

WuS - Complete Summary

Ruben Schenk, ruben.schenk@inf.ethz.ch

March 11, 2022

1 Introduction

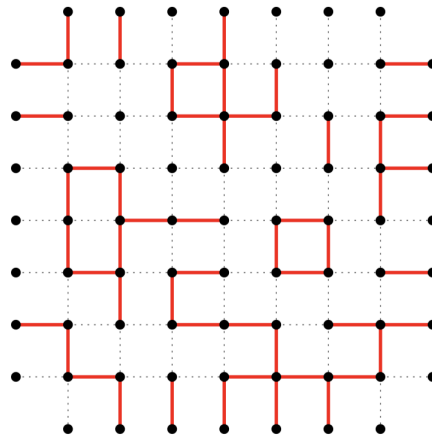
1.1 Percolation Theory

1.1.1 Overview

In physics and mathematics, **percolation theory** describes the behavior of clustered components in random networks. The common intuition is movement and filtering of fluids through porous materials, for example, filtration of water through soil and permeable rocks. In a network, let each node be a cell through which a fluid-like substance may transit to other cells. A network, i.e. a grid, then is a sponge-like substance and percolation is the determination of whether a substance introduced at one cell will reach the other side of the network (or grid).

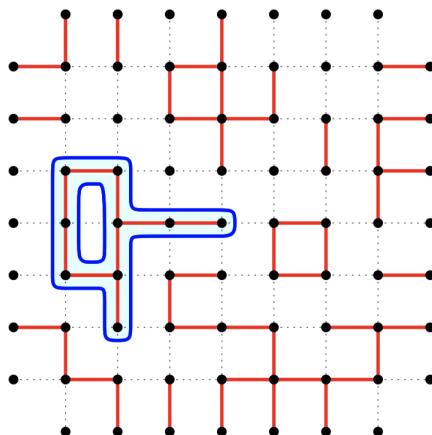
1.1.2 Percolation in a Box

Imagine a box (or grid) with vertices $V = \{-n, \dots, n\}^2$ and edges $E = \{e_1, \dots, e_N\}$. We introduce parameter p , with $0 \leq p \leq 1$. p denotes the probability that an edge e is *open* ($X_e = 1$). In other words, an edge e is *closed* ($X_e = 0$) with probability $1 - p$. The corresponding model could look something like this:



Note: If an edge is colored red, it means that it's open.

We denote an **open path** as a path consisting of open edges. A **cluster** is the connected component of $(V, \{e : X_e = 1\})$. The following figure shows an example of a cluster (marked in blue):



Theorem [Kesten, 1980]: For the percolation with parameter p we have:

$$\lim_{n \rightarrow \infty} \mathbb{P}[\bullet] = \begin{cases} 0, & \text{if } p < \frac{1}{2}, \\ 1, & \text{if } p > \frac{1}{2}. \end{cases}$$

where $\mathbb{P}[\bullet]$ denotes the probability that there exists an open path from the top to the bottom in an $n \times n$ box. Similarly, for the percolation with parameter p we have:

$$\mathbb{P}[\exists \text{ an infinite cluster}] = \begin{cases} 0, & \text{if } p < \frac{1}{2}, \\ 1, & \text{if } p > \frac{1}{2}. \end{cases}$$

1.2 Introduction to Probability

Probability is a mathematical language describing systems involving randomness. Probabilities are used for:

- *Describe random experiments* in the real world, such as coin flips, dice rolling, etc.
- *Express uncertainty.* For example, when a machine performs a measurement, the value is rarely exact. One may use probability theory in this context by saying that the value obtained is equal to the real value plus some small random error.
- *Decision-making.* Probability theory can be used to describe a system when only part of the information is known.
- *Randomized algorithms* in computer science. Sometimes, it is more efficient to add some randomness to perform an algorithm.
- *Simplify complex systems.* Examples include water molecules in water, cars on the highway, etc.

The **goal** of probability theory is to establish general theorems which describe the behavior of multiple random experiments. Example:

Theorem [Law of large numbers]:

$$X_i = \begin{cases} 0, & i^{th} \text{ throw is head,} \\ 1, & i^{th} \text{ throw is number.} \end{cases}$$

It holds, that:

$$\lim_{n \rightarrow \infty} \frac{X_1 + \dots + X_n}{n} = \frac{1}{2}.$$

2 Mathematical Framework

2.1 Probability Space

2.1.1 Sample Space

Assume we want to model a random experiment. The first mathematical object needed is the set of all possible outcomes of the experiment, denoted by Ω .

The set Ω is called the **sample space**. An element $\omega \in \Omega$ is called an **outcome** (or *elementary experiment*).

Example: If we throw a die, we have the following sample space:

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

2.1.2 Events

Previously, the set of **events** was always $\mathcal{P}(\Omega)$. In this class, we will work with more general sets of events $\mathcal{F} \subset \mathcal{P}(\Omega)$, called sigma algebras.

