Introduction to Probability

Kon Yi

February 26, 2026

Abstract

Lecture note of Introduction to Probability.

Contents

Chapter 1

Combinatorial Analysis

Lecture 1

24 Feb.

Definition 1.0.1 (Gerolamo Cardano (1501-1576) Basic Probability Model)**.**

- Sample space: set of all possible outcomes.
- Event: $E \subseteq S$, the set of outcomes we are interested in.
- Probability: $\mathbb{P}(E) \in [0, 1]$.

Remark 1.0.1. In (finite) uniform model,

$$\mathbb{P}(E) = \frac{|E|}{|S|}.$$

Example 1.0.1. Rolling a (fair) die, then what is the probability of getting a six?

Proof. $S = \{1, 2, 3, 4, 5, 6\}, E = \{6\}$, then $\mathbb{P}(E) = \frac{|E|}{|S|} = \frac{1}{6}$. ⊛

Example 1.0.2. If we roll a fair die, then what is the probability of rolling a prime?

Proof. $S = \{1, 2, 3, 4, 5, 6\}, E = \{2, 3, 5\}$, then $\mathbb{P}(E) = \frac{|E|}{|S|} = \frac{3}{6} = \frac{1}{2}$. ⊛

Example 1.0.3. Standard deck of 52 cards. Draw a random card, then what is the probability of getting an ace?

Answer. $\frac{1}{13}$. ⊛

Example 1.0.4. Roll two fair dice, what is the probability of the sum being 7?

Proof.

$$S = \{(1, 1), (1, 2), \ldots, (1, 6), (2, 1), (2, 2), \ldots, (6, 6)\},$$

and

$$E = \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\},$$

so $\mathbb{P}(E) = \frac{|E|}{|S|} = \frac{6}{36} = \frac{1}{6}$. ⊛

1.1 Counting Rule

Theorem 1.1.1 (Counting Rule)**.** If a set S is a disjoint union

$$S = S_1 \uplus S_2 \uplus \ldots \uplus S_n,$$

i.e. $S_i \cap S_j = \varnothing$ for all $i \neq j$, then

$$|S| = \sum_{i=1}^{n} |S_i|.$$

Example 1.1.1. Roll two fair dice. What is the probability of having at least one odd number?

Proof. $E = \{\text{at least one odd roll}\}$, then $E = E_1 \cup E_2$ where $E_1 = \{\text{first die is odd}\}$ and $E_2 = \{\text{second die is odd}\}$. However, $E_1 \cap E_2 \neq \varnothing$. Thus, instead, we define $E_1' = \{\text{first die is odd}\}$ and $E_2' = \{\text{first die is even and second die is odd}\}$, then we know

$$E = E_1' \uplus E_2',$$

so we have $|E| = |E_1'| + |E_2'|$, and $S = \{(x, y) : x, y \in [6]\}$, and we know

$$E_1' = \{(x, y) \mid x_1 \in \{1, 3, 5\}, y \in [6]\}$$

and

$$E_2' = \{(x, y) \mid x \in \{2, 4, 6\}, y \in \{1, 3, 5\}\},$$

which gives $|E_1'| = 18$ and $|E_2'| = 9$, and thus

$$\mathbb{P}(E) = \frac{|E|}{|S|} = \frac{18 + 9}{36} = \frac{3}{4}.$$

⊛

Theorem 1.1.2 (Product Rule)**.** If a set S is the Cartesian product of sets $S_1, S_2, \ldots, S_n$, i.e.

$$S = S_1 \times S_2 \times \cdots \times S_n = \{(a_1, a_2, \ldots, a_n) : \forall i \in [n], a_i \in S_i\},$$

then $|S| = |S_1| \times |S_2| \times \ldots |S_n| = \prod_{i=1}^{n} |S_i|$.

Remark 1.1.1. Informally, if a big chain can be broken into a sequence of smaller choice, then the total number of options is the product of the number of options for each small choice.

Example 1.1.2. Roll two fair dice. What is the probability that the sum is odd?

Proof. $S = \{(x, y) \mid x, y \in [6]\}$. Also, we know

$$E = \{\text{sum is odd}\} = \{\text{exactly one die is odd and the other is even}\} = E_1 \uplus E_2,$$

where $E_1 = \{\text{first is odd and second is even}\}$ and $E_2 = \{\text{first is even and second is odd}\}$. Note that

$$E_1 = \{1, 3, 5\} \times \{2, 4, 6\} \text{ and } E_2 = \{2, 4, 6\} \times \{1, 3, 5\},$$

so $|E_1| = |E_2| = 9$. By the sum rule, we know $|E| = |E_1| + |E_2| = 9 + 9 = 18$, and so

$$\mathbb{P}(E) = \frac{|E|}{|S|} = \frac{18}{36} = \frac{1}{2}.$$

⊛

Theorem 1.1.3 (Advanced Product Rule). If we are making a series of n choices, and for the i-th choice, we always have k_i options available, then the total number of options is

$$k_1 \times k_2 \times \cdots \times k_n = \prod_{i=1}^{n} k_i.$$

Example 1.1.3. Roll two fair dice. What is the probability that the sum is odd?

