

Probability

Jonathan Gai

January 25, 2022

Contents

1	Formal Setup	1
1.1	Examples of Probability Spaces	3
1.2	Combinatorial Analysis	4
1.3	Random Walks	5
1.4	Properties of Probability Measures	6

Lecture 1: Probability Space

20 Jan. 11:00

Example. If we have a die with outcomes $1, 2, \dots, 6$.

1. $\mathbb{P}(2) = \frac{1}{6}$
2. $\mathbb{P}(\text{multiple of } 3) = \frac{2}{6} = \frac{1}{3}$
3. $\mathbb{P}(\text{pair or a multiple of } 3) = \frac{4}{6} = \frac{2}{3}$

1 Formal Setup

We try to define a probability space rigorously in this section.

Definition 1.1 (Probability Space). We have the following,

1. Sample space Ω , a set of outcomes.
2. \mathcal{F} , a collection of subsets of Ω (called events).
3. \mathcal{F} is a σ -algebra if
 - (a) **F1:** $\Omega \in \mathcal{F}$
 - (b) **F2:** if $A \in \mathcal{F}$ then $A^c \in \mathcal{F}$
 - (c) **F3:** For all countable collections $\{A_n\}$ in \mathcal{F} , $\cup_n A_n \in \mathcal{F}$.

Given σ -algebra \mathcal{F} on Ω , function $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$ is a probability measure if

1. **P1:** The probability function is nonnegative.
2. **P2:** $\mathbb{P}(\Omega) = 1$
3. **P3:** For all countable collection $\{A_n\}$ of disjoint events in \mathcal{F} , we have

$$\mathbb{P}(\cup_n A_n) = \sum_{n=1}^{\infty} \mathbb{P}(A_n).$$

Then $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space.

Problem. Why $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$, not $\mathbb{P} : \Omega \rightarrow [0, 1]$?

We will justify the definition in the following examples.

Example. When Ω is finite or countable,

1. In general: $\mathcal{F} = \mathcal{P}(\Omega)$.
2. $\mathbb{P}(2)$ is shorthand for $\mathbb{P}(\{2\})$.
3. \mathbb{P} is determined by $\mathbb{P}(\{w\}), \forall w \in \Omega$.

Remark. When Ω is uncountable, a probability space behaves differently, as shown in the following example.

Example. If $\Omega = [0, 1]$, and we want to choose a real number, all equally likely.

If $\mathbb{P}\{0\} = \alpha > 0$, then $\mathbb{P}(\{0, 1, \frac{1}{2}, \dots, \frac{1}{n}\}) = n\alpha$. This cannot happen if n large, because we would have $\mathbb{P} > 1$. So $\mathbb{P}(\{0\}) = 0$ or undefined.

Example. When Ω is infinitely countable (e.g., $\Omega = \mathbb{N}$ or $\Omega = \mathbb{Q} \cap [0, 1]$), however, it is not possible to choose uniformly. Suppose it is possible, there are two possibilities

- If $\mathbb{P}(\{\omega\}) = \alpha \quad \forall \omega \in \Omega$,
 then $\mathbb{P}(\Omega) = \sum_{\omega \in \Omega} \mathbb{P}(\{\omega\}) = \infty$. \nexists
- If $\mathbb{P}(\{\omega\}) = 0 \quad \forall \omega \in \Omega$,
 then $\mathbb{P}(\Omega) = \sum_{\omega \in \Omega} \mathbb{P}(\{\omega\}) = 0$. \nexists

So it is not possible to have one such uniform probability space. But that's fine as there exists many other interesting probability measures on a infinite countably set.

Property. From the axioms, we want to prove the following properties of a probability space.

1. $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.

Proof. A, A^c disjoint. $A \cup A^c = \Omega$. So $\mathbb{P}(A) + \mathbb{P}(A^c) = \mathbb{P}(\Omega) = 1$ ■

2. $\mathbb{P}(\emptyset) = 0$

3. If $A \subseteq B$, then $\mathbb{P}(A) \leq \mathbb{P}(B)$.

4. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

1.1 Examples of Probability Spaces

Example. Here we list some concrete examples of probability spaces.

1. Ω finite, $\Omega = \{w_1, \dots, w_n\}$, \mathcal{F} = all subsets under uniform choice.

$$\mathbb{P} : \mathcal{F} \rightarrow [0, 1], \mathbb{P}(A) = \frac{|A|}{|\Omega|}. \text{ In particular: } \mathbb{P}(\{w\}) = \frac{1}{|\Omega|} \forall w \in \Omega.$$

2. If we are choosing without replacement n indistinguishable marbles that are labelled $\{1, \dots, n\}$. Pick $k \leq n$ marbles uniformly at random.

$$\text{Here we have } \Omega = \{A \subseteq \{1, \dots, n\}, |A| = k, |\Omega| = \binom{n}{k}.$$

3. If we have a well-shuffled deck of cards, and we uniformly chose permutation of 52 cards.

$$\Omega = \{\text{all permutations of 52 cards}\}. |\Omega| = 52!.$$

Then we have

$$\mathbb{P}(\text{first three cards have the same suit}) = \frac{52 \cdot 12 \cdot 11 \cdot 49!}{52!} = \frac{22}{425}.$$

Lecture 2: Finite Probability Space

22 Jan. 11:00

Example (Coincidental Birthday). There we have n people, what is the probability that at least two share a birthday? To be precise, we first make the following assumptions,

- No leap years; (365 days in a year)
- All birthdays are equally likely.

