models and data
- Can be grouped into two broad approaches:
  – Frequentist
  – Bayesian

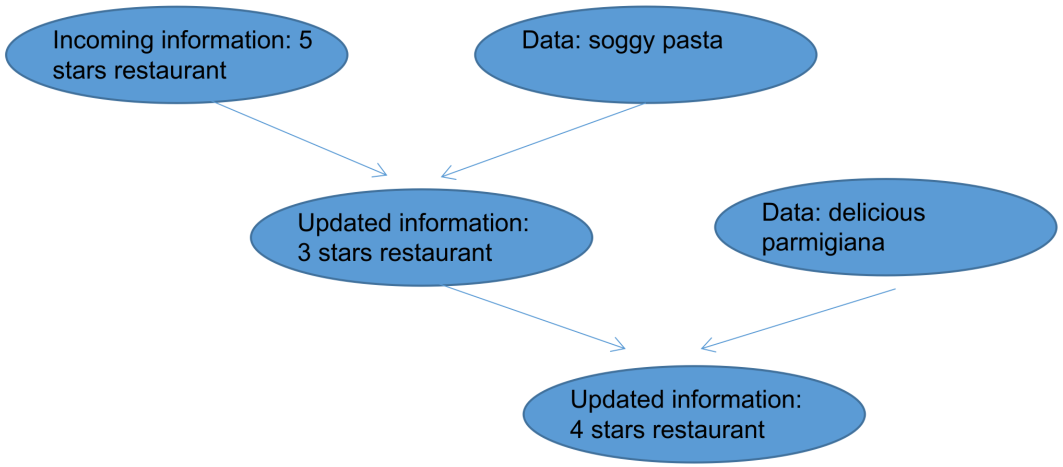
# Everyday thought process

- There's a new Italian restaurant in your neighborhood and on google reviews you see that it has 5 star (and of course you love Italian food).
- Thus, before going to the restaurant you believe that the food is going to be delicious.
- At your visit your pasta arrives in a soggy mess.
- So you update your belief and now conclude the restaurant is 3 stars rather than 5.
- You decide to try again and this time your aubergine parmigiana turns out to be exquisite.
- So you update your belief and now conclude the restaurant is 4 stars rather than 3.



# Everyday thought process

- There's a new Italian restaurant in your neighborhood and on google reviews you see that it has 5 star (and of course you love Italian food).
- Thus, before going to the restaurant you believe that the food is going to be delicious.
- At your visit your pasta arrives in a soggy mess.
- So you update your belief and now conclude the restaurant is 3 stars rather than 5.
- You decide to try again and this time your aubergine parmigiana turns out to be exquisite.
- So you update your belief and now conclude the restaurant is 4 stars rather than 3.



