**Mathematics for Computer Science (30470023)** Spring 2020

# Lecture 1 — Feb 17, 2020

*Instructor: Prof. Andrew C. Yao* *Scribes: Tianqi Yang*

## 1 Course information

### Grading

- Homework: 50%
- Mid-term Exam: 20%
- Final Exam: 30%

### Topics

- (Discrete) basic probability theory: $\sim 4$ weeks
- Graphs & Combinatorics: $\sim 5$ weeks
- Complexity & Cryptography (optional): $\sim 3$ weeks

## 2 Overview

In this first lecture we introduce discrete probability theory. We start from the definition of probability spaces and events, and then discuss some examples including the birthday paradox and Monty Hall problem. We apply the birthday paradox to the problem of allocating student IDs, and argue that random allocation needs lots of bits to avoid collisions.

## 3 Discrete probability theory

> *Fear? What should a man fear? It's all chance, chance rules our lives. Not a man on earth can see a day ahead, groping through the dark.*
>
> — Sophocles, *Oedipus Rex*

Our world is full of randomness. Take for example, when we roll two dices, it is almost impossible for us to tell the result with certainty in advance. However, we are still interested in how likely it is that the sum of the two outcomes is exactly 8. To answer this question, we need a measurement of such likelihood, which we call the *probability* of this event. To give a formal definition to this concept, we need to first introduce the *probability space*.

## 3.1 Probability space

**Definition 3.1 (Discrete probability space).** A discrete probability space is an ordered pair $\mathbb{P} = (\mathcal{U}, p)$ consisting of

- A *non-empty finite* set $\mathcal{U}$ describing all the possible outcomes of an execution. For every instance, the outcome must be an element in this set. This set is often called the sample space, ground set or universe of the probability space.
- A function $p : \mathcal{U} \to [0, 1]$ satisfying $\sum_{u \in \mathcal{U}} p(u) = 1$, describing the probability that an instance produces this outcome for any element in the sample space. ♢

**Definition 3.2 (Event).** An event $T$ is a subset of $\mathcal{U}$ (i.e., $T \subseteq \mathcal{U}$). The probability of the event $T$ is

$$\Pr\{T\} = \sum_{u \in T} p(u). \tag{1}$$

♢

**Remark 3.3.** We only talk about *discrete* probability spaces with a *finite* sample space in this course. In general, the sample space of a discrete probability space can also be *countable*. The above definition can be easily be generalized to this case.

In some sense, an event is a collection of possible outcomes. We say the event happens if the actual outcome is in it. Thus, the probability of it is exactly the sum over the probability of each individual possible outcome in this event. Sometimes, an event can be described by some property, i.e., the event can be defined as the set of possible outcomes that satisfy the property. However, not all events have a natural property.

**Example 3.4.** Take the previous example of throwing two *unbiased* dice. Let us resolve the probability that the sum of them is exactly 8. We first need to define the probability space $\mathbb{P} = (\mathcal{U}, p)$. Suppose that the outcomes of the two dice are $i$ and $j$ respectively, where $1 \le i, j \le 6$, then $(i, j)$ describes a possible outcome of the whole system. This suggests that we can define

$$\mathcal{U} = \{(i, j) \mid 1 \le i, j \le 6\}.$$

Since the dice are unbiased, the probability of any outcome $x \in \mathcal{U}$ is

$$p(x) = \frac{1}{|\mathcal{U}|} = \frac{1}{36}.$$

We need to calculate the probability that $i + j = 8$. Thus, the event $T$ is defined as the collection of possible outcomes satisfying $i + j = 8$, which is

$$\begin{aligned} T &= \{(i, j) \mid 1 \le i, j \le 6, i + j = 8\} \\ &= \{(2, 6), (3, 5), (4, 4), (5, 3), (6, 2)\}. \end{aligned}$$

Thus,

$$\Pr\{T\} = \sum_{u \in T} p(u) = \frac{|T|}{|\mathcal{U}|} = \frac{5}{36} \approx 14\%.$$

## 3.2 Birthday paradox

Let us consider the probability that there exist two out of 55 students in Yao Class having the same birthday, assuming that there are 365 days in a year. Intuitively, the probability should be considerably small, since 55 students only fill a small proportion of all 365 days in a year. However, this turns out to be not the case. We now examine this probability carefully.