Definition: A **sigma-algebra** is a subset $\mathcal{F} \subset \mathcal{P}(\Omega)$ satisfying the following properties:

1. $\Omega \in \mathcal{F}$
2. $A \in \mathcal{F} \implies A^C \in \mathcal{F}$
3. $A_1, A_2, \dots \in \mathcal{F} \implies \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$

Example: Following are some (non-) examples of sigma-algebras for $\Omega = \{1, 2, 3, 4, 5, 6\}$:

- $\mathcal{F} = \{\emptyset, \{1, 2, 3, 4, 5, 6\}\}$ is a sigma-algebra.
- $\mathcal{F} = \{\emptyset, \{1, 2\}, \{3, 4, 5, 6\}, \{1, 2, 3, 4, 5, 6\}\}$ is a sigma-algebra.
- $\mathcal{F} = \{\{1, 2, 3, 4, 5, 6\}\}$ is not a sigma-algebra because P2 is not satisfied.
- $\mathcal{F} = \{\emptyset, \{1, 2, 3\}, \{4, 5, 6\}, \{1\}, \{2, 3, 4, 5, 6\}, \Omega\}$ is not a sigma-algebra because P3 is not satisfied.

2.1.3 Probability Measure

Definition: Let Ω be a sample space, let \mathcal{F} be a sigma-algebra. A **probability measure** on (Ω, \mathcal{F}) is a map

$$\begin{aligned} \mathbb{P} : \mathcal{F} &\rightarrow [0, 1] \\ A &\mapsto \mathbb{P}[A] \end{aligned}$$

that satisfies the following two properties:

- **P1.** $\mathbb{P}[\Omega] = 1$.
- **P2. (countable additivity)** $\mathbb{P}[A] = \sum_{i=1}^{\infty} \mathbb{P}[A_i]$ if $A = \bigcup_{i=1}^{\infty} A_i$ (*disjoint union*).

2.1.4 Notion of Probability Space

Definition: Let Ω be a sample space, \mathcal{F} a sigma-algebra, and \mathbb{P} a probability measure. The triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called a **probability space**.

2.2 Examples of Probability Space

2.2.1 Example with Ω Finite

We discuss a particular type of probability spaces where the sample space Ω is an arbitrary **finite** set, and all the outcomes have the **same** probability $p_\omega = \frac{1}{|\Omega|}$.

Definition: Let Ω be a finite sample space. The **Laplace model** on Ω is the triple $(\Omega, \mathcal{F}, \mathbb{P})$, where:

- $\mathcal{F} = \mathcal{P}(\Omega)$,
- $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$ is defined by

$$\forall A \in \mathcal{F} \quad \mathbb{P}[A] = \frac{|A|}{|\Omega|}$$

Example: We consider $n \geq 3$ points on a circle, from which we select 2 at random. What is the probability that these two points selected are neighbors? We consider the Laplace model one

$$\Omega = \{E \subset \{1, 2, \dots, n\} : |E| = 2\}.$$

The event "the two points of E are neighbors" is given by

$$A = \{\{1, 2\}, \{2, 3\}, \dots, \{n-1, n\}, \{n, 1\}\}$$

and we have

$$\mathbb{P}[A] = \frac{|A|}{|\Omega|} = \frac{n}{\binom{n}{2}} = \frac{2}{n-1}.$$

2.2.2 Example with Ω Infinite Countable

Example: We throw a biased coin multiple times, at each throw, the coin falls on head with probability p , and it falls on tail with probability $1 - p$ (p is a fixed parameter in $[0, 1]$). We stop at the first time we see a tail. The probability that we stop exactly at time k is given by

$$p_k = p^{k-1}(1 - p).$$

For this experiment, one possible probability space is given by:

- $\Omega = \mathbb{N} \setminus \{0\} = \{1, 2, 3, \dots\}$
- $\mathcal{F} = \mathcal{P}(\Omega)$
- for $A \in \mathcal{F}$, $\mathbb{P}[A] = \sum_{k \in A} p_k$

2.3 Properties of Events

2.3.1 Operations on Events and Interpretation

The following propositions asserts that the different well-known set operations are allowed.

Proposition (Consequences of the definition): Let \mathcal{F} be a sigma-algebra on Ω . We have:

- **P4.** $\emptyset \in \mathcal{F}$
- **P5.** $A_1, A_2, \dots \in \mathcal{F} \implies \bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$
- **P6.** $A, B \in \mathcal{F} \implies A \cup B \in \mathcal{F}$
- **P7.** $A, B \in \mathcal{F} \implies A \cap B \in \mathcal{F}$

A short summary of the common set-operations is given below:

- A^C : A does not occur.
- $A \cap B$: A and B occur.
- $A \cup B$: A or B occurs
- $A \Delta B$: one and only one of A or B occurs
- $A \subset B$: If A occurs, then B occurs
- $A \cap B = \emptyset$: A and B cannot occur at the same time
- $\Omega = A_1 \cup A_2 \cup A_3$ with A_1, A_2, A_3 pairwise disjoint: for each outcome ω , one and only one of the events A_1, A_2, A_3 is satisfied.