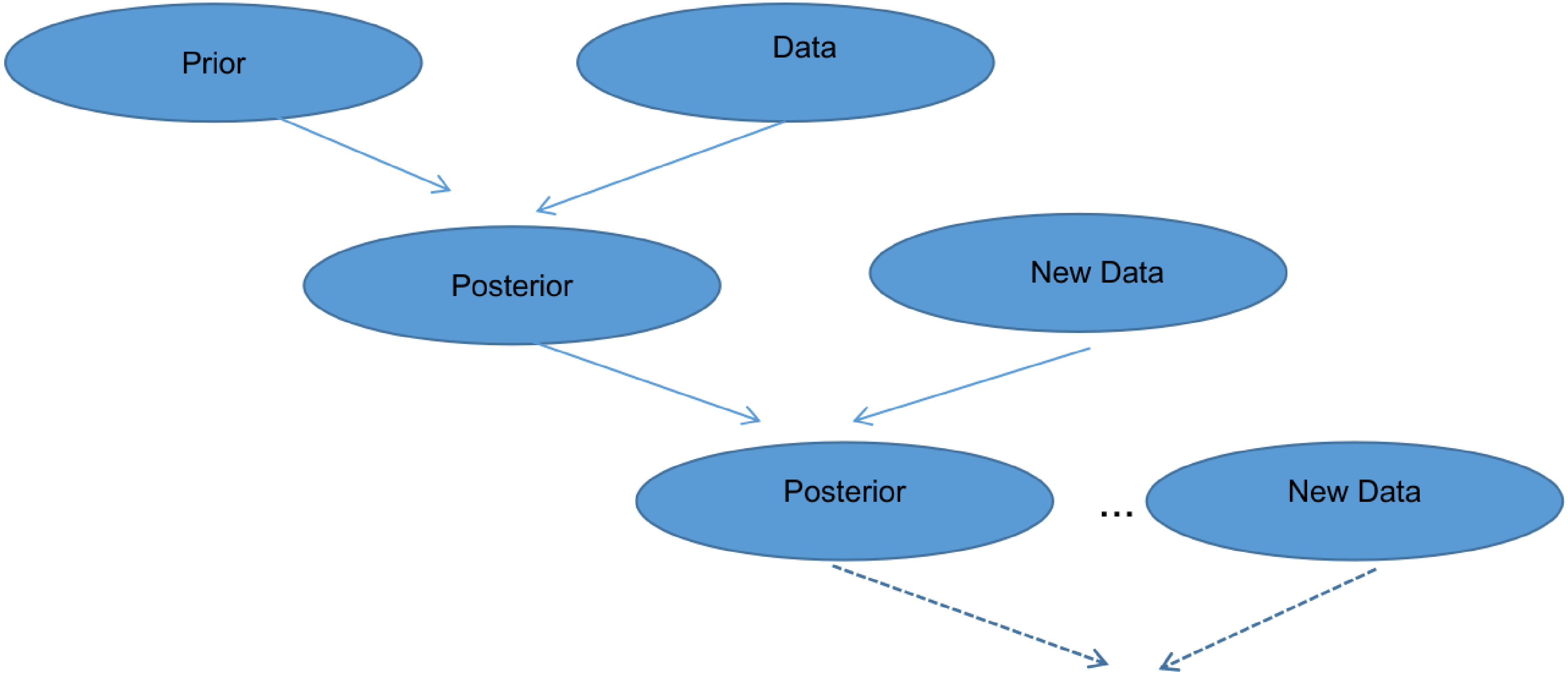

Natural Bayesian knowledge-building process of acknowledging your preconceptions, using data to update your knowledge, and repeating.

# Thought process of a physician

- A patient presents with a set of symptoms, concerned that they might have a certain disease
- The physician assesses the chance that the patient has this disease, based on
  - symptoms
  - family history
  - alternative explanations of symptoms
  - prevalence of disease
- The physician sends the patient for a diagnostic test
- The physician re-assesses the chance that the patient has this disease, taking account of
  - results of diagnostic test
  - reliability of diagnostic test
- The physician may send the patient for further diagnostic tests

**In pair: jot down the doodle for the physician knowledge-building process**

# Bayesian thinking




- This is the foundation of Bayesian statistics
- It can be applied to any research area

# Why Bayesian methods?

- Bayesian methods have been widely applied in many areas:
  - medicine / epidemiology
  - genetics
  - ecology
  - environmental sciences
  - social and political sciences
  - finance
  - archaeology
  - .....

- Motivations for adopting Bayesian approach vary:
  - natural and coherent way of thinking about science and learning
  - pragmatic choice that is suitable for the problem in hand

# Bayesian vs Frequentist

# Quiz time: are you a Frequentist or a Bayesian

1. When flipping a fair coin, we say that *the probability of flipping Heads is 0.5.* How do you interpret this probability?

a. If I flip this coin over and over, roughly 50% will be Heads.

b. Heads and Tails are equally plausible.

c. Both a and b make sense.

# Quiz time: are you a Frequentist or a Bayesian

1. When flipping a fair coin, we say that *the probability of flipping Heads is 0.5.* How do you interpret this probability?

a. If I flip this coin over and over, roughly 50% will be Heads.

b. Heads and Tails are equally plausible.

c. Both a and b make sense.

