Probability Overview

Many of these slides are derived from Seyong Kim, Tom Mitchell, William Cohen, Eric Xing. Thanks!

Random Variables

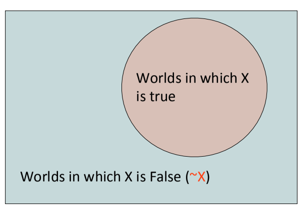
- Informally, X is a random variable if
  - X denotes something about which we are uncertain
  - perhaps the outcome of a randomized experiment
    - e.g. rolling a die
- Examples
  - X = True if a randomly drawn person from our class is female
    - binary
  - X = The hometown of a randomly drawn person from our class
    - multivalued
  - X = True if two randomly drawn persons from our class have same birthday
    - binary

Random Variables

- Define P(X) as “the fraction of possible worlds in which X is true” or “the fraction of times X holds, in repeated runs of the random experiment”
  - the set of possible worlds is called the sample space, S

Blue Rectangle: Sample space of all possible worlds (S)

Area = 1 (all possible things)



P(X) = Area of reddish oval
$0 < P(X) < 1$

A little formalism

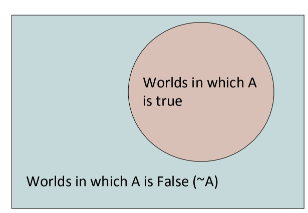
More formally, we have

- a sample space S (e.g., set of students in our class)
  - aka the set of possible worlds
- a random variable is a function defined over the sample space
  - Gender: S → { m, f } (binary, discrete)
  - Height: S → Real numbers (continuous)
- an event is a subset of S
  - e.g., the subset of S for which Gender=f
  - e.g., the subset of S for which (Gender=m) AND (eyeColor=blue)
- We are often interested in probabilities of specific events and of specific events conditioned on other specific events

The Axioms of Probability

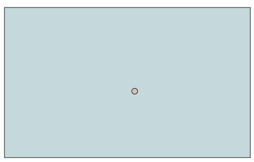
- Assume binary random variables A and B.
  - 0 <= P(A) <= 1
  - P(True) = 1
  - P(False) = 0
  - P(A or B) = P(A) + P(B) – P(A and B)

Visualizing Probability Axioms



Interpreting the axioms

- P(A) = 0



The area of A can't get any smaller than 0

And a zero area would mean no world could ever have A true

P(False) = 0

Interpreting the axioms

- P(A) = 1

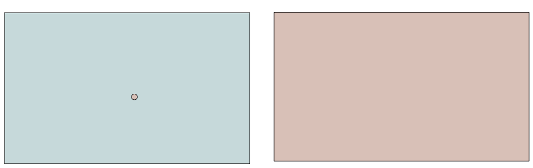
The area of A can't get any bigger than 1

And an area of 1 would mean all worlds will have A true

P(True) = 1

Interpreting the Axioms

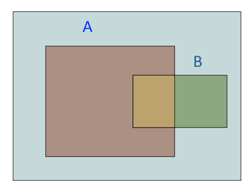
- 0 <= P(A) <= 1



Interpreting the axioms

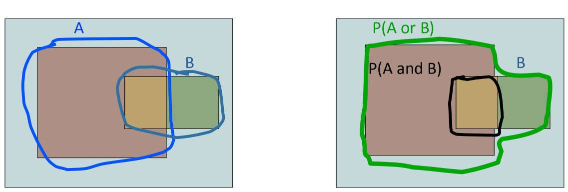
- P(A or B) = P(A) + P(B)

[WRONG! but why?]



Interpreting the axioms

- P(A or B) = P(A) + P(B) - P(A and B)



Simple addition and subtraction

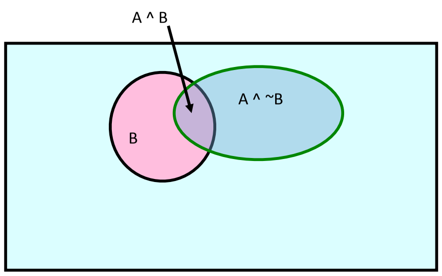
Another useful theorem

- 0 <= P(A) <= 1, P(True) = 1, P(False) = 0,
- P(A or B) = P(A) + P(B) - P(A and B)

➔ P(A) = P(A ^ B) + P(A ^ ~B)

Elementary Probability in Pictures

- P(A) = P(A ^ B) + P(A ^ ~B)



- P(A or B) = P(A ^ B) + P(A ^ ~B) + P(~A ^ B)

Interpreting the axioms

- 0 <= P(A) <= 1
- P(True) = 1
- P(False) = 0
- P(A or B) = P(A) + P(B) - P(A and B)

<u>Monotonicity</u>: if A is a subset of B, then P(A) <= P(B)

Extending the Axiom

- P(A or B or C) = ?