Lecture 03: Probability Basics

Overview

In today's lecture we shall continue with the basics of probability

Events

1. Let $\mathbb{X}$ be a random variable over the sample space Ω
2. An event E is a subset of Ω
3. The probability that the random variable $\mathbb{X}$ takes values in the set E is
$$\sum_{x \in \Omega \colon x \in E} \mathbb{P}[\mathbb{X} = x]$$
4. We succinctly represent this probability by the following expression
$$\mathbb{P}[\mathbb{X} \in E]$$

Chain Rule for Events

- Let $\mathbb{X}$ be a random variable over the sample space Ω
- Let $E_1, E_2, \ldots, E_k$ be events
- The probability that the random variable simultaneously satisfies all the events is represented by

$$\mathbb{P}\left[\mathbb{X} \in E_1, \ldots, \mathbb{X} \in E_k\right]$$

- The chain rule states that, the probability is equal to the following expression

$$\begin{aligned}
&\mathbb{P}\left[\mathbb{X} \in E_1\right] \\
&\times \mathbb{P}\left[\mathbb{X} \in E_2 | \mathbb{X} \in E_1\right] \\
&\times \mathbb{P}\left[\mathbb{X} \in E_3 | \mathbb{X} \in E_1, \mathbb{X} \in E_2\right] \\
&\vdots \\
&\times \mathbb{P}\left[\mathbb{X} \in E_k | \mathbb{X} \in E_1, \ldots, \mathbb{X} \in E_{k-1}\right]
\end{aligned}$$

Example Problem I

Consider the following experiment described in English

Experiment

Suppose there are n bins. Suppose we throw m balls into n bins such that each balls is thrown uniformly and independently at random into the n possible bins. What is the probability that all balls land in distinct bins?

Example Problem II

Let us formalize this problem

- The bins are numbered $\{1, \ldots, n\}$, represented by $[n]$
- Let $\mathbb{X} = (\mathbb{X}_1, \ldots, \mathbb{X}_m)$ be a joint distribution over the sample space $[n]^{\otimes m}$, where $\mathbb{X}_i$ represents the bin that the i-th ball falls into
- The random variable $\mathbb{X}_i$ is independent of the random variable

$$(\mathbb{X}_1, \ldots, \mathbb{X}_{i-1}, \mathbb{X}_{i+1}, \ldots, \mathbb{X}_m)$$

- The random variable $\mathbb{X}_i$ is a uniform random variable over $[n]$, i.e., for any $x \in [n]$, we have

$$\mathbb{P}[\mathbb{X}_i = x] = 1/n$$

Example Problem III

- Let E be the set of all $(x_1, \ldots, x_n) \in [n]^{\otimes m}$ such that all x_is are distinct
- We are interested in computing

$$\mathbb{P}\left[\mathbb{X} \in E\right]$$

Technique 1.

- Observe that $|\Omega| = n^m$
- For any $x \in \Omega$, we have $\mathbb{P}[\mathbb{X} = x] = 1/n^m$
- So, $\mathbb{P}[\mathbb{X} \in E] = |E|/n^m$
- We can see that $|E| = n(n-1)\cdots(n-m+1)$ (This is because the first coordinate can take n possible values, the second coordinate can take $(n-1)$ possible values, and so on.)
- So, we have

$$\mathbb{P}[\mathbb{X} \in E] = \frac{n(n-1)\cdots(n-m+1)}{n^m}$$

Example Problem V

Technique 2.

- For $i \in \{1, \ldots, m\}$ define the event E_i as follows
  "Ball i falls in a different bin from the balls $\{1, 2, \ldots, i-1\}$"
- We are interested in the probability

$$\mathbb{P}[\mathbb{X} \in E] = \mathbb{P}[\mathbb{X} \in E_1, \ldots, \mathbb{X} \in E_m]$$

- The main observation is the following

$$\mathbb{P}\left[\mathbb{X} \in E_i | \mathbb{X} \in E_1, \ldots, \mathbb{X} \in E_{i-1}\right] = \left(1 - \frac{i-1}{n}\right)$$

- By chain rule, we have

$$\mathbb{P}[\mathbb{X} \in E] = 1 \cdot \left(1 - \frac{1}{n}\right) \cdot \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{m-1}{n}\right)$$

1. Let $\mathbb{X}$ be a random variable over the sample space Ω
2. Let E_1 and E_2 be two events

Lemma

If $E_1 \subseteq E_2$, then

$$\mathbb{P}\left[\mathbb{X} \in E_1\right] \leqslant \mathbb{P}\left[\mathbb{X} \in E_2\right]$$

Proof.

