

Recitation/Discussion 1

Probability Basics and Combinatorics

accompaniment to MATH 170E: Introduction to Probability and Statistics 1: Probability
at the University of California, Los Angeles

Allen Lin

In this recitation, we will review properties of probability and use those properties to calculate probability of events. Then we will review definitions of permutations and combinations in combinatorics and use them to enumerate and count objects.

Exit Survey

Please fill out the exit survey.

This is so I can get feedback and know what to focus on next time!

Properties of Probability

For an event A from a sample space S we want to associate A to a number $P(A)$ that describes the frequency or the likelihood of the event occurring. For example, if $S = \{1, 2, 3, 4, 5, 6\}$ is the sample space of outcomes of a standard die and $A = \{2, 4, 6\}$ is the event of rolling an even number, we want $P(A)$ to encode the frequency of A or the likelihood of A occurring. Intuitively, the number of outcomes of A is half the number of total outcomes in S , so we want or should expect $P(A) = 1/2$.

Definition 1. For any set X , the *power set* of X , denoted by $\mathcal{P}(X)$, is the set of subsets of X .

Example 1. If $S = \{1, 2, 3\}$, then

$$\mathcal{P}(S) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}.$$

Example 2. Recall that $\emptyset \subseteq S$ for any set S , so $\emptyset \in \mathcal{P}(S)$ for any subset S . In particular, if $S = \emptyset$, then $\mathcal{P}(S) = \mathcal{P}(\emptyset) = \{\emptyset\}$.

A *set function* is a function that associates a real number to every set. Now we have all the necessary tools to define the notion of probability.

Definition 2. Let S be a sample space. Then P denotes the *probability* if it is a real-valued set function $P: \mathcal{P}(S) \rightarrow [0, 1]$ such that

1. $P(A) \geq 0$;
2. $P(S) = 1$;

Of course, this is assuming that each outcome is equally likely.

Sometimes, the power set of X is also denoted by 2^X . Can you guess why? As a hint, look at the next few examples and calculate the size of the power set of X and compare it with the size of X .

This is an important example!

3. if $\{A_k\}_{k \in \mathbb{N}}$ is a collection of pairwise disjoint events (i.e., $A_i \cap A_j = \emptyset$ for $i \neq j$), then

$$P\left(\bigcup_{n=1}^k A_k\right) = \sum_{n=1}^k P(A_k),$$

This is what we call a finite union.

and

$$P\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} P(A_n).$$

This is what we call a countable union. A countable union is a union over sets indexed by $\mathbb{N}$.

In the sequel, we let S denote the sample space and A^c denote the complement $S \setminus A$ of A . We now recall the important properties of probability.

This theorem relates the probability of an event to its complement. This is useful for calculating.

Theorem 3. For any event A , $P(A^c) = 1 - P(A)$.

This is extremely useful in calculating probability. Sometimes it is easier to calculate the probability that an event doesn't happen compared to the probability that an event does happen.

Hint. Write $S = A \cup A^c$.

Proof. Note that $A \cap A^c = \emptyset$, so they are disjoint. Moreover, $S = A \cup A^c$, so we use properties 2 and 3 to obtain

$$1 = P(S) = P(A) + P(A^c),$$

so $P(A^c) = 1 - P(A)$. ■

This next theorem is fundamental and is used to show that the probability of mutually exclusive events is 0.

Theorem 4. $P(\emptyset) = 0$.

This should make sense intuitively. $\emptyset$ represents the set with no possibilities (or sometimes an impossible set), so the probability that the event occurs should be 0.

Hint. How can you use the previous theorem?

Proof. Take $A = \emptyset$ in the previous theorem, noting $A^c = S$, to obtain

$$1 = P(A^c) = 1 - P(A).$$

Thus $P(A) = 0$. ■

The next theorem is known as *monotonicity* and is useful when trying to bound the probability of an event.

Theorem 5. Let A and B be events. If $A \subseteq B$, then $P(A) \leq P(B)$.

Intuitively, $A \subseteq B$ implies that if $x \in A$, then $x \in B$. In particular, this means that whenever event A happens, event B also happens, so the probability that B happens should be at least as big as that of A .

Hint. Write B as a disjoint union of A and some other set.

Proof. We write $B = A \cup (B \cap A^c)$ and note that $A \cap (B \cap A^c) = \emptyset$, so

$$P(B) = P(A) + P(B \cap A^c) \geq P(A)$$

using property 1. ■

We have the following naive bound on the probability of any event.

