Math Camp - Probability Notes

Kevin McAlister[1]

[1]*Department of Political Science, University of Michigan, Ann Arbor 48109*

1 What is probability?

- Define research question.
- How do I answer that question?
- Question: How many chicken fingers can I fit in my mouth?
  - Easy to answer. Cook a bunch of chicken fingers and see how many I can fit in my mouth.
  - Directly observable answer to this question.
- Question: How many squirrels are there on campus?
  - I can see squirrels running around campus. But, can I count them all?
  - What if the squirrels are hiding when I'm counting?
  - How do I know that the squirrel that I'm counting next is one that I haven't counted before?
  - Uncertain.
- Probability is an axiomatic way to talk about uncertainty.

- What things are uncertain?
  - How many students on campus sleep more than 7 hours per night? I can't ask every student. Students are unlikely to be able to answer correctly 100 percent of the time.
  - How many countries have functional democracies? How to define a functional democracy.
  - What states will cast electoral votes for Trump in 2020? Predicting the future. States change over time.
- Prior beliefs vs. Posterior beliefs
  - Rank the members of the U.S. Senate from most liberal to most conservative.
  - Prior belief: Republicans are all more conservative than Democrats.
  - Observe: Most Republicans vote against a bill. Most Democrats vote for a bill. Ben Nelson (D) votes against the bill. Olympia Snowe (R) votes for the bill.
  - Question: Is Olympia Snowe more liberal than Ben Nelson?
  - Answer: Maybe! There is some evidence this is the case.
  - Quantify the level of "maybe-ness"
- Probability is a measure of beliefs about how likely an event is to occur.
  - Will it rain today?
  - The ground is wet this morning. Did it rain last night?
- Is probability frequency of occurrence? Sort of... but not always.
  - Question: Will the sun explode tonight?
  - Answer: Probably not. But, it could.
  - Beliefs!

2 The rules of the game

- Probability is a consistent way to think about a set of uncertain events.
- Consistency requires rules!
- Refer to a set of connected random outcomes as an **experiment**.
- An experiment is a random study where all possible outcomes are known.
  - Will it rain tonight? Outcomes: $\{ItRained, ItDidn'tRain\}$
  - How many minutes does it take to boil a pot of water? Outcomes: $\mathbb{R}^+$.
  - Who will be in the 2022 World Cup? 48 teams in tournament.
    Outcome: $\{\{Qatar\}, Egypt, England, ...\}$
    How many elements are there in this probability space? $\binom{\#ofCountries}{48}$
  - How many flips of a fair coin until 5 tails? Outcomes: $\{5, 6, 7, ..., \infty\}$
  - Result of two flips of a fair coin. Outcomes: $\{HH, HT, TH, TT\}$
- A formal definition:
  - Consider an experiment. Let Ω be the sample set a.k.a the outcome set. Ω is the set of all possible outcomes.
  - Roll 2 three-sided dice. What is the sample set?
    $\Omega = \{\{1,1\}, \{1,2\}, \{1,3\}, \{2,1\}, \{2,2\}, \{2,3\}, \{3,1\}, \{3,2\}, \{3,3\}\}$
- Given a sample space, we are interested in observing **events**.
- An event is simply a collection of elements from the sample set.
- Let A be the event that I get at least one heads in 2 flips of a fair coin. What elements of the sample set belong to A?
  Answer: $A = \{HT, TH, HH\}$
- Events are not necessarily distinct. Let B be the event that I get at least one tails in 2 flips.
  Answer: $B = \{HT, TH, TT\}$
  - What is the intersection of A and B ($A \cap B$)?
  - What is the union of A and B ($A \cup B$)?
  - What is the complement of A (A^C)?
  - Are A and B exhaustive events ($A \cup B = \Omega$)?
- Given events, we can formally define probability.
- Definition: Probability is a (real-valued) set function $P(\cdot)$ that assigns to each event A in the sample space Ω a number $P(A)$, called the probability of the event A, such that the following hold:
  1. The probability of any event A must be nonnegative - $P(A) \geq 0$
  2. The probability of the sample space must be equal to 1 - $P(\Omega) = 1$
  3. Given *mutually exclusive* or *disjoint* events, that is $A_i \cup A_j = \emptyset$ for $i \neq j$, the probability of the union of events is equal to the sum of the probabilities of the individual events - $P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3)$
     - Note that this extends trivially to countably infinite sets - $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$