2. An election is coming up and a pollster claims that candidate A has a 0.9 probability of winning. How do you interpret this probability?

a. If we observe the election over and over, candidate A will win roughly 90% of the time.

b. Candidate A is much more likely to win than to lose.

c. The pollster's calculation is wrong. Candidate A will either win or lose, thus their probability of winning can only be 0 or 1.

# Quiz time: are you a Frequentist or a Bayesian

1. When flipping a fair coin, we say that *the probability of flipping Heads is 0.5*. How do you interpret this probability?

a. If I flip this coin over and over, roughly 50% will be Heads.

b. Heads and Tails are equally plausible.

c. Both a and b make sense.

2. An election is coming up and a pollster claims that candidate A has a 0.9 probability of winning. How do you interpret this probability?

a. If we observe the election over and over, candidate A will win roughly 90% of the time.

b. Candidate A is much more likely to win than to lose.

c. The pollster's calculation is wrong. Candidate A will either win or lose, thus their probability of winning can only be 0 or 1.

3. Consider two claims. (1) Zuofu claims that he can predict the outcome of a coin flip. To test his claim, you flip a fair coin 10 times and he correctly predicts all 10. (2) Kavya claims that she can distinguish natural and artificial sweeteners. To test her claim, you give her 10 sweetener samples and she correctly identifies each. In light of these experiments, what do you conclude?

a. You're more confident in Kavya's claim than Zuofu's claim.

b. The evidence supporting Zuofu's claim is just as strong as the evidence supporting Kavya's claim.

# Quiz time: are you a Frequentist or a Bayesian

1. When flipping a fair coin, we say that *the probability of flipping Heads is 0.5*. How do you interpret this probability?

a. If I flip this coin over and over, roughly 50% will be Heads.

b. Heads and Tails are equally plausible.

c. Both a and b make sense.

2. An election is coming up and a pollster claims that candidate A has a 0.9 probability of winning. How do you interpret this probability?

a. If we observe the election over and over, candidate A will win roughly 90% of the time.

b. Candidate A is much more likely to win than to lose.

c. The pollster's calculation is wrong. Candidate A will either win or lose, thus their probability of winning can only be 0 or 1.

3. Consider two claims. (1) Zuofu claims that he can predict the outcome of a coin flip. To test his claim, you flip a fair coin 10 times and he correctly predicts all 10. (2) Kavya claims that she can distinguish natural and artificial sweeteners. To test her claim, you give her 10 sweetener samples and she correctly identifies each. In light of these experiments, what do you conclude?

a. You're more confident in Kavya's claim than Zuofu's claim.

b. The evidence supporting Zuofu's claim is just as strong as the evidence supporting Kavya's claim.

4. Suppose that during a recent doctor's visit, you tested positive for a very rare disease. If you only get to ask the doctor one question, which would it be?

a. What's the chance that I actually have the disease?

b. If in fact I don't have the disease, what's the chance that I would've gotten this positive test result?

# Question 1

- There is a crucial difference in how Frequentist and Bayesian interpret probability.

In the Bayesian philosophy, a probability measures **the relative plausibility** of an event.

The frequentist philosophy is so named for its interpretation of probability as the **long-run relative frequency** of a **repeatable** event.

So for question 1

- a Bayesian would conclude that Heads and Tails are equally likely
- a Frequentist would conclude that if we flip the coin over and over and over, roughly 1/2 of these flips will be Heads

# Question 2

- Since the election is a one-time event, the long-run relative frequency concept of observing the election over and over does not work
- In a very strict sense a Frequentist might say that the pollster is wrong. Since the candidate will either win or lose, their win probability must be either 1 or 0.
- A less extreme Frequentist interpretation, though a bit awkward, is more reasonable: in long-run hypothetical repetitions of the election, i.e., elections with similar circumstances, candidate A would win roughly 90% of the time.
- A Bayesian would not be bound by the concept of long-run relative frequency and would simply acknowledge that the candidate A has a higher probability of winning the election.

