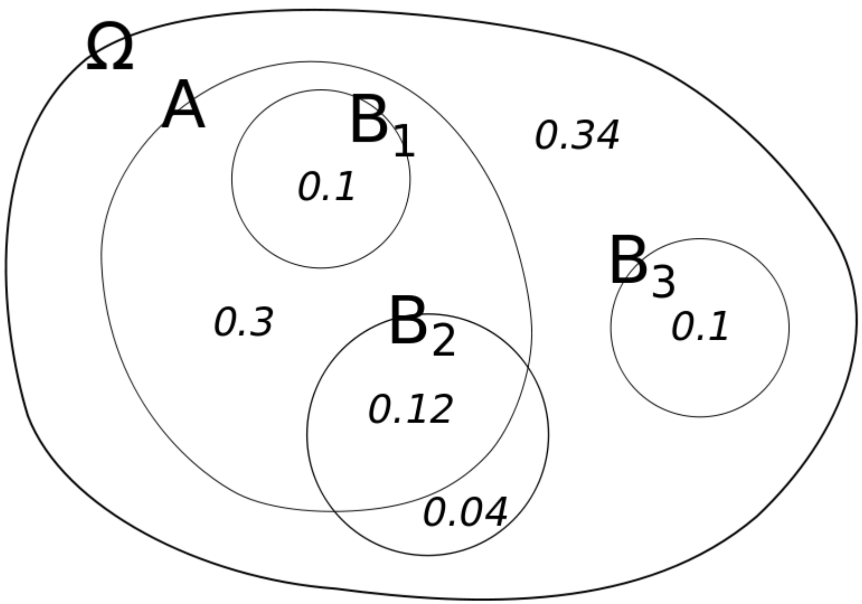
Introduction to Bayes' Law

Pamela Wu
IBB2014

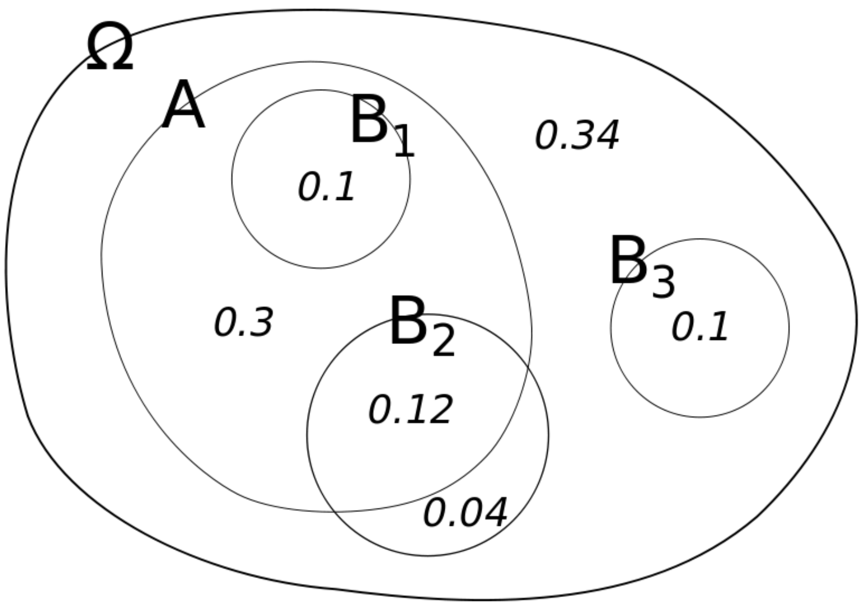
Conditional Probability



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

- If A and B are events, what is the fraction of B that is also A?
- Also written as: P(A|B)
- It is calculated by dividing the intersection (∩) of A and B by the probability of B.
- What is P(A|B1)?
- What is P(A|B2)?
- What is P(A|B3)?

