Think Bayesian

Giorgio Corani - SUPSI

Bayesian Data analysis and Probabilistic Programming

Credits

- Chap. 1 of *Bayes Rules! An Introduction to Applied Bayesian Modeling*
  - https://www.bayesrulesbook.com/chapter-1.html

The big picture

- We continuously update our knowledge about the world as we accumulate experience or collect data.
- The Bayesian approach rigorously models the knowledge-building process, in which you update your knowledge on the basis of new data.
- Knowledge (or ignorance) is expressed by means of a probability distribution on the variables of interest.

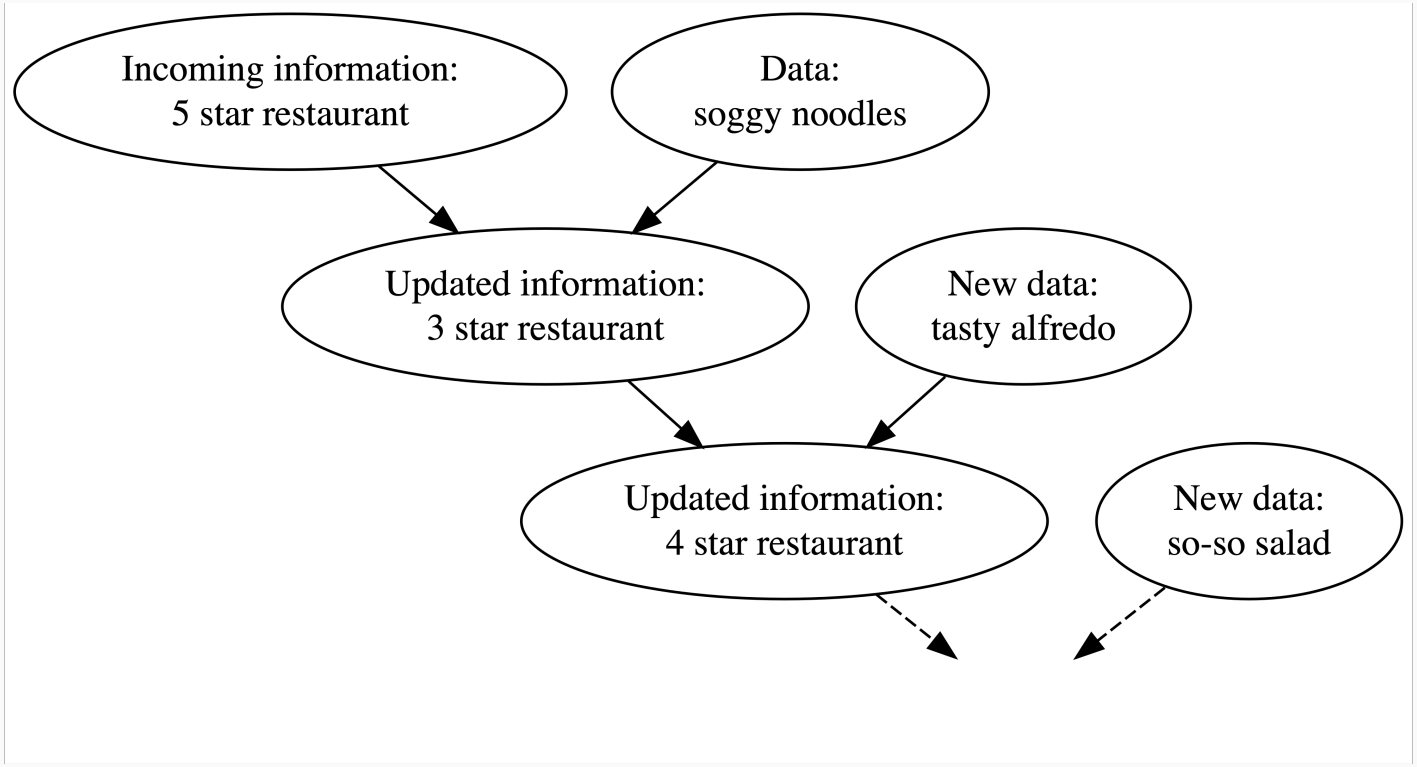
Changing opinion based on experience

- Suppose there's a new restaurant with 5-star online rating.
- Prior to enter the restaurant, you expect that it will be delicious.
- On your first visit, the pasta is poorly cooked.
- You weigh the high online rating against your poor meal (which might have just been a fluke), and you update your opinion knowledge: this is a medium, not an excellent restaurant.

Changing opinion based on experience

- In your second meal at the restaurant you're pleased with your dinner and increase your personal restaurant's rating to good.
- You continue to visit the restaurant, collecting edible data and updating your knowledge each time.
- After enough visits, you have your own informed opinion.

