

Bayesian Inference

Emily Jane McTavish

Life and Environmental Sciences
University of California, Merced

`ejmctavish@ucmerced.edu`, `twitter:snacktavish`

(With thanks to Mark Holder, and Paul Lewis for slides!)

Bayesian statistics

- ▶ Suppose that you are worried that you might have a rare disease.

- ▶ Suppose that you are worried that you might have a rare disease.
- ▶ You decide to get tested.

- ▶ Suppose that you are worried that you might have a rare disease.
- ▶ You decide to get tested.
- ▶ Test is correct 99 percent of the time. (in other words, if you have the disease, it shows that you do with 99 percent probability, and if you don't have the disease, it shows that you do not with 99 percent probability).

- ▶ Suppose that you are worried that you might have a rare disease.
- ▶ You decide to get tested.
- ▶ Test is correct 99 percent of the time. (in other words, if you have the disease, it shows that you do with 99 percent probability, and if you don't have the disease, it shows that you do not with 99 percent probability).
- ▶ This disease is actually quite rare, occurring randomly in the general population in only one of every 10,000 people.

If your test results come back positive, what are your chances that you actually have the disease?

Example from Su, Francis E., et al. "Medical Tests and Bayes' Theorem." Math Fun Facts.
<<http://www.math.hmc.edu/funfacts>>.

If your test results come back positive, what are your chances that you actually have the disease?

- ▶ .99
- ▶ .90
- ▶ .10
- ▶ .01

If your test results come back positive, what are your chances that you actually have the disease?

- ▶ .99
- ▶ .90
- ▶ .10
- ▶ .01

Congratulations, there is less than 1 percent chance that you have the disease!

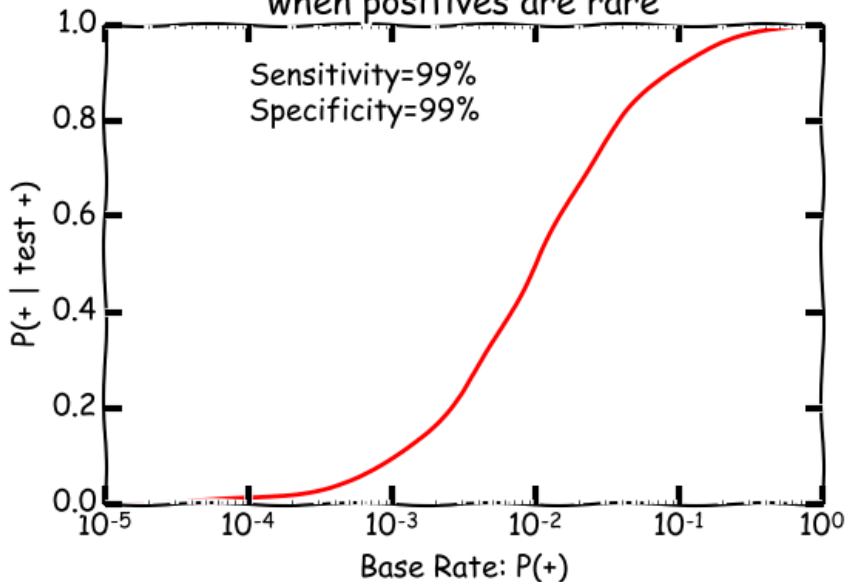
We can think about this problem in terms of conditional probabilities.

We can think about this problem in terms of conditional probabilities.

- ▶ Event A is the event you have this disease, and event B is the event that you test positive.
- ▶ $P(B|notA)$ is the probability of a "false positive": that you test positive even though you don't have the disease.
- ▶ In this case: $P(B|A) = .99$
- ▶ but $P(A) = .0001$.
- ▶ $P(B)$ may be derived by conditioning on whether event A does or does not occur:
- ▶ $P(B) = P(B|A)P(A) + P(B|notA)P(notA)$
- ▶ or $.99*.0001 + .01*.9999$.

Thus the probability that you have the disease is less than 1 percent.