Proof. We know $S = \{(x, y) : x, y \in [6]\}$, and $E = \{(x, y) \in S : 2 \nmid x + y\}$.

- First question: Which roll is odd?
- Second question: What is the first roll? How many options?
- Third question: What is the second roll? How many options?

For the first question, we have 2 options, the first and the second. For the second question, we know there are 3 options since we need the first die to be even or odd, and similarly we know the second roll also has 3 options. Hence, $|E| = 2 \times 3 \times 3 = 18$, and thus $\mathbb{P}(E) = \frac{1}{2}$ since $|S| = 36$. ⊛

1.1.1 Permutations

Claim 1.1.1. There are

$$n! = n \times (n-1) \times (n-2) \times \cdots \times 1$$

ways to order n distinct elements.

Proof. Use the advanced product rule. For the first option, we have n choices, and the second has $n-1$ options, and so on. ⊛

1.1.2 Combinations

Question. How many subsets of size r are there of an n -element set?

Definition 1.1.1. The binomial coefficient $\binom{n}{r}$, "n choose r" counts the number of r-element subsets of an n-element set.

Claim 1.1.2. $\forall 0 \leq r \leq n$, we have

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

Proof. We count the number of ways of ordering all n items in two different ways. (Double counting)

- First method: Direct permutation, which has $n!$ ways.
- Second method:
  - Step 1: Choose which elements will be in the front r, which has $\binom{n}{r}$ elements.
  - Step 2: Order these r elements, which has $r!$ methods.
  - Step 3: Order the remaining $n-r$ elements, which has $(n-r)!$ methods.

  Thus, by advanced product rule, we know

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

$\circledast$

Observation. For all $0 \leq r \leq n$,

$$\binom{n}{r} = \binom{n}{n-r}.$$

Proof.

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

$\blacksquare$

Observation. Choosing a subset of r elements is equivalent to choose the $n-r$ elements that don't go in the subset. In fact, it means $\binom{n}{r} = \binom{n}{n-r}$.

Proposition 1.1.1 (Pascal's identity)**.** $\forall 1 \leq r \leq n$,

$$\binom{n+1}{r} = \binom{n}{r} + \binom{n}{r-1}.$$

Proof.

$$\binom{n}{r} + \binom{n}{r+1} = \frac{n!}{r!(n-r)!} + \frac{n!}{(r+1)!(n-(r+1))!} = \frac{(n+1)!}{r!(n+1-r)!}.$$

$\blacksquare$

Lecture 2

26 Feb.

Another proof for Pascal's identity. Since we know

$$\binom{n}{r} = \#\text{ of subsets of size } r \text{ from a set of size } n\text{, say } [n] = \{1, 2, \ldots, n\}.$$

Let $S = \binom{[n]}{r}$ be the set of these subsets.

Notation. For any set X,

$$\binom{X}{r} = \{Y \subseteq X, |Y| = r\},$$

so $\left|\binom{X}{r}\right| = \binom{|X|}{r}$.

Let

$$S_1 = \left\{y \in \binom{[n]}{r} : n \notin Y\right\} \text{ and } S_2 = \left\{y \in \binom{[n]}{r} : n \in Y\right\},$$

then $S = S_1 \cup S_2$, and thus $|S| = |S_1| + |S_2|$. Now since $|S_1| = \binom{n-1}{r}$ and $|S_2| = \binom{n-1}{r-1}$, so we know

$$\binom{n}{r} = |S| = |S_1| + |S_2| = \binom{n-1}{r} + \binom{n-1}{r-1}.$$

$\blacksquare$

Remark 1.1.2. Pascal's identity gives a recursive formula for computing $\binom{n}{r}$, which is much simpler than $\binom{n}{r} = \frac{n!}{r!(n-r)!}$.

Theorem 1.1.4 (Binomial Theorem)**.** For any integer $n > 0$ and any $x, y \in \mathbb{R}$,

$$(x+y)^n = \sum_{r=0}^{n} \binom{n}{r} x^r y^{n-r}.$$

Proof. Note that

$$(x+y)^n = \underbrace{(x+y)(x+y)\dots(x+y)}_{n \text{ times}}.$$

Each monomial comes from choosing one term from each factor(x or y), taken the from $x^r y^{n-r}$, where r is the number of factors from which we choose x, $0 \le r \le n$. Note that the coefficient of $x^r y^{n-r}$ is $\binom{n}{r}$ for all r, so

$$(x+y)^n = \sum_{r=0}^{n} \binom{n}{r} x^r y^{n-r}.$$

■

Corollary 1.1.1. Total number of subsets of an n-element set is 2^n.

