Rules of Probability

Bayesian Modeling for Socio-Environmental Data

Chris Che-Castaldo, Mary B. Collins, N. Thompson Hobbs

August 2017

Road map

- Rules of probability
  - Conditional probability
  - Independence
  - The law of total probability
- Factoring joint probabilities

Why do we need to know this stuff?

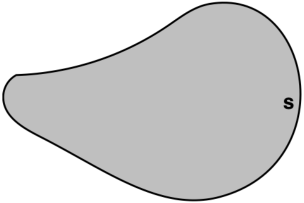
1. **Conditional probability** foundational for all the inferences that we make.
2. **The law of total probability** is the denominator of Bayes' Theorem.
3. **Factoring** joint distributions is how we deal with complexity, reducing high dimensional problems.
4. **Independence** allows us to simplify fully factored joint distributions.

Random variables

- are quantities governed by chance.
- have a specific value called an *event* or *outcome*.
- are summarized by probability distributions.
- *Bayesians treat every unobserved quantity as a random variable*.

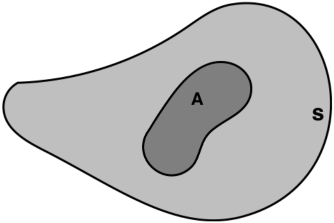
S=Sample Space

- The set of all possible events or outcomes of an experiment or survey.
- The sample space, S has a specific area.

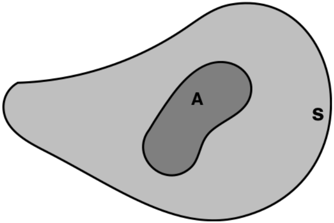


Events in S

- Can define and event, A.
- The area of event A is less than S.



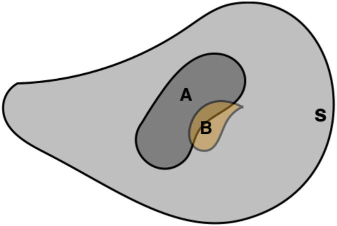
What is the probability of event A?



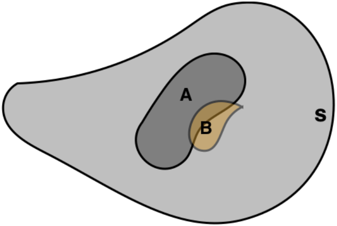
$\Pr(A) = \frac{\text{Area of } A}{\text{Area of } S}$

Conditional Probability

Conditional probability is a measure of the probability of an event given that another event has occurred.

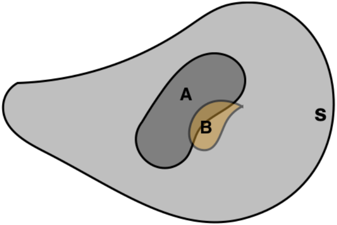


What is the probability of event A, given that event B has occurred?



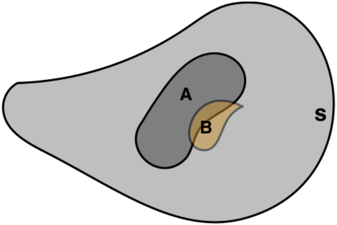
$\Pr(A|B) =$ prob' of A conditional on knowing B has occurred

What is the probability of event A, given that event B has occurred?



$$Pr(B|A) = \frac{\text{Joint Prob}}{\text{Prob of A}} = \frac{\Pr(A \cap B)}{\Pr(A)} = \frac{\Pr(A, B)}{\Pr(A)}$$

What is the probability of event A, given that event B has occurred?



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

We will make lots of use of the rearrangement of this equation

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

$$\Pr(A,B) = \Pr(A|B)\Pr(B) \text{ and equivalently,} \tag{1}$$

$$\Pr(A,B) = \Pr(B|A)\Pr(A) \tag{2}$$

Conditional probabilty

True or False?