# Question 3

- There is a crucial difference in how Frequentist and Bayesian consider information.

In the Bayesian philosophy, we have two sources of information: the **prior** and the **data**; they are combined into the **posterior**.

The Frequentist philosophy based all the inference on data alone.

- Zuofu and Kavya's claims are equally likely in the Frequentist approach, despite Zuofu's being quite ridicolous

- A Bayesian analysis gives voice to our prior knowledge.
  - Here, our experience on Earth suggests that Zuofu is probably overstating his abilities but that Kavya’s claim is reasonable.
  - Since the data is consistent with our prior, we’re even more certain that Kavya is a sweetener savant. However, given its inconsistency with our prior experience, we are chalking Zuofu’s psychic achievement up to simple luck.

# Question 4

- Though the answers to both questions would be helpful, having to choose I would go for answer to (a)
- Interestignly, only a Bayesian analysis would be able to answer (a), as it **assesses the uncertainty of the hypothesis in light of the observed data**.
- A Frequentist analysis **assesses the uncertainty of the observed data in light of an assumed hypothesis**.

A Bayesian hypothesis test seeks to answer: In light of the observed data, what's the chance that the hypothesis is correct?

A frequentist hypothesis test seeks to answer: If in fact the hypothesis is incorrect, what's the chance I'd have observed this, or even more extreme, data?

**For an other example of an analysis done using the Bayesian and the Frequentist approach see recording 1**

# Components of a Bayesian analysis

# A Bayesian analysis

A clinical trial is carried out to collect evidence about an unknown *treatment effect*

## Conventional analysis

- p-value for $H_0$: treatment effect is zero
- Point estimate and CI as summaries of size of treatment effect

**Aim is to learn what this trial tells us about the treatment effect**

# A Bayesian analysis

A clinical trial is carried out to collect evidence about an unknown *treatment effect*

Conventional analysis

- p-value for $H_0$: treatment effect is zero
- Point estimate and CI as summaries of size of treatment effect

**Aim is to learn what this trial tells us about the treatment effect**

Bayesian analysis

- Inference is based on probability statements summarising the posterior distribution of the treatment effect

Asks: **how should this trial change our opinion about the treatment effect?**

# Components of a Bayesian analysis

A clinical trial is carried out to collect evidence about an unknown *treatment effect*. The Bayesian analyst needs to explicitly state

- a reasonable opinion concerning the plausibility of different values of the treatment effect *excluding* the evidence from the trial (the prior distribution)
- the support for different values of the treatment effect based *solely* on data from the trial (the likelihood),

and to combine these two sources to produce

- a final opinion about the treatment effect (the posterior distribution)

# Components of a Bayesian analysis

A clinical trial is carried out to collect evidence about an unknown *treatment effect*. The Bayesian analyst needs to explicitly state

- a reasonable opinion concerning the plausibility of different values of the treatment effect *excluding* the evidence from the trial (the prior distribution)
- the support for different values of the treatment effect based *solely* on data from the trial (the likelihood),

and to combine these two sources to produce

- a final opinion about the treatment effect (the posterior distribution)

The final combination is done using **Bayes theorem** (and only simple rules of probability), which essentially weights the likelihood from the trial with the relative plausibilities defined by the prior distribution

One can view the Bayesian approach as a formalisation of the process of learning from experience

# Bayesian inference: the posterior distribution

Posterior distribution forms basis for all inference --- can be summarised to provide

- point and interval estimates of Quantities of Interest (QOI), e.g. treatment effect, small area estimates, ...
- point and interval estimates of any function of the parameters
- probability that QOI (e.g. treatment effect) exceeds a critical threshold
- prediction of QOI in a new unit
- prior information for future experiments, trials, surveys, ...
- inputs for decision making
- ...

# References

Blangiardo, M. et al. (2015). *Spatial and spatio-temporal Bayesian models with R-INLA*. John Wiley & Sons.

Johnson, A. A. et al. (2022). *Bayes Rules!: An Introduction to Applied Bayesian Modeling*. CRC Press.