The universe of our probability space is all the possible assignments of a birthday to each student. Formally,

$$\mathcal{U} = \{(i_1, i_2, \ldots, i_{55}) \mid \forall 1 \le j \le 55,\ 1 \le i_j \le 365\}.$$

Suppose that the probability is uniform, then each $x \in \mathcal{U}$ has $p(x) = \frac{1}{|\mathcal{U}|}$. We now consider the cases when there are two students share the same birthday. The corresponding event is

$$T = \{(i_1, i_2, \ldots, i_{55}) \in \mathcal{U} \mid \exists j_1 \neq j_2,\ i_{j_1} = i_{j_2}\}.$$

We count on the complement of $T$, which represents the cases that all 55 students have distinct birthdays. In this case, the number of ways to assign a birthday to the first student is 365. Then the birthday of the second student has 364 possible choices, since it cannot coincide with the first. Similarly, there are $365 - i + 1$ choices for the $i$-th student. Therefore,

$$\Pr\{T\} = \frac{|T|}{|\mathcal{U}|} = 1 - \frac{|U \setminus T|}{|\mathcal{U}|} = 1 - \frac{365 \cdot 364 \cdot 363 \cdot \cdots \cdot 311}{365^{55}}.$$

This formula is intractable to evaluate directly, so we need some appropriate approximation. Notice that when $x$ is close to 0, $e^x \approx 1 + x$. Substituting this in the above equation gives us

$$\begin{aligned}
\Pr\{T\} &= 1 - \left(1 - \frac{1}{365}\right)\left(1 - \frac{2}{365}\right) \cdot \left(1 - \frac{54}{365}\right) \\
&= 1 - e^{-\frac{1}{365}} e^{-\frac{2}{365}} \cdot e^{-\frac{54}{365}} \\
&= 1 - e^{-\frac{1}{365}(1+2+\cdots+54)} \\
&\approx 98.3\%.
\end{aligned}$$

Although some error might be introduced by the approximation, this is acceptable. In fact, the precise result is 98.6%, which is not far from our rough one.

**Remark 3.5.** The approximation method we use here is an example of back-of-the-envelope calculation, which stands for the ways of making rough evaluation of complex equations available with only simple calculations. See Wikipedia if you wish to learn more: https://en.wikipedia.org/wiki/Back-of-the-envelope_calculation.

We can see that this probability is extremely high, which is counter-intuitive in some sense. This is called the birthday paradox. In fact, even 23 students can make this probability exceed 50%.

## 3.3 Application on random student IDs

Now we assume that in the university, student IDs are allocated randomly. Specifically, each student gets a random $m$-bit number as their student ID. How many bits are needed to ensure no collisions?

Let us suppose that there are $n = 40,000$ students. Applying similar arguments, since there are $M = 2^m$ possible $m$-bit numbers, the probability of collision is

$$\Pr\{\text{collision}\} = 1 - \left(1 - \frac{1}{M}\right)\left(1 - \frac{2}{M}\right)\ldots\left(1 - \frac{n-1}{M}\right) \approx 1 - e^{-\frac{n^2}{2M}} \approx \frac{n^2}{2M}.$$

If we need to ensure that the probability is not greater than $10^{-20}$, then roughly 90 bits are needed. So we can see that due to the birthday paradox, random allocation is really a bad idea.

## 4 Monty Hall problem

The Monty Hall problem is a famous puzzle based on an American television show. In this problem, there are 3 closed doors, with a gift behind one of them. The door with gift is pre-determined by the host uniformly randomly. Now the guest should make a guess, say door NO 1 without loss of generality. Then the host will open another door different from his choice, say door NO 2, which does not have a gift. Now the guest has a chance to change his decision. He can choose to switch to the other closed door NO 3, or stick to his initial choice. The objective of the guest is to find out the door with the gift.

The problem is that whether or not he should switch his choice. We can analyze this with tedious calculation, but here is a sophisticated method. Suppose that the initial guess is correct, which has probability 1/3, then sticking to the initial choice can give the correct answer. However, if the guess is incorrect initially, which has probability 2/3, the door which is neither chosen by the initial guess nor opened by the host is the correct answer. Thus, switching can give the guest a higher probability to win.

**Acknowledgements:** I would like to thank Jiatu Li, Mengdi Wu and Boyang Chen who proofread this note, and gave tons of useful advice.