3 Counting and Simple Probability Exercises

- Probability is intuitive!
- Fancy counting.
- Counting can help define the sample set.
- How do we define probability for a given experiment?
- Information about the sample space combined with information about the outcomes.
- Start simple. Uniform random experiments.
- What is the probability that the card at the top of the deck is the ace of spades?
  - What makes sense here?
  - Is the outcome random?
  - Is the outcome uniformly random?
  - 52 possible outcomes.
  - One outcome of interest. Our **event** of interest is that the top card is the ace of spades.
  - 1/52
- Two steps. Determine how many possible outcomes there are. Determine how many outcomes are in our event of interest.
- Multiplication rule: Given a set of experiments, multiply the number of outcomes in each experiment to determine the number of possible outcomes in the composite experiment.
- Example: I have 3 shirts and two pairs of jeans. I pick a pair of jeans and a shirt at random each day. What is the probability that I wear my Motorhead t-shirt and black jeans on Wednesday?
  - Each shirt and each pair of jeans is selected uniformly at random.
  - How many possible outcomes? 2 jeans $\times$ 3 shirts. 6 possible outfits. 1 outfit is in my event of interest. 1/6.
- A useful generalization. Say that I want to fill n positions with n different items. How many unique ways can I arrange the n items?
- There are n items that can take the first position. $n-1$ for the second. $n-2$ for the third. And on and on and on. Think of these as a series of random experiments.
- Number of arrangements: $n \times n-1 \times n-2 \times ... \times 2 \times 1 = n!$
- More card experiments. I have an ace of each suit. What is the probability that the ace of spades is in the second position?

  - How many outcomes? $4! = 4 \times 3 \times 2 \times 1 = 24$. There are 24 permutations of the 4 aces in a deck of cards.
  - How many outcomes are in my event of interest? If the ace of spades takes the second position, then there are three open positions. $3! = 6$ outcomes in event.
  - Uniform random experiment. $6/24 = 1/4$
  - Intuitive outcome. If there are 4 possibilities for the second spot and I don't care about the rest of the outcomes, then there is a $1/4$ chance that the second spot is occupied by the ace of spades.
  - Either approach gives the correct answer.
- The question being asked is important.
- How many different license plates with 8 characters can be generated from the 36 alphanumeric characters if any character can be repeated any number of times? $36 \times 36 \times 36 \times 36 \times 36 \times 36 \times 36 \times 36 = 36^8$
- How many different license plates with 8 characters can be generated from the 36 alphanumeric characters if no character can be repeated within a single license plate? $36 * 35 * 34 * 33 * 32 * 31 * 30 * 29$
- The first case is *sampling with replacement*. The second case is *sampling without replacement*.
- What if each license plate had 24 characters and allowed no repeats? We know there $36 * 35 * 34 * ... * 14 * 13$ possible outcomes. But, this is tedious to write.
- New notation. How many permutations of 8 characters can we make from a set of 36 possible characters?
- Write in factorial notation. $\frac{36!}{(36-8)!} = \frac{36*35*34*...*3*2*1}{28*27*26*...*3*2*1}$
- nPr: $= \frac{n!}{(n-r)!}$
  - Randomly arrange characters into 36 slots (36!). This is the number of possible arrangements of all 36 characters. This goes in the numerator.
  - I don't care about any placements beyond the first 8 selections. Figure out how many irrelevant arrangements exist. This goes in the denominator.
- More complicated question: I have 3 extra tickets to see Dane Cook in Windsor. I want to give these tickets to my friends. I have 20 friends. How many different unique ways can I distribute these three tickets to my 20 friends?
- New wrinkle: Order doesn't matter. If I invite Jess then Kiela then Erin, isn't that the same as inviting Erin then Kiela then Jess?
- How do we count when order doesn't matter?
  - Intuition: Randomly place 20 friends into 20 slots (20!). This is the numerator.

- How many parts of the arrangements don't I care about? I don't care about who is ranked 4 - 20. 10 spots. 17! arrangements that I don't care about.
- For the arrangements I do care about, how many ways can they be repeated? 3! ways that the meaningful arrangements are repeated (i.e. Erin, Jess, Kiela ; Erin, Kiela, Jess; Kiela, Jess, Erin; Kiela, Erin, Jess; Erin, Jess, Kiela; Erin, Kiela, Jess). Divide those out too.
- This leaves us with $\frac{20!}{3!17!}$ possible **combinations**.

