January 31, 2018

CS 361: Probability & Statistics

Probability

Probability theory

Probability

- Reasoning about uncertain situations with formal models
- Allows us to compute probabilities
- Experiments will be our data generating process

Outcomes

- If we toss a fair coin a bunch of times we expect about the same number of heads and tails
- If we roll a die, we don't expect to see one number more often than any other
- We can formally state the set of **outcomes** (heads, tails) we expect from an **experiment** (flipping the coin)

Outcomes

- Tossing a fair coin once: {H, T}
- Tossing a die: {1, 2, 3, 4, 5, 6}
- Tossing two coins: {HH, HT, TH, TT}

Sample space

- The **sample space** is the set of all possible outcomes of an experiment, written Ω

Example

- Three playing cards: King (K), Queen (Q), Knight (N). One is turned over randomly
- What is the sample space?
- {K, Q, N}

Example

- Suppose we flip a card, turn it back over, rearrange the cards and flip another. What is the sample space?
- {KK, KQ, KN, QQ, QK, QN, NN, NK, NQ}

Example

- A couple decides to have children until they have both a boy and a girl or until they have three children
- What is the sample space?
- {BG, GB, BBG, GGB, BBB, GGG}

Example: Monty Hall

- There are three doors: door #1, door #2, door #3. Behind two of them are goats, behind one is a car. The goats are indistinguishable
- If we open the doors and note what we observe the sample space is {CGG, GCG, GGC}

Monty hall #2

- Consider the Monty Hall scenario with distinguishable goats, one male and one female
- {CFM, CMF, FCM, FMC, MCF, MCF}
- Notice how there are more outcomes here

More family planning

- A couple decides to have children
- They decide to have children until a girl and then a boy is born
- What is the sample space here?
- The set of all strings that end in GB and contain no other GBs
- As a regular expression: B*G+B
- Somewhere between two and infinite children

Sample spaces

- Each run of an experiment has exactly one outcome
- The set of all possible outcomes is the sample space
- We will need to think of sample spaces to think rigorously about probability
- Sample spaces can be finite or infinite

Probability

- We might like to think of how often we will see each particular outcome *A* if we repeat an experiment over and over

$$\lim_{N \to \infty} \frac{\#\{A\}}{N}$$

- If our experiment is flipping a coin and we repeat it a large number of times
- We probably expect the relative frequencies of heads and tails to be non-negative and we want the frequency of either outcome to be at most 1
- Finally we want the relative frequencies of heads and tails to add up to 1

Probability

More generally, for an experiment with sample space Ω intuition tells us that we want each outcome A to satisfy

$$0 \leq \underbrace{P(A)}_{\text{probability of outcome A}} \leq 1$$

And we also want the following

$$\sum_{A_i \in \Omega} P(A_i) = 1$$

Example

- If we have a biased coin where $P(H) = 1/3$ and $P(T) = 2/3$ and we toss it three million times, how many times will we expect to see heads?
- We will see close to a million heads and 2 million tails

Example

- We often look at experiments where each outcome is equally likely
- In the example earlier with 3 cards, a King, Queen, and Knight where we turn one over at random. If each is equally likely, what probability should we assign to each card?

Example

- Recall the Monty Hall setup: 3 doors, 2 with a goat, and one with a car
- If we open the first door, what is the probability that we see a goat? What is the probability we see the car?
- $P(\text{car}) = 1/3$, $P(\text{goat}) = 2/3$

Example

- Recall the Monty Hall setup with distinguishable goats
- What is the probability we find a female goat behind door #1?
- P(female goat) = 1/3

Events

- Our formalization of probability will be about rules to assign probability to **sets** of outcomes
- For example, we might flip a coin three times
- Our sample space, the set of all outcomes, is {HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}
- "Getting two tails" might be something we are interested int he probability of and is the set of outcomes {HTT, THT, TTH}
- We give sets of outcomes a special name, **events**, and the theory we develop will concern the probabilities of events

Events

Remember: events are sets of outcomes

The set of all outcomes Ω is therefore an event and has probability 1

Likewise the empty set is an event (the set of no outcomes) and has probability 0

Events

Since events are sets, we want it to be the case that if A and B are events things like $A \cup B$,$A \cap B$,$A - B$, and $A^c = \Omega - A$ are also events

Example: Roll a die, let A={1,2,5} and B={2,4,6}

What are $A \cup B$, $A \cap B$,$A - B$, and $A^c = \Omega - A$ in this case?