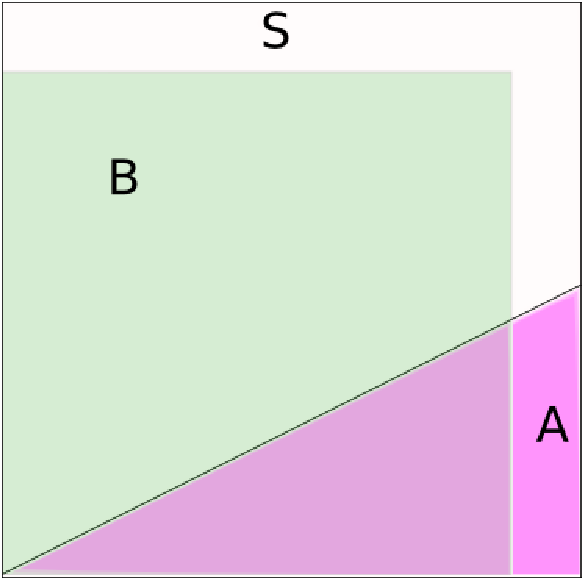
$\Pr(B|A) = \Pr(A|B)$

What happens if event A doesn't tell us anything about event B?

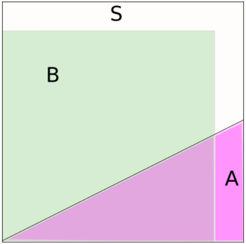
*In this case, events A and B are said to be **independent***

Events are independent if and only if...

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

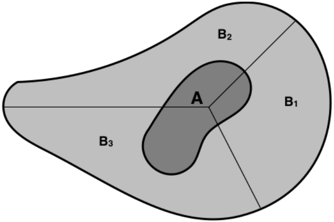

S
B
A

Assuming independence, the joint probability of event A and event B



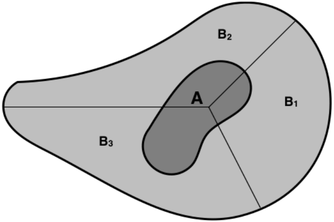
$$\Pr(A, B) = \Pr(A)\Pr(B)$$

The Law of Total Probability



We can define a set of events $\{B_n : n = 1, 2, 3, ...\}$, which taken together define the entire sample space, $\sum_n B_n = S$.

What is the probability of event A?



$\Pr(A) = \sum_n \Pr(A|B_n)\Pr(B_n)$ (discrete case)

$\Pr(A) = \int \Pr(A|B)\Pr(B)dB$ (continuous case)

Prob

Have a prob' where there is theta 1 through 4... write out the full expression

The Chain Rule

In probability theory, the chain rule (also called the general product rule) permits the calculation of any member of the joint distribution of a set of random variables using only conditional probabilities.

$$\Pr(z_1, z_2, ..., z_n) = \Pr(z_n|z_{n-1}, ..., z_1) ... \Pr(z_3|z_2, z_1) \Pr(z_2|z_1) \Pr(z_1)$$

Notice the pattern here.

- z's can be scalars or vectors.
- Sequence of conditioning doesn't matter.
- When we build models, we choose a sequence that makes sense.

Factoring joint probabilities

- The rules of probability allow us to take complicated joint distributions of random variables and break then down into chunks.
- Chunks can then be analyzed one at a time as if all other random variables were known and constant.
- Provide a usable graphical and mathematical foundation, which is *critical* for accomplishing the model specification step in the general modeling process.

Consider a Bayesian Network (represented by a directed acyclic graph or DAG)

- Bayesian networks specify how joint distributions are factored into conditional distributions using nodes to represent RV's and arrows to represent dependencies.
- Nodes at the heads of arrows *must* be on the left hand side of the conditioning symbols;
- Nodes at the tails of arrows are on the right hand side of the conditioning symbols.
- Any node at the tail of an arrow without and arrow leading into it must be expressed unconditionally.

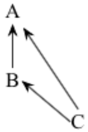
Factoring with DAGs

$\Pr(A, B) =$

Factoring with DAGs

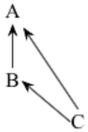
$\Pr(A, B) = \Pr(A|B)\Pr(B)$

Factoring with DAGs



$\Pr(A, B, C) =$

Factoring with DAGs



$$\Pr(A, B, C) = \Pr(A|B, C)\Pr(B|C)\Pr(C)$$

Generalizing

$$\Pr(z_1, ..., z_n) = \prod_{i=1}^{n} \Pr(z_i | \{P_i\})$$

$\{P_i\}$ is the set of parents of node z_i

Work on lab

Complete parts I-VI