$$\begin{aligned}
\mathbb{P}\left[\mathbb{X} \in E_2\right] &= \sum_{x \in \Omega:\ x \in E_2} \mathbb{P}\left[\mathbb{X} = x\right] \\
&= \sum_{x \in \Omega:\ x \in E_1} \mathbb{P}\left[\mathbb{X} = x\right] + \sum_{x \in \Omega:\ x \in E_2 \setminus E_1} \mathbb{P}\left[\mathbb{X} = x\right] \\
&\geqslant \mathbb{P}\left[\mathbb{X} \in E_1\right]
\end{aligned}$$

□

Lemma

$$\mathbb{P}\left[\mathbb{X} \in E_1, \mathbb{X} \in E_2\right] \leqslant \mathbb{P}\left[\mathbb{X} \in E_1\right]$$

Proof.

$$\begin{aligned}\mathbb{P}\left[\mathbb{X} \in E_1, \mathbb{X} \in E_2\right] &= \mathbb{P}\left[\mathbb{X} \in E_1\right] \cdot \mathbb{P}\left[\mathbb{X} \in E_1 \middle| \mathbb{X} \in E_2\right] \\ &\leqslant \mathbb{P}\left[\mathbb{X} \in E_1\right] \cdot 1 \\ &= \mathbb{P}\left[\mathbb{X} \in E_1\right]\end{aligned}$$

□

Probability Inequalities III

Lemma (Union Bound)

$$\mathbb{P}\left[\mathbb{X} \in E_1 \cup E_2\right] \leqslant \mathbb{P}\left[\mathbb{X} \in E_1\right] + \mathbb{P}\left[\mathbb{X} \in E_2\right]$$

Proof.

$$\begin{aligned}
\mathbb{P}\left[\mathbb{X} \in E_1 \cup E_2\right] &= \sum_{x \in \Omega\colon x \in E_1 \cup E_2} \mathbb{P}\left[\mathbb{X} = x\right] \\
&= \sum_{x \in \Omega\colon x \in E_1} \mathbb{P}\left[\mathbb{X} = x\right] + \sum_{x \in \Omega\colon x \in E_2 \setminus E_1} \mathbb{P}\left[\mathbb{X} = x\right] \\
&= \mathbb{P}\left[\mathbb{X} \in E_1\right] + \mathbb{P}\left[\mathbb{X} \in E_2 \setminus E_1\right] \\
&\leqslant \mathbb{P}\left[\mathbb{X} \in E_1\right] + \mathbb{P}\left[\mathbb{X} \in E_2\right]
\end{aligned}$$

□

Expected Outcome

- Suppose $\Omega \subseteq \mathbb{R}$
- For the purposes of this course we shall restrict to discrete Ω (for example, the set of all natural numbers)
- The expected outcome of the random variable $\mathbb{X}$, represented by $\mathbb{E}[\mathbb{X}]$, is the following quantity

$$\mathbb{E}[\mathbb{X}] = \sum_{x \in \Omega} x \cdot \mathbb{P}[\mathbb{X} = x]$$

For example: Suppose $\mathbb{X}$ is a random variable over the sample space $\mathbb{N}$ (the set of all natural numbers). The probability $\mathbb{P}[\mathbb{X} = i] = 1/2^i$.
So, the expected outcome is defined to be

$$\mathbb{E}[\mathbb{X}] = \sum_{x \in \Omega} x\mathbb{P}[\mathbb{X} = x] = \sum_{x=1}^{\infty} x \cdot \frac{1}{2^x} = 2$$

Linearity of Expectation

Theorem (Linearity of Expectation)

Suppose $(\mathbb{X}_1, \mathbb{X}_2)$ *is a joint distribution over the sample space* $\Omega_1 \times \Omega_2$ *The following holds.*

$$\mathbb{E}[\mathbb{X}_1 + \mathbb{X}_2] = \mathbb{E}[\mathbb{X}_1] + \mathbb{E}[\mathbb{X}_2]$$

We emphasize that this result holds even if $\mathbb{X}_1$ and $\mathbb{X}_2$ are correlated!

Think: The proof is left as an exercise.

Indicator Variables

- Let $\mathbb{X}$ be a random variable over the sample space Ω
- Let $E \subseteq \Omega$ be an event
- Let $f \colon \Omega \to \{0, 1\}$ be the following function

$$f(x) = \begin{cases} 1, & x \in E \\ 0, & x \notin E \end{cases}$$

- The random variable $f(\mathbb{X})$ is referred to as the "indicator variable for the event E"
- It is represented by $\mathbf{1}_{\{E\}}$

Lemma

$$\mathbb{E}\left[\mathbf{1}_{\{E\}}\right] = \mathbb{P}\left[\mathbb{X} \in E\right]$$

The proof is left as an exercise.