Probability and events: axioms

Probabilities of events are non-negative: $P(A) \geq 0$

Every experiment has an outcome: $P(\Omega) = 1$

The probability of disjoint events is additive: if for all i and j we have $A_i \cap A_j = \emptyset$ then

$$P(\cup_i A_i) = \sum_i P(A_i)$$

Any function P that maps sets to real numbers and satisfies these is called a probability function or probability

Some properties

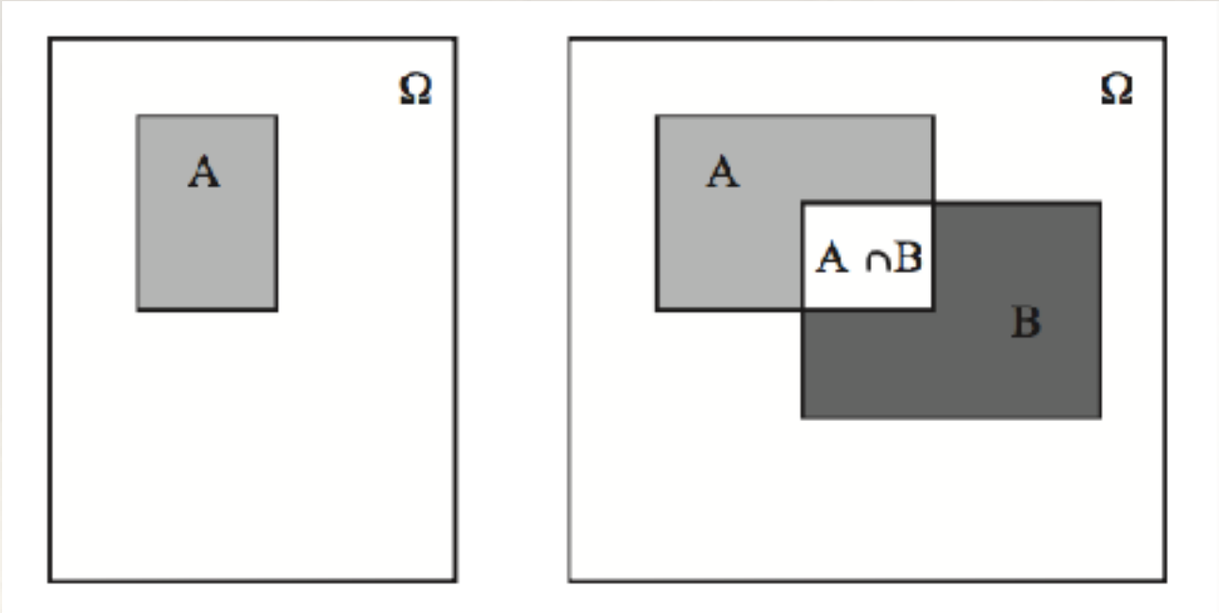
- These properties follow from the simple axioms on the last slide

Useful Facts: 4.3 *Probability of events*

- $P(\mathcal{A}^c) = 1 - P(\mathcal{A})$
- $P(\emptyset) = 0$
- $P(\mathcal{A} - \mathcal{B}) = P(\mathcal{A}) - P(\mathcal{A} \cap \mathcal{B})$
- $P(\mathcal{A} \cup \mathcal{B}) = P(\mathcal{A}) + P(\mathcal{B}) - P(\mathcal{A} \cap \mathcal{B})$
- $P(\cup_1^n \mathcal{A}_i) = \sum_i P(\mathcal{A}_i) - \sum_{i<j} P(\mathcal{A}_i \cap \mathcal{A}_j) + \sum_{i<j<k} P(\mathcal{A}_i \cap \mathcal{A}_j \cap \mathcal{A}_k) + \ldots (-1)^{(n+1)} P(\mathcal{A}_1 \cap \mathcal{A}_2 \cap \ldots \cap \mathcal{A}_n)$

Probability and size

- It can be helpful to think of the probability of an event as measuring its "size"
- We can then use diagrams to reason about certain properties. Which properties are illustrated below?



Counting and probability

In many cases, we will use counting in order to compute probabilities

If we are interested in an event $\mathcal{E}$, a set of outcomes, we might think of writing its probability as follows

$$P(\mathcal{E}) = \sum_{O \in \mathcal{E}} P(O)$$

Thus, if each outcome in Ω has the same probability then we have

$$P(\mathcal{E}) = \frac{\text{Number of outcomes in } \mathcal{E}}{\text{Total number of outcomes in } \Omega}$$

Example

- Roll two fair dice and consider the total: what is the probability of getting an odd number?
- There are 36 outcomes in the sample space and each of them are equally likely giving each a probability of 1/36
- 18 of the 36 outcomes result in an odd number
- So the probability is 18/36=1/2

Example

- If we shuffle a pack of cards and draw one at random, what is the probability that it is a red 10?
- There are 52 cards, each an equally likely outcome in the sample space
- Our event of interest is the set containing the two outcomes where a red 10 is drawn
- The probability is 2/52=1/26

Example

- Roll two dice, add the two numbers, what in the probability the result is divisible by 5?
- There are 36 equally likely outcomes as before
- There are 4 ways to get a 5 with two dice and 3 ways to get a 10, so the probability is 7/36

Example

- We stop three random people on the street and ask them on what day of the week they were born. What is the probability that they were born on days of the week that are in succession? E.g. Monday-Tuesday-Wednesday or Saturday-Sunday-Monday
- The sample space has 7^3 outcomes each equally likely
- Our event has 7 possible outcomes
- So our probability is 7/7^3 = 1/49