Theorem 6. For any event A , $P(A) \leq 1$.

Hint. How can you use the previous theorem?

Proof. Since $A \subseteq S$, apply the previous theorem with $B = S$ to get $P(A) \leq P(S) = 1$. ■

This means that $0 \leq P(A) \leq 1$ (i.e., the probability of any event is between 0 and 1, inclusive). You can use this to check that your calculation is right; for example, if you obtain $P(A) = 3/2$ or something negative, something went wrong.

Finally, this last theorem is known as *the inclusion-exclusion principle*.

Theorem 7. Let A and B be events. Then

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Hint. To motivate the proof, draw a Venn diagram.

Proof. Note that $A \cup B = A \cup (A^c \cap B)$ and $A \cap (A^c \cap B) = \emptyset$, so

$$P(A \cup B) = P(A) + P(A^c \cap B).$$

On the other hand, $B = (A \cap B) \cup (A^c \cap B)$ and $(A \cap B) \cap (A^c \cap B) = \emptyset$, so

$$P(B) = P(A \cap B) + P(A^c \cap B).$$

Combining the two equations gives

$$P(A \cup B) = P(A) + P(B) - P(A \cap B),$$

as required. ■

This can be generalized to any number of events.

Theorem 8. For events $A_1, \dots, A_n$,

$$\begin{aligned} P\left(\bigcup_{k=1}^n A_k\right) &= \sum_{k=1}^n P(A_k) - \sum_{1 \leq i < j \leq n} P(A_i \cap A_j) \\ &\quad + \sum_{1 \leq i < j < k \leq n} P(A_i \cap A_j \cap A_k) - \dots + (-1)^{n+1} P\left(\bigcap_{k=1}^n A_k\right). \end{aligned}$$

There are better bounds like the Chebyshev bound and Chernoff bounds that are tighter than this and used in randomized algorithms.

Since $P(A \cap B) \geq 0$, we obtain $P(A \cup B) \leq P(A) + P(B)$, and this is an equality if and only if $P(A \cap B) = 0$. This is the *union bound* and is applied everywhere in randomized algorithms.

Draw a Venn diagram for $n = 3$ and see how this works!

Now let's apply these properties.

Exercise 1. Toph and Aang are playing a dice game in which a pair of dice is rolled repeatedly. Toph wins if a sum of 5 is rolled before any sum greater than or equal to 8, and Aang wins if any sum greater than or equal to 8 is rolled first. We will find the probability that Toph wins the game. First, calculate

- the probability that a sum of 5 is rolled on any given roll of the pair of dice and
- the probability that a sum of any number greater than or equal to 8 is rolled on any given roll of the pair of dice.

Solution. Here $S = \{1, \dots, 6\}^2$, so there are 36 total possibilities. Let A be the event that a sum of 5 is rolled. There are four possible ways to roll a 5: $A = \{(1, 4), (2, 3), (3, 2), (4, 1)\}$, so $P(A) = 4/36 = 1/9$. Let B be the event that a sum of any number greater than or equal to 8 is rolled. Then

$$B = \underbrace{\{(2, 6), (3, 5), (4, 4), (5, 3), (6, 2)\}}_{\text{rolling a 8}} \cup \underbrace{\{(3, 6), (4, 5), (5, 4), (6, 3)\}}_{\text{rolling a 9}} \\ \cup \underbrace{\{(4, 6), (5, 5), (6, 4)\}}_{\text{rolling a 10}} \cup \underbrace{\{(5, 6), (6, 5)\}}_{\text{rolling a 11}} \cup \underbrace{\{(6, 6)\}}_{\text{rolling a 12}}$$

for a total of $5 + 4 + 3 + 2 + 1 = 15$ possibilities. Thus $P(B) = 15/36 = 5/12$.

Exercise 2. Calculate the probability that neither Toph nor Aang wins on a given roll.

Solution. Toph wins if A occurs, and Aang wins if B occurs. Note that A and B are mutually exclusive, i.e., $A \cap B = \emptyset$, so $P(A \cap B) = P(\emptyset) = 0$.

$$\begin{aligned} P\left((A \cup B)^c\right) &= 1 - P(A \cup B) \\ &= 1 - P(A) - P(B) + P(A \cap B) \\ &= 1 - \frac{1}{9} - \frac{5}{12} \\ &= \frac{17}{36}. \end{aligned}$$

Exercise 3. Let E_n be the event that a sum of 5 occurs on the n th roll and a sum of neither 5 nor any number greater than or equal to 8 occurs on the first $n - 1$ rolls. Compute $P(E_n)$.