Even a **very good** test
will be wrong most of the time
when positives are rare





Covid-19 lateral flow tests

Calculator for interpreting test results

This calculator demonstrates how interpreting a covid-19 lateral flow device (LFD) result varies according to the pre-test probability, and the sensitivity and specificity of the LFD used. LFDs are best for identifying people with active infection. We recommend choosing values for test performance and pre-test probability that reflect somebody having covid-19 and being infectious (see article below).

Adjust these values to update graphic below

Pre-test probability

Likelihood of having covid-19

0.10

%

Test sensitivity

The proportion of patients with covid-19 who have a positive test

67.00

%

Test specificity

The proportion of patients without covid-19 who have a negative test

99.95

%

If 100 000 people were tested with these values

100 have covid-19

True positive



67 people who test positive have covid-19

99 900 do not have covid-19

False positive



50 people who test positive do not have covid-19

57.26%

Probability of having covid-19 if test is positive

False negative



33 people who test negative have covid-19

True negative



99 850 people who test negative do not have covid-19

0.03%

Probability of having covid-19 if test is negative

+

Lateral flow test positive

-

Lateral flow test negative

Disclaimer: This infographic is not a validated clinical decision aid. This information is provided without any representations, conditions, or warranties that it is accurate or up to date. BMJ and its licensors assume no responsibility for any aspect of treatment administered with the aid of this information. Any reliance placed on this information is strictly at the user's own risk. For the full disclaimer wording see BMJ's terms and conditions: <http://www.bmj.com/company/legal-information/>

© 2021 BMJ Publishing Group Ltd.

<https://www.bmj.com/content/373/bmj.n1411>

Bayes' theorem (aka Bayes' rule or Bayes law) helps us find the probability of event A given event B, $P(A|B)$, in terms of the probability of B given A, $P(B|A)$, and the probabilities of A and B:

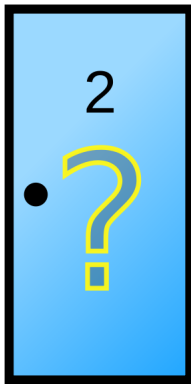
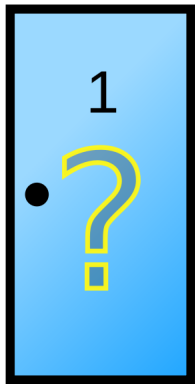
$$P(A|B) = \frac{P(A)P(B|A)}{P(B)}$$



Reverend Bayes

Monte Hall problem

Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No. 1, and the host, who knows what's behind the doors, opens another door, say No. 2, which has a goat. He then says to you, "Do you want to pick door No. 3?" Is it to your advantage to switch your choice?



Monte Hall problem

Exercise:

- ▶ Find a partner.
- ▶ Take 3 playing cards - two red (goats), and one black (car).
- ▶ Play the Monte Hall game. (take turns as the host and the contestant)
- ▶ Record if you got the car or the goat, and if you chose to switch or not. Run at least 10 games per tactic, and add your scores to the spreadsheet.
- ▶ Try out other variants! (what if the host doesn't know which door has the car? What if you have 3 goats and a car?)

Applying Bayes theorem to the Monty Hall problem:

Assume we pick Door 1 and then Monty shows us a goat behind Door 2.

Now let A be the event that the car is behind Door 1.

B be the event that Monty shows us a goat behind Door 2.

$$\Pr(A | B) = \frac{\Pr(B | A) \times \Pr(A)}{\Pr(B)}$$

<http://angrystatistician.blogspot.com/2012/06/>

[bayes-solution-to-monty-hall.html](http://angrystatistician.blogspot.com/2012/06/bayes-solution-to-monty-hall.html)

<https://brilliant.org/wiki/monty-hall-problem/>

The tricky calculation is $\Pr(B)$.

- ▶ We initially chose Door 1.
- ▶ if the car is behind Door 1, Monty will show us a goat behind Door 2 half the time, $1/3 \times 1/2$

The tricky calculation is $\Pr(B)$.