- Combinations appear very frequently in probability theory.
- If I flip a coin three times, what is the probability that I get 2 heads? 3! possible arrangements of heads and tails. 1 spot I don't care about. 2 ways that the meaningful part of the arrangment can be premuted.
- Formal definition: $nCr = \frac{n!}{r!(n-r)!} = \binom{n}{r}$
- Why worry about this? Back to uniform random experiments.
- Assume that I given a ticket to each of my twenty friends with equal probability. What is the probability that Kiela and Jess and Erin get tickets?
- $\binom{20}{3}$ disjoint outcomes. 1 possible outcome. $\frac{1}{\binom{20}{3}}$

4 Conditional Probability, Independence, and Bayes Rule

- Let's start with a motivating example. We want to detect ballot boxes that have fraudulent votes. Given that we cannot directly observe fraudulence, we have to estimate the probability that a ballot box is fraudulent. Research question: **Given that I observe that 90% of the votes in a ballot box are for the current majority party**, what is the probability that the ballot box contains fraudulent votes?
- For a sample space, I may care about *conditional* events.
- Given that A occurs, what is the probability that B occurs?
- Perhaps more important than standard probability for the social sciences. Inferences are made conditional on observed data.
- Conditional probability is a way of partitioning a sample space.
- Draw it out! Use Venn diagrams.
- A conditional probability is still a probability. The rules of the game still apply.

- Denote the probability that A occurs **given** that B has occurred as $P(A|B)$.
- Now, we can define *disjoint* events in a formal way. If A and B are disjoint events, then $P(A|B) = 0$.
- How do we define conditional probabilities for a set of events? The formal calculation:

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

- Draw it out. Remember the rules of the game. If I observe than an event occurs, normalize the sample space to fit the rules again.
- Intuitive! Renormalizing over partition of the probability space.
- An introductory example: I put 4 white balls and 3 red balls in a jar and draw two balls. Given that the first ball was red, what is the probability that the second ball is red?
  - Let's approach this like a counting problem. First, is this a uniform random experiment?
  - Label each of the balls with a number (W1,W2,W3,W4,R1,R2,R3). How many unique doublets can we pull from the jar? Remember that this is an example of a 2 step experiment without replacement. What is the probability of any distinct outcome? (42 possible outcomes)
  - How many outcomes have a red ball drawn first? What is the probability that a red ball is drawn first? Does this make sense? (18 outcomes)
  - Of those outcomes, how many have a red ball drawn second? (6 outcomes) Divide this value by the total number of outcomes that have a red ball drawn first. This is the probability that B is red given that A is red.
- Up until now, we've had to rely on uniform random experiments to define probability in complex settings. Each disjoint event has an equal probability of occurrence. This is easily extended to cases with different probabilities specified *a priori*. Now, we can start working with non-standard events (i.e. joint events, dependent events, etc.).
- IMPORTANT DEFINITION!!!!! Chain rule for probabilities:

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

- Extends to many events:

$$P(A \cap B \cap C) = P(A|B, C)P(B|C)P(C)$$

- Note that this is a direct consequence of the definition of conditional probability:

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

- This makes sense. $P(B)$ gives the mass of the partition of the sample space. $P(A|B)$ is the mass of the partitioned space where A and B occur.
- An example: I have a drawer with 4 red socks, 6 brown socks, and 8 green socks. I select two socks at random. What is the probability that I select two green socks given that I have selected two socks of the same color?
  - What is the intersection of selecting two green socks and selecting two socks of the same color?
  - Start by determining the probability that I select two socks of the same color. Are selecting two red socks and selecting two green socks disjoint events?
  - $P(2G|2S) = \frac{P(2G)}{P(2G)+P(2R)+P(2B)}$
  - Uniform random experiment. We can enumerate all possible outcomes. But, there's an easier way.
  - Suggestions?
  - $P(2G) = \frac{8}{18} * \frac{7}{17}, P(2R) = \frac{4}{18} * \frac{3}{17}, P(2B) = \frac{6}{18} * \frac{5}{17}$
  - $\frac{56}{56+12+30} = \frac{56}{98}$
- Conditional probability is key to understanding relationships between events. Conditional probability gives a way to define *joint probability* (i.e. $P(A,B)$).
- Joint probability is a probability measure that tells us the likelihood that two events will occur. We've already defined this formally.
- Joint probability and conditional probability are explicitly linked - $P(A,B) = 0 \rightarrow P(A|B) = 0$
- When is this the case?
- What about when two events are not dependent on one another? For example, I roll a die twice. What is the probability that I roll a 1 on the first roll and a 2 on the second roll?
- Define *independent* events:

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

- Independence is intuitive, but can be misleading.

- Gambler's fallacy - I won the last game of Blackjack. The dealer reshuffles the deck. What is the probability that I win the next game? The same as the first one.
- Hands of Blackjack are independent. Knowing the outcome of the last game gives me no info about the result of the next game.
- Independence implies identical partition sizes in conditional events:

$$P(A, B) = P(A|B)P(B) = P(A)P(B)$$

- A key identity in probability theory.