Conditional Probability

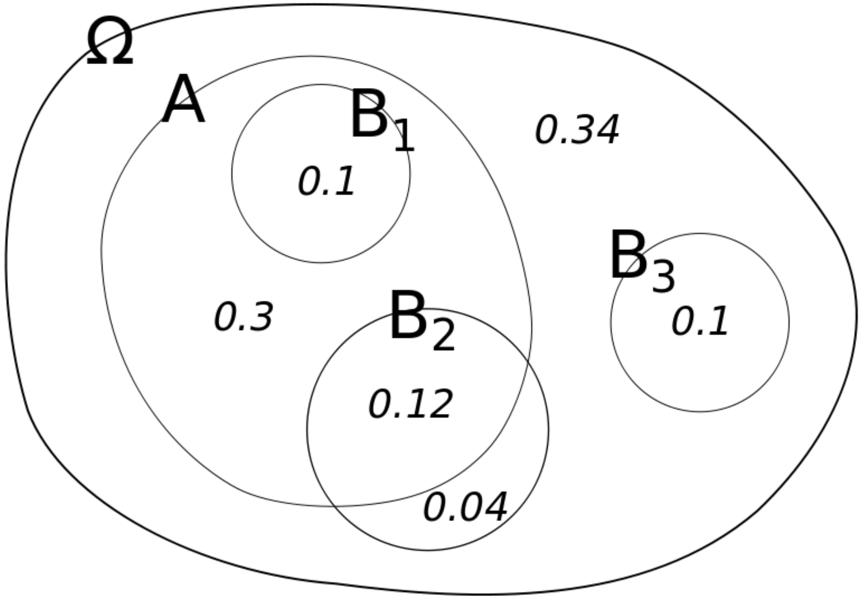


$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

- If A and B are events, what is the fraction of B that is also A?
- Also written as: P(A|B)
- It is calculated by dividing the intersection (∩) of A and B by the probability of B.
- What is P(A|B1)? 1.0
- What is P(A|B2)? 0.75
- What is P(A|B3)? 0.0

Conditional Probability

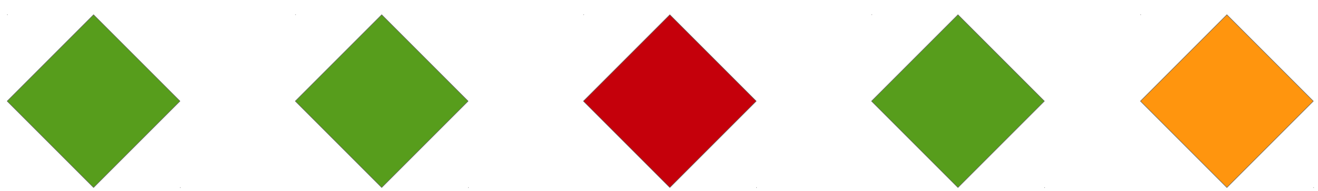
- Some properties of conditional probabilities:
  - $0 \leq P(A|B) \leq 1$
  - If $A \cap B$ is null, $P(A|B) = 0$
  - $P(A|A) = 1$



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

Frequentist vs. Bayesian approaches to statistics

- Frequentists think in terms of sample spaces
- Probability is a proportion of the total sample space
  - If I roll a dice, each roll is an element of the sample space and I expect 1/6 of the members of the sample space to be a 2.
  - In a population of frogs, each frog is an element of the sample space, and the number of green frogs divided by the total is the probability of finding a green frog.



