Lisa Yan and Jerry Cain
CS109

Section #1 Analytic Probability

Based on a handout by previous CS109 CAs Tim Gianitsos, Alex Tsun

Overview of Section Materials

The warmup questions provided will help students practice concepts introduced in lectures. The section problems are meant to apply these concepts in more complex scenarios similar to what you will see in problem sets and quizzes.

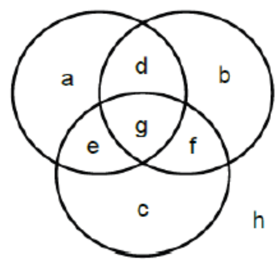
1 Warmups

1.1 Lecture 1: Counting

The Inclusion Exclusion Principle for three sets is:

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$$

Explain why in terms of a venn-diagram.



We want a, b, c, d, e, f, & g to each be accounted for exactly once. If we add A, B, and C together, we properly count a, b, & c exactly once. However we will double count d, e, & f, and triple count g. Removing $A \cap B$, $A \cap C$, & $B \cap C$ will reduce the counts for d, e, & f down to one. However, g went from being counted 3 times too many to being counted zero times. The expression corresponding to g is $A \cap B \cap C$, so we add it in to allow g to have a count of exactly one.

1.2 Lecture 2 Warmup: Permutations and Combinations

Suppose there are 7 blue fish, 4 red fish, and 8 green fish in a large fishing tank. You drop a net into it and end up with 6 fish. What is the probability you get 2 of each color?

$$\frac{\binom{7}{2}\binom{4}{2}\binom{8}{2}}{\binom{19}{6}}$$

1.3 Lecture 3 Warmup: Axioms of Probability

For each of the four statements below, evaluate True or False.

$$P(A|B) + P(A^C|B) = 1 \qquad P(A|B) + P(A|B^C) = 1 \qquad P(A \cap B) + P(A \cap B^C) = 1$$

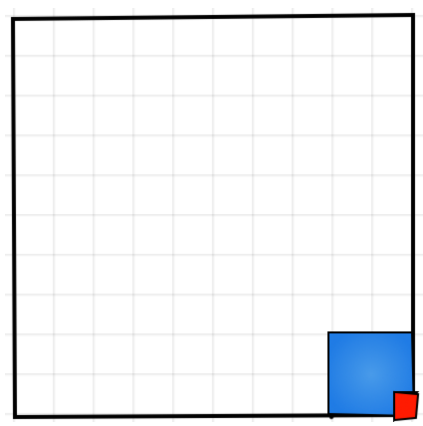
$$P(A) = 0.4 \wedge P(B) = 0.6 \implies A = B^C$$

T F F
F

1.4 *Lecture 4 Warmup: Conditional Probability and Bayes*

Bayes Theorem is $P(H|E) = P(E|H) * P(H)/P(E)$ where H can be thought of as a hypothesis and E as evidence. This equation can be notoriously counter intuitive. Draw a diagram where $P(E|H) = 1$ and $P(E^C|H^C)$ is close to 1, but $P(H|E)$ is still close to 0. How can we interpret this?

Draw the sample space as a square. Draw the (E)vidence as a tiny square inside taking up less than 5% of the sample space. Lastly, draw an even tinier dot for the (H)ypothesis that is completely inside of E. H should be less than 5% of E, (thus less than 0.25% of the sample space).



$P(E|H)$ is 1 because H is inside of E. $P(E^C|H^C)$ is close to 1 because, since both H and E are small compared to the sample space, H^C and E^C will be comparable in size to each other. $P(H|E)$ is close to zero because H is less than 5% of E.

As an example, assume 100% of NBA players are tall, and 99% of non-NBA players are not tall. We still can't infer for certain that a randomly selected tall person is in the NBA.

In general, people assume that $P(H|E) \approx P(E|H)$ without realizing it. Bayes Theorem says $P(H|E) = P(E|H) * P(H)/P(E)$. Therefore this assumption is only true if $P(H)$ / $P(E)$ is close to one - otherwise our intuitions are violated.

2 Problems

2.1 *Lecture 1 Generative Processes: The Birthday Problem*

Preamble: When solving a counting problem, it can often be useful to come up with a generative process, a series of steps that "generates" examples. A correct generative process to count the elements of set A will (1) *generate every element of* A and (2) *not generate any element of* A *more than once*. If our process has the added property that (3) *any given step always has the same number of possible outcomes*, then we can use the product rule of counting.

Example: Say we want to count the number of ways to roll two (distinct) dice where one die is even and one die is odd. Our process could be: (1) choose a number for the first die, (2) choose a number of opposite parity for the second die. Since the first step has 6 options and the second step has 3 options regardless of the outcome of the first step, the number of possibilities is $6 * 3 = 18$.

Problem: Assume that birthdays happen on any of the 365 days of the year with equal likelihood (we'll ignore leap years).

a. What is the probability that of the n people in class, at least two people share the same birthday?