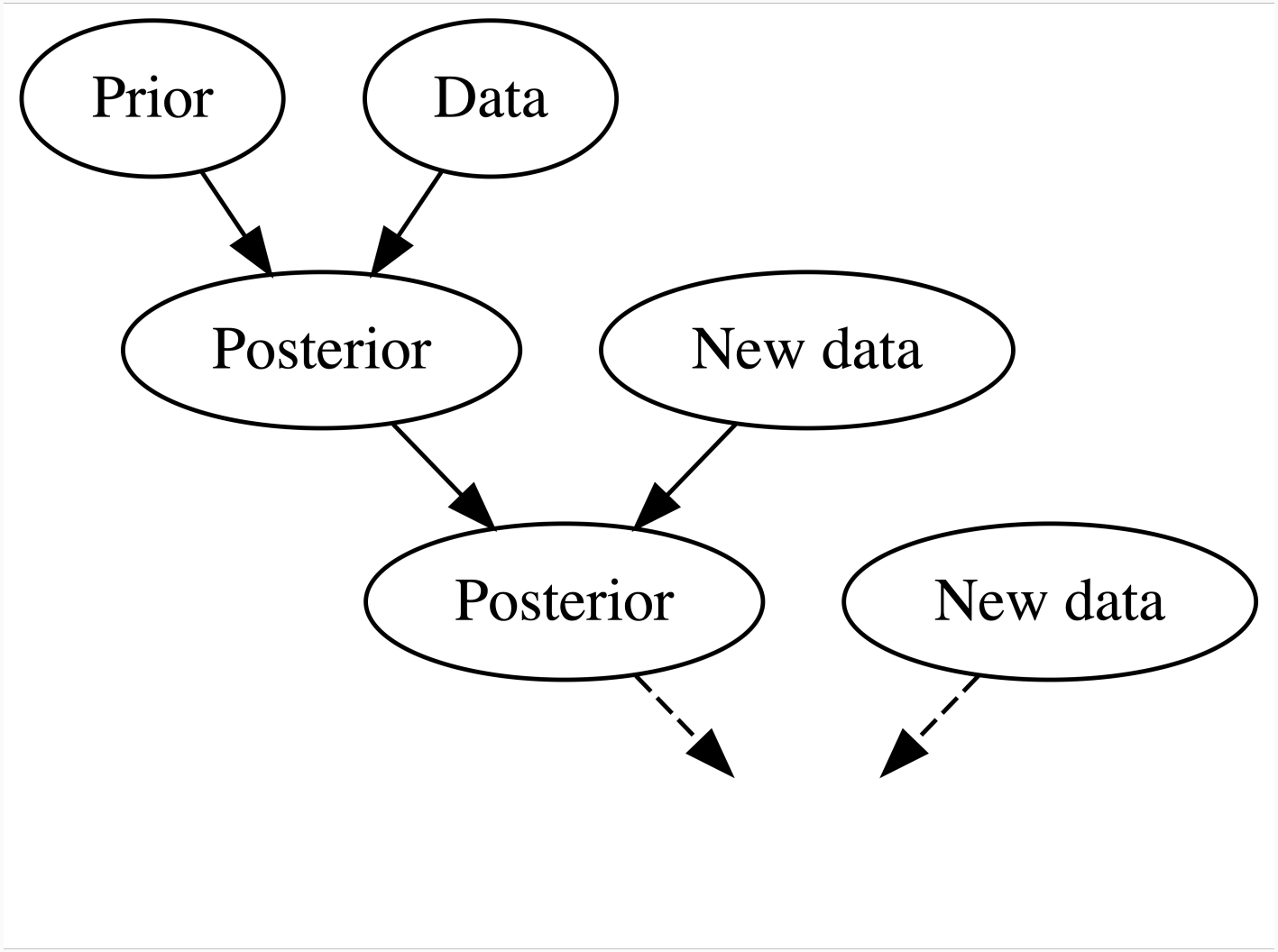
Building knowledge



Building knowledge

- An environmental scientist analyses of the human role in climate change.
- He carries a degree of incoming or prior information based on previous research and experience.
- In light of this information he interprets new data, considering both to develop an updated understanding (posterior information).
- He continues to refine this information as he gathers new evidence.

Building knowledge as a Bayesian process



Your turn

- Think of a recent situation in which you changed your mind.
- As with the Italian restaurant example, make a diagram that includes your prior information, your new data that helped you change your mind, and your posterior conclusion.

Bayesian interpretation of probability

- The probability of a coin landing Heads is 0.5.
  - How do you interpret this probability?
- If I flip this coin over and over, roughly 50% will be Heads (*frequentist* interpretation).
- Heads and Tails are equally plausible in the next flip (*Bayesian* interpretation).

Bayesian vs frequentist probability

- According to both Bayesian and frequentist probability, the probability of observing Heads on a fair coin flip is 1/2. The difference is in their interpretation:
- Bayesian: probability measures the relative plausibility of an event.
- The frequentist interpretation of probability is the long-run relative frequency of a repeatable event.

Bayesian vs frequentist intepretation

- 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?
  - If we observe the election over and over, candidate A will win roughly 90% of the time.
  - Candidate A is much more likely to win than to lose.

Asking questions

- Imagine that you tested positive for a disease; you can make a single question to the doctor.

- 1) what's the probability that I have the disease?

- 2) if I do not have the disease, what's the chance that I would've gotten the positive result?

Asking questions

- Bayesian analysis would answer the first question, by assessing the uncertainty of the hypothesis, in light of the data.
- A frequentist analysis would answer the second question, by assesssin the uncertainty of the data in light of an assumed hypothesis.

Probability of having the disease, given the positive test

	test pos	test neg	total
disease	3	1	4
no disease	9	87	96
total	12	88	100

- Only 3 of the 12 people that tested positive have the disease; there's only a 25% chance that you have the disease.
- This is the Bayesian answer: it measures the uncertainty of an hypothesis, given the observed data.

Probability of testing positive, given that I do not have the disease

	test pos	test neg	total
disease	3	1	4
no disease	9	87	96
total	12	88	100

- From the frequentist standpoint, since disease status isn't repeatable, the probability you have the disease is either 1 or 0 – you have it or you don't.
- Medical testing is repeatable. You can get tested for the disease over and over and over.
- Thus, a frequentist analysis asks: If I don't actually have the disease, what's the chance that I would've tested positive?
- Since only 9 of the 96 people without the disease tested positive,