- ▶ We initially chose Door 1.
- ▶ if the car is behind Door 1, Monty will show us a goat behind Door 2 half the time, $1/3 \times 1/2$
- ▶ If the car is behind Door 2, Monty never shows us a goat behind Door 2. $1/3 \times 0$

The tricky calculation is $\Pr(B)$.

- ▶ We initially chose Door 1.
- ▶ if the car is behind Door 1, Monty will show us a goat behind Door 2 half the time, $1/3 \times 1/2$
- ▶ If the car is behind Door 2, Monty never shows us a goat behind Door 2. $1/3 \times 0$
- ▶ Finally, if the car is behind Door 3, Monty shows us a goat behind Door 2 every time. $1/3 \times 1$

Thus, $\Pr(B) = 1/3 \times 1/2 + 1/3 \times 0 + 1/3 \times 1 = 1/2$.

Applying Bayes theorem to the Monty Hall problem:

Assume we pick Door 1 and then Monty shows us a goat behind Door 2.

Now let A be the event that the car is behind Door 1.

B be the event that Monty shows us a goat behind Door 2.

$$\begin{aligned}\Pr(A | B) &= \frac{\Pr(B | A) \times \Pr(A)}{\Pr(B)} \\ &= \frac{1/2 \times 1/3}{1/3 \times 1/2 + 1/3 \times 0 + 1/3 \times 1} \\ &= 1/3.\end{aligned}$$

<http://angrystatistician.blogspot.com/2012/06/bayes-solution-to-monty-hall.html>

We saw a goat behind door 2. The car is either behind Door 1 or Door 3, since the probability that it is behind Door 1 is $1/3$ and the sum of the two probabilities must equal 1, the probability the car is behind Door 3 is $1 - 1/3 = 2/3$.

Key aspect of this problem:

The added information is not THAT Monty Hall revealed a goat, it is that he picked a door specifically because it had a goat.



An R script to simulate the Monte Hall problem by Corey Chivers,
(<https://bayesianbiologist.com/>)
[https://raw.githubusercontent.com/McTavishLab/GradPhylo/
master/docs/scripts/MonteHall.R](https://raw.githubusercontent.com/McTavishLab/GradPhylo/master/docs/scripts/MonteHall.R)

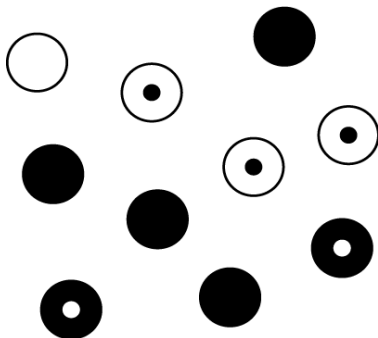
HW: What is the probability you get the car after switching if there are 10 doors? (1 car, 9 goats, everything else the same.)

Advantages of Bayesian inference:

- ▶ integrating prior information
- ▶ marginalizing inferences over parameter values

Joint probabilities

B = Black S = Solid
W = White D = Dotted



$$\begin{aligned}\Pr(B) &= 0.6 & \Pr(S) &= 0.5 \\ \Pr(W) &= 0.4 & \Pr(D) &= 0.5\end{aligned}$$

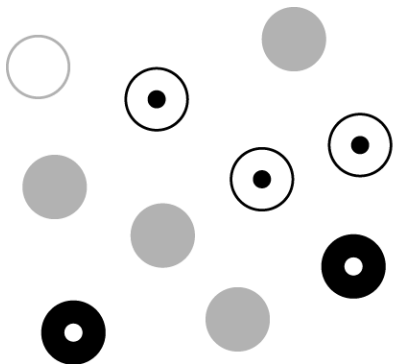
$$\Pr(\bullet\odot) = \Pr(B, D) = 0.2$$

$$\Pr(\bullet\bullet) = \Pr(B, S) = 0.4$$

$$\Pr(\odot\odot) = \Pr(W, D) = 0.3$$

$$\Pr(\odot\bullet) = \Pr(W, S) = 0.1$$

Conditional probabilities



$$\Pr(B|D) = \frac{2}{5} = 0.4$$

Hide all solid marbles
(leaving 5 with dot)