Hint. Events are *independent* if and only if the outcome of any one event does not affect the rest. Argue that the rolls are independent and use the fact that if $X_1, \dots, X_n$ are independent events, then

$$P\left(\bigcap_{k=1}^n X_k\right) = \prod_{k=1}^n P(X_k).$$

Solution. The rolls are independent because any roll does not affect the outcome of the next roll. Then

$$\begin{aligned} P(E_n) &= \underbrace{P\left((A \cap B)^c\right)}_{\text{first roll}} \times \cdots \times \underbrace{P\left((A \cap B)^c\right)}_{(n-1)\text{th roll}} \times \underbrace{P(A)}_{n\text{th roll}} \\ &= \frac{17}{36} \times \cdots \times \frac{17}{36} \times \frac{1}{9} \\ &= \frac{1}{9} \times \left(\frac{17}{36}\right)^{n-1}. \end{aligned}$$

Exercise 4. Compute $\sum_{n=1}^{\infty} P(E_n)$ and argue that this is the probability that Toph wins the game.

Solution. Toph wins the game if E_n occurs for some $n \in \mathbb{N}$. Thus

$$\text{event that Toph wins} = \bigcup_{n=1}^{\infty} E_n,$$

and these events are clearly mutually exclusive. Thus

$$\begin{aligned} P(\text{event that Toph wins}) &= P\left(\bigcup_{n=1}^{\infty} E_n\right) \\ &= \sum_{n=1}^{\infty} P(E_n) \\ &= \sum_{n=1}^{\infty} \frac{1}{9} \left(\frac{17}{36}\right)^{n-1} \\ &= \frac{1}{9} \sum_{n=0}^{\infty} \left(\frac{17}{36}\right)^n \\ &= \frac{1}{9} \cdot \frac{1}{1 - 17/36} \\ &= \frac{4}{19}, \end{aligned}$$

by the inclusion-exclusion principle and by the sum of a geometric series with ratio $r = 17/36$.

Exercise 5. Roll a fair six-sided die three times. Let

$$\begin{aligned} A_1 &= \{1 \text{ or } 2 \text{ on first roll}\}, \\ A_2 &= \{3 \text{ or } 4 \text{ on second roll}\}, \\ A_3 &= \{5 \text{ or } 6 \text{ on third roll}\}. \end{aligned}$$

- Compute $P(A_1 \cup A_2 \cup A_3)$ using the inclusion-exclusion principle.
- Compute $P(A_1 \cup A_2 \cup A_3)$ without using the inclusion-exclusion principle.

Hint. Argue that the events are independent.

Solution. The rolls are independent because any roll does not affect the outcome of the next roll. For the first task, we use the inclusion-exclusion principle to obtain

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3) &= P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) \\ &\quad - P(A_1 \cap A_3) - P(A_2 \cap A_3) \\ &\quad + P(A_1 \cap A_2 \cap A_3) \\ &= \frac{1}{3} + \frac{1}{3} + \frac{1}{3} - \frac{1}{9} - \frac{1}{9} - \frac{1}{9} + \frac{1}{27} \\ &= \frac{19}{27}. \end{aligned}$$

For the second, we see that

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3) &= 1 - P\left((A_1 \cup A_2 \cup A_3)^c\right) \\ &= 1 - P\left(A_1^c \cap A_2^c \cap A_3^c\right) \\ &= 1 - P\left(A_1^c\right) P\left(A_2^c\right) P\left(A_3^c\right) \\ &= 1 - \left(1 - \frac{1}{3}\right)^3 \\ &= \frac{19}{27}. \end{aligned}$$

Combinatorics

The main counting technique in determining the number of outcomes is the *multiplication principle*: If there are n_1 choices for the first position, n_2 choices for the second position, and in general n_k choices for the k th position, then there are a total of $n_1 n_2 \cdots n_k$ possible arrangements for k positions.

You can derive everything in this section from this principle, so it's important to remember.

Exercise 6. How many possible orders of boba drinks are there if there are 5 choices of flavor, 8 choices of topping, 3 choices of sugar level, and 3 choices of size?

Solution. Using the multiplication principle we get $5 \cdot 8 \cdot 3 \cdot 3 = 360$.