We have the probability space

$$\Omega = \{1, \dots, 365\}^n$$

$$\mathcal{F} = \mathcal{P}(\Omega)$$

$$A = \{\text{at least 2 people share birthday}\}$$

$$A^c = \{\text{all } n \text{ birthdays are different}\}.$$

So we have the probability

$$\begin{aligned}\mathbb{P}(A^c) &= \frac{365 \times 364 \times \dots \times (365 - n + 1)}{365^n}, \\ \mathbb{P}(A) &= 1 - \frac{365 \times 364 \times \dots \times (365 - n + 1)}{365^n}.\end{aligned}$$

Remark.

- We note several special n values,

$$\begin{aligned}n = 22 & : \quad \mathbb{P}(A) \approx 0.479 \\ n = 23 & : \quad \mathbb{P}(A) \approx 0.507 \\ n \geq 366 & : \quad \mathbb{P}(A) = 1\end{aligned}$$

- The probability of birthday is not equal in real life though. It is more likely to be born about 9 months after christmas.
- Sometimes it would be easier to calculate the probability of the complement of an event.

1.2 Combinatorial Analysis

If Ω is a finite set such that $|\Omega| = n$,

Problem. How many ways to partition Ω into k disjoint subsets $\Omega_1, \dots, \Omega_k$ with $|\Omega_i| = n_i$ ($\sum_{i=1}^k n_i = n$)?

The total number of ways M is

$$\begin{aligned}M &= \binom{n}{n_1} \binom{n - n_1}{n_2} \binom{n - n_1 - n_2}{n_3} \dots \binom{n - n_1 - n_2 - \dots - n_{k-1}}{n_k} \\ &= \binom{n}{n_1} \binom{n - n_1}{n_2} \binom{n - n_1 - n_2}{n_3} \dots \binom{n}{n_k} \\ &= \frac{n!}{n_1!(n - n_1)!} \times \frac{(n - n_1)!}{n_2!(n - n_1 - n_2)!} \times \dots \times \frac{(n - n_1 - n_2 - \dots - n_{k-1})!}{n_k!0!} \\ &= \frac{n!}{n_1!n_2! \dots n_k!} \\ &= \binom{n}{n_1, n_2, \dots, n_k}\end{aligned}$$

which is called the *multinomial coefficient*, and denoted by the last term in the equations.

Remark. The ordering of the subsets do matter in this setting.

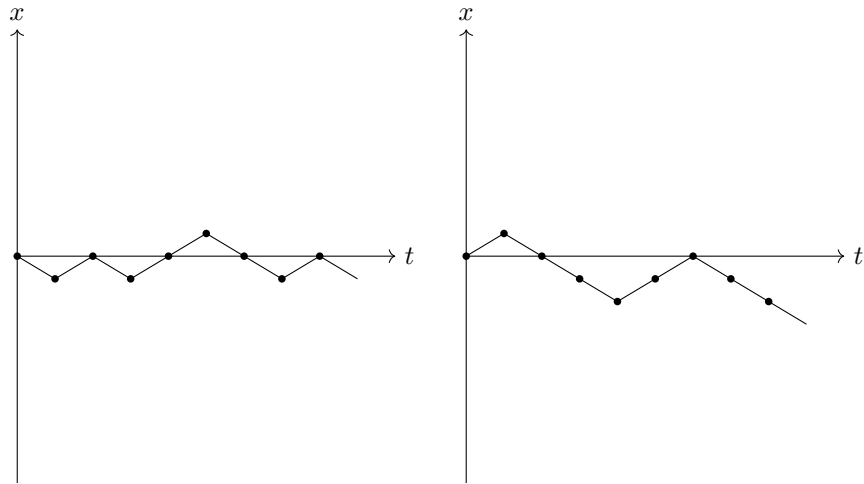


Figure 1: Random Walks

1.3 Random Walks

We have the following uniform probability space

$$\Omega = \{(x_0, x_1, \dots, x_n) \mid x_0 = 0, |x_k - x_{k-1}| = 1, k = 1, \dots, n\},$$

$$|\Omega| = 2^n.$$

Problem. What's $\mathbb{P}(x_n = 0)$ and $\mathbb{P}(x_n = n)$?

We have $\mathbb{P}(x_n = n) = \frac{1}{2^n}$.

When n is odd, $\mathbb{P}(x_n = 0) = 0$ because after every step the value changes parity. To find the probability when n is even, we need to choose $\frac{n}{2}$ ks for which $x_k = x_{k-1} + 1$, and the rest $x_k = x_{k-1} - 1$. So

$$\begin{aligned} \mathbb{P}(x_n = 0) &= 2^{-n} \binom{n}{n/2} \\ &= \frac{n!}{2^n \left[\left(\frac{n}{2}\right)!\right]}. \end{aligned}$$

Problem. What happens when n is large?