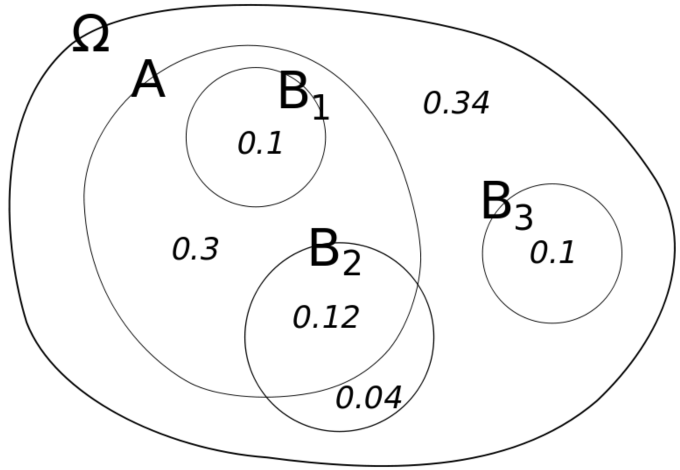
Frequentist vs. Bayesian approaches to statistics

- This way of thinking becomes problematic when you are talking about likelihoods, since all possibilities for an individual do not exist all at once to be counted up.
  - For a given email message, what is the likelihood that it is spam?
  - Given a person's family and medical history, what is the likelihood that she will get lung cancer?
- The problem is that all email messages and all potential cancer patients are (more or less) unique, so it doesn't make sense to lump them all together. For example, both are email messages:
  - Hey Pam, can you cat-sit Nigerian Prince? I OD'd on Viagra so I'm out of commish for a few days.
  - 50% off on storeyouveneverheardof.com! Act fast now with promo code IDGAF2014

Frequentist vs. Bayesian approaches to statistics

- If there were multiple universes, you could use the frequentist approach to findng out what is the likelihood of a person with a given medical history getting cancer because you would just count all of the versions of that person who get cancer compared to all that didn't get cancer.
- These approaches are not necessarily mutually exclusive. You can use the frequentist approach to calculate the evidence for the likelihoods for the Bayesian approach.

Derivation of Bayes' Law



$P(A \cap B) = P(A|B) \times P(B) = P(B|A) \times P(A)$

$P(A|B2) \times P(B2) = 0.12/(0.12+0.04) \times (0.12+0.04)$

$P(B2|A) \times P(A) = 0.12/(0.12+0.3+0.1) \times (0.12+0.3+0.1)$

$P(A|B2) \times P(B2) = P(B2|A) \times P(A) = 0.12$

$P(A|B) \times P(B) = P(B|A) \times P(A)$

Bayes' Law

P(A|B) = P(B|A) x P(A) / P(B)

- It is means of updating beliefs based on new evidence because what you observe (A) is thought to be conditionally dependent on various pieces of evidence (B).
- For example, if you see someone at a distance, they are 50/50 likely to be a woman, but if you notice the person has long hair, and women are more likely to have long hair, you can now say that it is more likely that you are seeing a woman in the distance.
- P(long hair|woman) is the proportion of women with long hair.

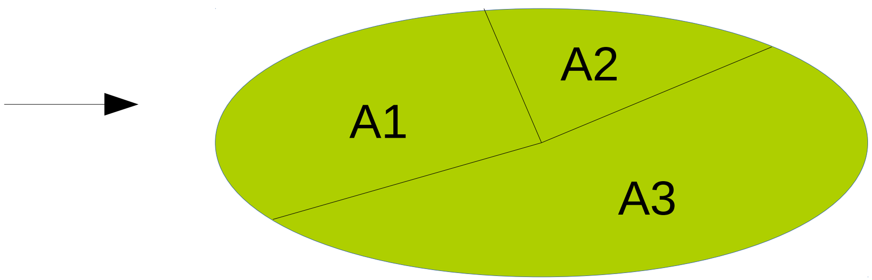
P(woman|long hair) = P(long hair|woman) x P(woman) / P(long hair)

Bayes' Law

P(A|B) = P(B|A) x P(A) / P(B)

- Sometimes you don't know the denominator directly P(B), so you need to calculate it from what you know. Think of the evidence B as divided up into the **mutually exclusive** outcomes that they predict (A1, A2, etc)