Exercise 7. Use the multiplication principle to show that the number of arrangements of n distinct objects in n different positions is $n!$.

Solution. There are n choices for the first position, $n - 1$ choices for the second position, and in general $n - k$ choices for the k th position, so

$$n \times (n - 1) \times \cdots \times 1 = n!$$

total positions.

Definition 9. A *permutation* of n objects is any arrangement of the n objects.

Corollary 10. There are $n!$ permutations of n objects.

In more generality, what about the number of arrangements of n distinct objects in r different positions?

Exercise 8. Use the multiplication principle to show that the number of arrangements of n distinct objects in r different positions is ${}_nP_r = n!/(n - r)!$.

Solution. There are n choices for the first position, $n - 1$ choices for the second position, and so on, until there are $n - r + 1$ for the r th position, so

$$n \times (n - 1) \times \cdots \times (n - r + 1) = \frac{n!}{(n - r)!}$$

total positions.

Definition 11. A *permutation of n objects taken r at a time* is any arrangement of the n objects in r positions.

Corollary 12. There are ${}_nP_r = n!/(n - r)!$ permutations of n objects taken r at a time.

When counting, keep in mind whether you are sampling with replacement or sampling without replacement. Using the multiplication principle, there are n^r samples of size r taken from a set of n objects when sampling with replacement, whereas there are only ${}_nP_r$ samples of size r taken from a set of n objects when sampling without replacement.

The convention $0! = 1$ means that there is only one permutation of 0 objects.

Sampling with replacement occurs when an object is selected and replaced, whereas the object is not replaced in sampling without replacement.

Exercise 9. Compute the number of functions $\{1, \dots, n\} \rightarrow \{1, \dots, m\}$.

Solution. A function assigns every element in $X = \{1, \dots, n\}$ to an element in $Y = \{1, \dots, m\}$. For every element in X there are m choices in Y ; there are n elements in X , so using the multiplication principle we have m^n functions.

Exercise 10. Compute the number of injections $\{1, \dots, n\} \rightarrow \{1, \dots, m\}$.

Solution. We assume $n \geq m$; otherwise, the number of injections is 0. Injections can be determined by first arranging the m objects and then reading off the first n positions. There are ${}_mP_n$ such arrangements.

Sometimes, we want to disregard the order of the arrangements and count them as one. For example, in a poker hand of five cards, any arrangement of the five cards constitute the same hand. To counteract this, we divide the total number of permutations of 52 cards taken 5 at a time by $5!$ to get $52!/(5! \cdot 47!) = 2598960$.

In general, from ${}_nP_r$ permutations of n objects taken r at a time, we divide by $r!$ because any permutation of those r objects constitute the same unordered arrangement, and there are $r!$ such permutations. We denote this by ${}_nC_r = \binom{n}{r} = n!/(r!(n-r)!)$.

Definition 13. A combination of n objects taken r at a time is any arrangement of the n objects taken r positions, without order.

Corollary 14. There are ${}_nC_r = n!/(r!(n-r)!)$ combinations of n objects taken r at a time.

Exercise 11. Show algebraically and combinatorially that

$$\binom{n}{k} = \binom{n}{n-k}$$

for positive integers $k \leq n$.

Proof. Algebraically, we have

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n!}{(n-k)!k!} = \binom{n}{n-k}.$$

Combinatorially, the number of ways to choose k objects out of n is equivalent to the number of ways to not choose $n-k$ objects out of n . ■

Exercise 12. Prove combinatorially that

$$\sum_{k=0}^n \binom{n}{k} = 2^n.$$

Hint. Consider the set $S = \{1, \dots, n\}$ and its power set $\mathcal{P}(S) = 2^S$.

Proof. The number of subsets of $S = \{1, \dots, n\}$ of size k is $\binom{n}{k}$ since this is a combination of n objects taken k at a time, without order. As the size k ranges from 0 to n , we attain all the subsets of S , which there are $2^{|S|} = 2^n$ many. Thus, adding the number of subsets of size k for each $k = 0, \dots, n$ gives

$$\sum_{k=0}^n \binom{n}{k} = 2^n,$$

as required. ■

These numbers $\binom{n}{k}$ are also known as binomial coefficients because of their appearance in the binomial theorem. In the expansion $(a+b)^n$, to calculate the coefficient of the product $a^k b^{n-k}$, we can choose a from k factors and b from the other $n-k$ factors, leaving a total of $\binom{n}{k}$ total ways of choosing.

You can also prove this using induction.