It is much easier to calculate the probability that no one shares a birthday. Let our sample space, S be the set of all possible assignments of birthdays to the students in section. By the assumptions of this problem, each of those assignments is equally likely, so this is a good choice of sample space. We can use the product rule of counting to calculate $|S|$:

$$|S| = (365)^n$$

Our event E will be the set of assignments in which there are no matches (i.e. everyone has a different birthday). We can think of this as a generative process where there are 365 choices of birthdays for the first student, 364 for the second (since it can't be the same birthday as the first student), and so on. Verify for yourself that this process satisfies the three conditions listed above. We can then use the product rule of counting:

$$|E| = (365) \cdot (364) \cdot \dots \cdot (365 - n + 1)$$

$$\begin{aligned} P(\text{birthday match}) &= 1 - P(\text{no matches}) \\ &= 1 - \frac{|E|}{|S|} \\ &= 1 - \frac{(365) \cdot (364) \dots (365 - n + 1)}{(365)^n} \end{aligned}$$

An common misconception is that the size of the event E can be computed as $|E| = \binom{365}{n}$ by choosing n distinct birthdays from 365 options. However, outcomes in this event (n unordered distinct dates) cannot recreate any outcomes in the sample space $|S| = 365^n$ (n distinct dates, one for each distinct person). However if we compute the size of event E as $|E| = \binom{365}{n} n!$ (equivalent to the number above), then we can assign the n birthdays to each person in a way consistent with the sample space.

Interesting values: $(n = 13 : p \approx 0.19)$, $(n = 23 : p \approx 0.5)$, $(n = 70 : p \geq 0.99)$.

b. What is the probability that this class contains exactly one pair of people who share a birthday?

We can use the same sample space, but our event is a little bit trickier. Now E is the set of birthday assignments in which exactly two students share a birthday and the rest have different birthdays. One generative process that works for this is (1) choose the two students who share a birthday, (2) choose $n - 1$ birthdays in the same manner as in part a (i.e. one for the pair of students and one for each of the remaining students). We then have:

$$P(\text{exactly one match}) = \frac{|E|}{|S|} = \frac{\binom{n}{2}(365) \cdot (364) \cdot \ldots \cdot (365 - n + 2)}{(365)^n}$$

Many other generative processes work for this problem. Try to think of some other ones and make sure you get the same answer!

2.2 Lecture 2 Permutations and Combinations: Flipping Coins

Preamble: One thing that students often find tricky when learning combinatorics is how to figure out when a problem involves permutations and when it involves combinations. Naturally, we will look at a problem that can be solved with both approaches. Pay attention to what parts of your solution represent distinct objects and what parts represent indistinct objects.

Problem: We flip a fair coin n times, hoping (for some reason) to get k heads.

a. How many ways are there to get exactly k heads? Characterize your answer as a *permutation* of H's and T's.

We want to know the number of sequences of n H's and T's such that there are k H's and $n-k$ T's. This is the same as permuting n objects of which one set of k is indistinguishable and one set of $n-k$ is indistinguishable. Using our formula for the permutation of indistinguishable objects, we get $\frac{n!}{k!(n-k)!}$

b. For what x and y is your answer to part a equal to $\binom{x}{y}$? Why does this *combination* make sense as an answer?

Our answer to part a is equal to $\binom{n}{k}$. This makes sense because we can come up with a valid sequence by *choosing* k flips to come out to heads (and implicitly define the other $n-k$ to be tails). The answer is also equivalent to $\binom{n}{n-k}$ for which the same logic applies except with choosing flips to be tails.

c. What is the probability that we get exactly k heads?

If we define our sample space to be all possible sequences of flips, then our event is the number of sequences where we get exactly k heads, meaning that $|E|$ is (conveniently) the answer to the previous two parts. Our probability is then $\frac{|E|}{|S|} = \frac{\binom{n}{k}}{2^n}$.

2.3 Lecture 4 Bayes Rule: Song Identification

Preamble: In this class, seeing a problem written in English can often throw you off of its scent. In this problem, we will practice translating a problem from English to equations and then applying Bayes Rule, which you learned this week.

Problem: Shazam is an application which can predict what song is playing. Based on the frequency of requests it's been getting these days, Shazam has found that:

- 80% of songs are Hold Up by Beyonce
- 20% of songs are Can't Get Used to Losing You by Andy Williams

When a request is made, Shazam receives an audio sample that it uses to update its belief. From one particular audio sample, S, Shazam estimates that:

- S would have a 50% chance of appearing if Hold Up were playing.
- S would have a 90% chance of appearing if Can't Get Used to Losing You were playing.

What is the updated probability that the song is Hold Up given the audio sample heard? HINT: Define variables and write all of the information we have given to you in terms of those variables.

Let X_1 be the event that the song is Hold Up and let X_2 be the event that the song is Can't Get Used to Losing You. We can write the information from the problem as: $P(X_1) = 0.8$, $P(X_2) = 0.2$, $P(S|X_1) = 0.5$, $P(S|X_2) = 0.9$. We are looking for $P(X_1|S)$, which we can tackle with Bayes theorem:

$$P(X_1|S) = \frac{P(S|X_1)P(X_1)}{P(S)}$$

Since there are only two songs, we can expand the denominator using the law of total probability:

$$\begin{aligned} P(X_1|S) &= \frac{P(S|X_1)P(X_1)}{P(S|X_1)P(X_1) + P(S|X_2)P(X_2)} \\ &= \frac{0.50 \cdot 0.80}{0.50 \cdot 0.80 + 0.90 \cdot 0.20} \\ &\approx 0.69 \end{aligned}$$