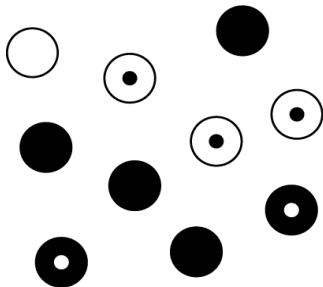
Of those left, 2 are black

Bayes' rule

$\Pr(B, D)$

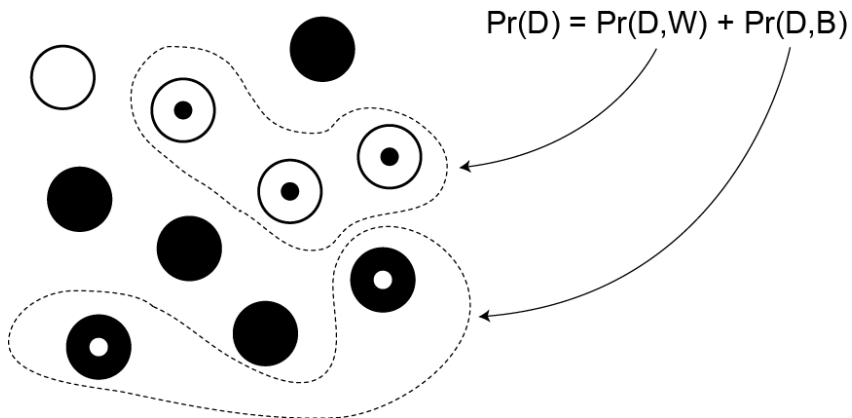
$$\Pr(D) \Pr(B|D) = \Pr(B) \Pr(D|B)$$

$$\frac{1}{2} \times \frac{2}{5} = \frac{3}{5} \times \frac{1}{3}$$



$$\begin{aligned} \Pr(B|D) &= \frac{\Pr(B) \Pr(D|B)}{\Pr(D)} \\ &= \frac{\frac{3}{5} \times \frac{1}{3}}{\frac{1}{2}} = \frac{2}{5} \end{aligned}$$

Probability of "Dotted"



Bayes' rule (cont.)

$$\begin{aligned}\Pr(B|D) &= \frac{\Pr(B) \Pr(D|B)}{\Pr(D)} \\ &= \frac{\Pr(D, B)}{\Pr(D, B) + \Pr(D, W)}\end{aligned}$$

$\Pr(D)$ is the **marginal probability** of being dotted
To compute it, we **marginalize over colors**

Bayes' rule (cont.)

It is easy to see that $\Pr(D)$ serves as a *normalization constant*, ensuring that $\Pr(B|D) + \Pr(W|D) = 1.0$

$$\Pr(B|D) = \frac{\Pr(D, B)}{\Pr(D, B) + \Pr(D, W)} \longleftarrow \Pr(D)$$

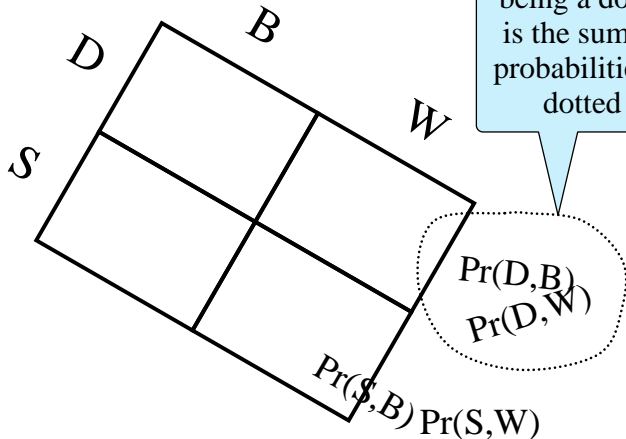
$$\Pr(W|D) = \frac{\Pr(D, W)}{\Pr(D, B) + \Pr(D, W)} \longleftarrow \Pr(D)$$

$$\Pr(B|D) + \Pr(W|D) = \frac{\cancel{\Pr(D, B)} + \cancel{\Pr(D, W)}}{\cancel{\Pr(D, B)} + \cancel{\Pr(D, W)}} = 1$$