Proof. Apply the product rule, for each of the elements, ask if it is in the subset or not. We know there are 2 options per question, so the product rule gives 2^n subsets. ■

Another proof. Let S be the set of all subsets. Then for $0 \le r \le n$, let $S_r \subseteq S$ be the subsets of those subsets of size r. Then,

$$S = S_0 \uplus S_1 \uplus \dots \uplus S_n,$$

and by sum rule we know

$$|S| = \sum_{r=0}^{n} |S_r| = \sum_{r=0}^{n} \binom{n}{r} = \sum_{r=0}^{n} \binom{n}{r} 1^r 1^{n-r} (1+1)^n = 2^n.$$

■

Corollary 1.1.2. Let X be an n-element set, $n \ge 1$, then the number of even-sized subsets of X is equal to the number of odd-sized subsets of X.

An approach for odd n and even r.

$$\# \text{ of even-sized subsets} = \sum_{\substack{r \text{ even} \\ 0 \le r \le n}} \binom{n}{r} = \sum_{\substack{r \text{ even} \\ 0 \le r \le n}} \binom{n}{n-r}.$$

Now if n is odd and r even, then $n - r$ is odd, so

$$\sum_{\substack{r \text{ even} \\ 0 \le r \le n}} \binom{n}{n-r} = \sum_{\substack{r \text{ odd} \\ 0 \le r \le n}} \binom{n}{r}.$$

■

Proof. Want to show

$$\sum_{r \text{ even}} \binom{n}{r} - \sum_{r \text{ odd}} \binom{n}{r} = 0,$$

i.e.

$$\sum_{r=0}^{n} \binom{n}{r} (-1)^r = 0.$$

Note that

$$\sum_{r=0}^{n} \binom{n}{r} (-1)^r = \sum_{r=0}^{n} \binom{n}{r} (-1)^r 1^{n-r} = (1 + (-1))^n = 0.$$

■

Example 1.1.4. In poker, we are dealt a hand of five cards from a standard deck. What is the probability of getting a full house (three-of-a-kind and a pair)?

Proof. The sample space is the subsets of 5 cards from deck. If D means the deck of cards, then $S = \binom{D}{5}$ and $|S| = \binom{52}{5}$. Also, we know $E = \{\text{full houses}\} = \{\text{triple} + \text{pair}\}$, so we can first choose the triple then choose the pair, and for triple and pair, we have to choose suits and values. Thus, we have $13 \times \binom{4}{3}$ options for the triple and $12 \times \binom{4}{2}$ options for the pair. Thus, number of full houses is $13 \times \binom{4}{3} \times 12 \times \binom{4}{2}$, and thus

$$\mathbb{P}(\text{full houses}) = \frac{13 \times \binom{4}{3} \times 12 \times \binom{4}{2}}{\binom{52}{5}}.$$

⊛

1.1.3 Choosing with repetition

As previously seen.

$$\begin{aligned}\binom{n}{r} &= \#\text{ of subsets of size } r \text{ of a set of size } n \\ &= \#\text{ of ways of choosing } r \text{ items out of } n \text{ without order and without repetition.}\end{aligned}$$

Question. What if repetition is allowed? How many ways can we choose r items out of n, with repitition but without order?

Let x_i be the number of times the i-th element is chosen, then we want to count the number of $(x_1, x_2, \dots, x_n)$ pairs satisfying

$$x_i \geq 0, \quad x_i \in \mathbb{Z}, \quad \sum_{i=1}^{n} x_i = r.$$

We can use a method call **stars and bars drawing** to count the number of such pairs. We represent our choices with stars, and use bars to separate the different elements. For example, $(1, 0, 2, 0, 0)$, which is a possible pair, and it corresponds to

$$*||**||,$$

and every possible pair corresponds to a diagram like this, and it is a bijection, while there are r stars and $n-1$ bars in each diagram, so there are $\binom{n+r-1}{n-1}$ possible pairs.

Example 1.1.5. There is a probability course with 77 students. Professor chooses 60 students to pass. How many options does the professor have.

Answer. $\binom{77}{60}$. ⊛

Example 1.1.6. What if the professor instead needs to assign grades to the students. Professor decides to distributes 4500 points between the 77 students. How many ways are there of doing this?

Proof. Let x_i be the number of points to student i, then $x_i \geq 0$ and $\sum_{i=1}^{77} x_i = 4500$. Thus, the number of solutions is $\binom{4576}{4500}$. ⊛

Example 1.1.7. What if every student should receive at least 10 points?

Proof. Now we have a restriction of $x_i \geq 10$ for all i, so we can define $y_i = x_i - 10$ for all i, and thus we want $y_i \geq 0$ for all i and

$$\sum_{i=1}^{77} y_i = \sum_{i=1}^{77} (x_i - 10) = 3730,$$

which means the number of ways of distribution is $\binom{3806}{3730}$.

$\circledast$

Chapter 2

Axiom of Probability

Suppose we roll a fair die 100 times and are interested in the sum of the rolls.

> **Question.** What is the sample space?

We have to define the definition of sample space first.

> **As previously seen.** The sample space is a set S of all possible outcomes, and sometimes denoted by Ω, and the events is a subset $E \subseteq S$.

Appendix