2.4 Properties of Probability Measures

2.4.1 Direct Consequences of the Definition

Proposition: Let \mathbb{P} be an arbitrary measure on (Ω, \mathcal{F}) . We have:

- **P3.** $\mathbb{P}[\emptyset] = 0$.
- **P4. (additivity)** Let $k \geq 1$. let A_1, \dots, A_k be k pairwise disjoint events, then $\mathbb{P}[A_1 \cup \dots \cup A_k] = \mathbb{P}[A_1] + \dots + \mathbb{P}[A_k]$.
- **P5.** Let A be an event, then $\mathbb{P}[A^C] = 1 - \mathbb{P}[A]$.
- **P6.** If A and B are two events (not necessarily disjoint), then $\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A \cap B]$.

2.4.2 Useful Inequalities

Proposition (Monotonicity): Let $A, B \in \mathcal{F}$, then

$$A \subset B \implies \mathbb{P}[A] \leq \mathbb{P}[B].$$

Proposition (Union bound): Let A_1, A_2, \dots be a sequence of events (not necessarily disjoint), then we have

$$\mathbb{P}\left[\bigcup_{i=1}^{\infty} A_i\right] \leq \sum_{i=1}^{\infty} \mathbb{P}[A_i].$$

Remark: The union bound also applies to a *finite* collection of events.

2.4.3 Continuity Properties of Probability Measures

Proposition: Let (A_n) be an increasing sequence of events (i.e. $A_n \subset A_{n+1}$ for every n). then

$$\lim_{n \rightarrow \infty} \mathbb{P}[A_n] = \mathbb{P}\left[\bigcup_{n=1}^{\infty} A_n\right]. \quad (\text{increasing limit})$$

Let (B_n) be a decreasing sequence of events (i.e. $B_n \supset B_{n+1}$ for every n). Then

$$\lim_{n \rightarrow \infty} \mathbb{P}[B_n] = \mathbb{P}\left[\bigcap_{n=1}^{\infty} B_n\right]. \quad (\text{decreasing limit})$$

Remark: By monotonicity, we have $\mathbb{P}[A_n] \leq \mathbb{P}[A_{n+1}]$ and $\mathbb{P}[B_n] \geq \mathbb{P}[B_{n+1}]$ for every n . Hence the limits in the proposition are well defined as monotone limits.

2.5 Conditional Probabilities

Definition (Conditional probability): Let $(\Omega, \mathcal{F}, \mathbb{P})$ be some probability space. Let A, B be two events with $\mathbb{P}[B] > 0$. The **conditional probability of A given B** is defined by

$$\mathbb{P}[A | B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}.$$

Remark: $\mathbb{P}[B | B] = 1$.

Proposition: Let $\Omega, \mathcal{F}, \mathbb{P}$ be some probability space. Let B be an event with positive probability. Then $\mathbb{P}[\cdot | B]$ is a probability measure on Ω .

Proposition (Formula of total probability): Let B_1, \dots, B_n be a partition of the sample space Ω with $\mathbb{P}[B_i] > 0$ for every $1 \leq i \leq n$. Then, one has

$$\forall A \in \mathcal{F} : \mathbb{P}[A] = \sum_{i=1}^n \mathbb{P}[A | B_i] \mathbb{P}[B_i].$$

Here, a *partition* B_i is such that $\Omega = B_1 \cup \dots \cup B_n$ and the events are pairwise disjoint.

Proposition (Bayes formula): Let $B_1, \dots, B_n \in \mathcal{F}$ be a partition of Ω with $\mathbb{P}[B_i] > 0$ for every i . For every event A with $\mathbb{P}[A] > 0$, we have

$$\forall i = 1, \dots, n : \mathbb{P}[B_i | A] = \frac{\mathbb{P}[A | B_i] \cdot \mathbb{P}[B_i]}{\sum_{j=1}^n \mathbb{P}[A | B_j] \cdot \mathbb{P}[B_j]}.$$

2.6 Independence

2.6.1 Independence of Events

Definition (Independence of two events): Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Two events A and B are said to be **independent** if

$$\mathbb{P}[A \cap B] = \mathbb{P}[A] \cdot \mathbb{P}[B].$$

Remark: If $\mathbb{P}[A] \in \{0, 1\}$, then A is independent of every event, i.e. $\forall B \in \mathcal{F} : \mathbb{P}[A \cap B] = \mathbb{P}[A] \cdot \mathbb{P}[B]$. Furthermore we might also state, that A is independent of B if and only if A is independent of B^C .

Proposition: Let $A, B \in \mathcal{F}$ be two events with $\mathbb{P}[A], \mathbb{P}[B] > 0$. Then the following are equivalent:

- $\mathbb{P}[A \cap B] = \mathbb{P}[A] \cdot \mathbb{P}[B]$ (A and B are independent)
- $\mathbb{P}[A | B] = \mathbb{P}[A]$ (the occurrence of B has no influence on A)
- $\mathbb{P}[B | A] = \mathbb{P}[B]$ (the occurrence of A has no influence on B)

Definition: Let I be an arbitrary set of indices. A collection of events $(A_i)_{i \in I}$ is said to be **independent** if

$$\forall J \subset I \text{ infinite} : \mathbb{P}\left[\bigcap_{j \in J} A_j\right] = \prod_{j \in J} \mathbb{P}[A_j].$$