We next present Stirling's Formula, and we adopt the following notation for the time being.

Notation. If (a_n) , b_n are two sequences, we say $a_n \sim b_n$ as $n \rightarrow \infty$ if $\frac{a_n}{b_n} \rightarrow 1$ as $n \rightarrow \infty$.

Theorem 1.1 (Stirling's Formula).

$$n! \sim \sqrt{2\pi n} n^{n+\frac{1}{2}} e^{-n} \quad \text{as } n \rightarrow \infty.$$

We also have the weaker version

$$\log(n!) \sim n \log n.$$

Lecture 3

25 Jan. 11:00

Proof. We have

$$\log(n!) = \log 2 + \log 3 + \dots + \log n.$$

So

$$\begin{aligned} \int_1^n \log x dx &\leq \log(n!) \leq \int_1^{n+1} \log x dx \\ \underbrace{n \log n - n + 1}_{n \log n} &\leq \log(n!) \leq \underbrace{(n+1) \log(n+1) - n}_{n \log n}. \end{aligned}$$

$\log(n!)$ is sandwiched between the lower and upper integrals, so $\log(n!)$ must be approximately $n \log n$ as well. In this calculation, these facts helped

1. $\log x$ is increasing, so it's easier to be bounded by the integrals.
2. $\log x$ has a nice integral. So the integrals have closed forms.

■

(Ordered) Compositions

Definition 1.2. A *composition* of m with k parts is sequence (m_1, \dots, m_k) of non-negative integers with $\sum_{i=1}^k m_i = m$.

We use stars and bars. There are m stars and $k - 1$ bars, and

$$\#\text{Compositions} = \binom{m+k-1}{m}.$$

1.4 Properties of Probability Measures

Recall Definition (1.1). We prove the following properties.

Property.

1. Countable sub-additivity

Let $(A_n)_{n \geq 1}$ sequence of events in \mathcal{F} . Then

$$\mathbb{P}(\cup_{n \geq 1} A_n) \leq \sum_{n \geq 1} \mathbb{P}(A_n).$$

Proof. We rewrite $\cup_{n \geq 1}$ as a disjoint union.

Define $B_1 = A_1$ and $B_n = A_n \setminus (A_1 \cup \dots \cup A_{n-1})$.

So

- $\cup_{n \geq 1} B_n = \cup_{n \geq 1} A_n$,
- $(B_n)_{n \geq 1}$ disjoint (by construction),
- $B_n \subseteq A_n \implies \mathbb{P}(B_n) \leq \mathbb{P}(A_n)$.

And we have

$$\mathbb{P}(\cup_{n \geq 1} A_n) = \mathbb{P}(\cup_{n \geq 1} B_n) = \sum_{n \geq 1} \mathbb{P}(B_n) = \sum_{n \geq 1} \mathbb{P}(A_n).$$

■

2. Continuity $(A_n)_{n \geq 1}$ increasing sequence of events in \mathcal{F} that is $A_n \subseteq A_{n+1}$ for all n .

In fact, $\lim_{n \rightarrow \infty} \mathbb{P}(A_n) = \mathbb{P}(\cup_{n \geq 1} A_n)$.

Proof. We reuse the B_n s, and we have

- $\cup_{k=1}^n B_k = A_n$,
- $\cup_{n \geq 1} B_n = \cup_{n \geq 1} A_n$.

So we have

$$\mathbb{P}(A_n) = \sum_{k=1}^n \mathbb{P}(B_k) \rightarrow \sum_{k \geq 1} \mathbb{P}(B_k) = \mathbb{P}(\cup_{n \geq 1} B_n) = \mathbb{P}(\cup_{n \geq 1} A_n).$$

■

3. Inclusion-Exclusion Principle

Background: $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

Similarly, for $A, B, C \in \mathcal{F}$,

$$\begin{aligned} \mathbb{P}(A \cup B \cup C) &= \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(B \cap C) \\ &\quad - \mathbb{P}(C \cap A) + \mathbb{P}(A \cap B \cap C). \end{aligned}$$

The full Inclusion-Exclusion principle statement is the following. Let $A_1, \dots, A_n \in \mathcal{F}$, then

$$\begin{aligned} \mathbb{P}(\cup_{i=1}^n A_i) &= \sum_{i=1}^n \mathbb{P}(A_i) - \sum_{1 \leq i_1 < i_2 \leq n} \mathbb{P}(A_{i_1} \cap A_{i_2}) + \dots \\ &\quad + (-1)^{n+1} \mathbb{P}(A_1 \cap \dots \cap A_n) \\ &= \sum_{\substack{I \subseteq \{1, \dots, n\} \\ I \neq \emptyset}} (-1)^{|I|+1} \mathbb{P}(\cap_{i \in I} A_i). \end{aligned}$$