Green
Circle
Is B



$P(B) = P(B \cap A1) + P(B \cap A2) + P(B \cap A3)$

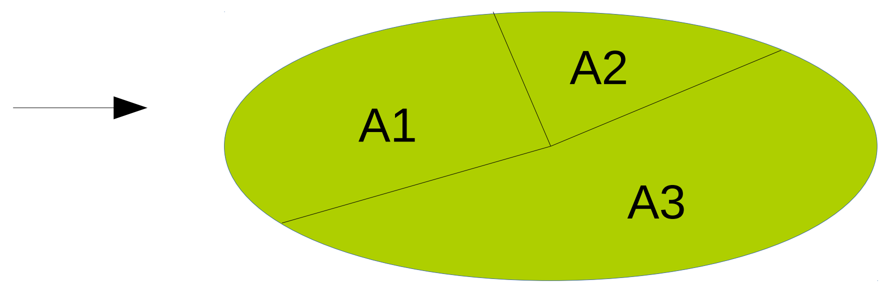
Since $P(B \cap A) = P(B|A) \times P(A)$:

$P(B) = P(B|A1) \times P(A1) + P(B|A2) \times P(A2) + P(B|A3) \times P(A3)$

Bayes' Law

P(A|B) = P(B|A) x P(A) / P(B)

Green
Circle
Is B

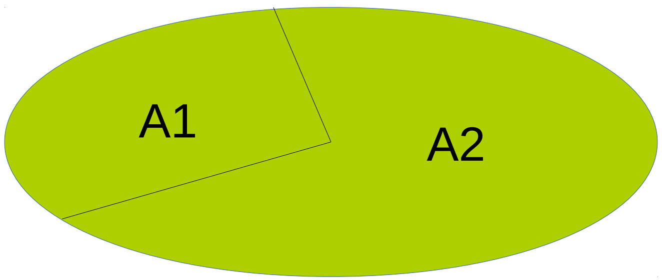


$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B|A1) \times P(A1) + P(B|A2) \times P(A2) + P(B|A3) \times P(A3)}$$

Bayes' Law

P(A|B) = P(B|A) x P(A) / P(B)

Green
Circle
Is B

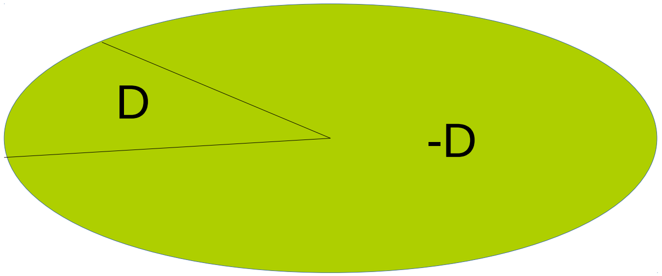


$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B|A1) \times P(A1) + P(B|A2) \times P(A2)}$$

Disease given test

$$P(D|+T) = \frac{P(+T|D) \times P(D)}{P(+T|D) \times P(D) + P(+T|-D) \times P(-D)}$$

Green
Circle
Is all tests



The disease affects 1 in 10,000 people
The test is 99% accurate

(If patient has disease, it will be positive 99/100 times
If the patient does not have the disease, it will be negative 99/100 times)

Fill out the equation above based on this information

Disease given test

$$P(D|+T) = \frac{P(+T|D) \times P(D)}{P(+T|D) \times P(D) + P(+T|-D) \times P(-D)}$$

The disease affects 1 in 10,000 people
The test is 99% accurate

(If patient has disease, it will be positive 99/100 times
If the patient does not have the disease, it will be negative 99/100 times)

$P(D) = 1/10000 = 0.0001$
$P(+T|D) = 99/100 = 0.99$
$P(+T|-D) = 1 – 0.99 = 0.01$
$P(-D) = 1 – 0.0001 = 0.9999$

Therefore,
$P(D|+T) = 0.0098$

Out of 101 patients with a positive test, only one has the disease