Joint probabilities

	B	W
D	$\Pr(D,B)$	$\Pr(D,W)$
S	$\Pr(S,B)$	$\Pr(S,W)$

Marginalizing over colors



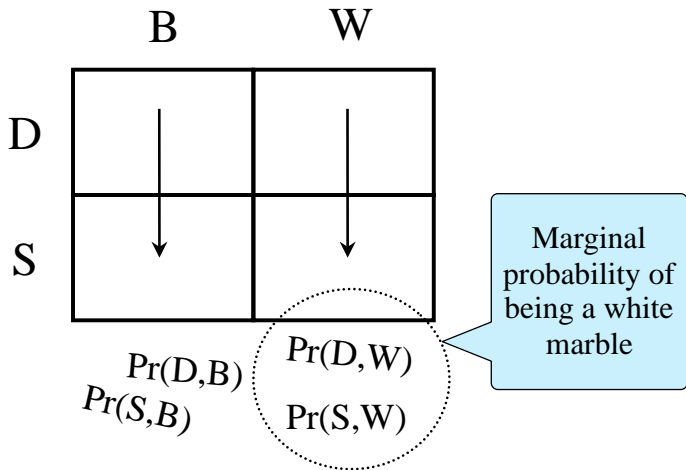
Marginal probabilities

	B	W	
D			$\Pr(D) = \text{marginal probability of being dotted}$
			$\Pr(D,B) + \Pr(D,W)$
S			
			$\Pr(S,B) + \Pr(S,W)$
			$\Pr(S) = \text{marginal probability of being solid}$

Joint probabilities

	B	W
D	$\Pr(D,B)$	$\Pr(D,W)$
S	$\Pr(S,B)$	$\Pr(S,W)$

Marginalizing over "dottedness"



Bayes' rule (cont.)

$$\begin{aligned}\Pr(B|D) &= \frac{\Pr(B) \Pr(D|B)}{\Pr(D, B) + \Pr(D, W)} \\&= \frac{\Pr(B) \Pr(D|B)}{\Pr(B) \Pr(D|B) + \Pr(W) \Pr(D|W)} \\&= \frac{\Pr(B) \Pr(D|B)}{\sum_{\theta \in \{B, W\}} \Pr(\theta) \Pr(D|\theta)}\end{aligned}$$

Bayes' rule in Statistics

$$\Pr(\theta|D) = \frac{\Pr(D|\theta) \Pr(\theta)}{\sum_{\theta} \Pr(D|\theta) \Pr(\theta)}$$

D refers to the "observables" (i.e. the **Data**)

θ refers to one or more "unobservables"

(i.e. **parameters** of a model, or the **model itself**):

- *tree model* (i.e. tree topology)
- *substitution model* (e.g. JC, F84, GTR, etc.)
- *parameter* of a substitution model (e.g. a branch length, a base frequency, transition/transversion rate ratio, etc.)
- *hypothesis* (i.e. a special case of a model)
- a *latent variable* (e.g. ancestral state)

Bayes' rule in statistics

The diagram illustrates Bayes' rule with the following components and annotations:

- Likelihood of hypothesis θ** : An arrow points from this text to the blue box containing $\Pr(D|\theta)$.
- Prior probability of hypothesis θ** : An arrow points from this text to the orange box containing $\Pr(\theta)$.
- Posterior probability of hypothesis θ** : An arrow points from this text to the purple box containing $\Pr(\theta|D)$.
- Marginal probability of the data (marginalizing over hypotheses)**: An arrow points from this text to the green box containing the denominator $\sum_{\theta} \Pr(D|\theta) \Pr(\theta)$.

$$\Pr(\theta|D) = \frac{\Pr(D|\theta) \Pr(\theta)}{\sum_{\theta} \Pr(D|\theta) \Pr(\theta)}$$

Next class we will discuss how to estimate posterior probabilities, and apply Bayesian statistics to phylogenetic data.