Lecture 03: Probability Basics
In today’s lecture we shall continue with the basics of probability
1. Let $X$ be a random variable over the sample space $\Omega$.

2. An event $E$ is a subset of $\Omega$.

3. The probability that the random variable $X$ takes values in the set $E$ is

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

4. We succinctly represent this probability by the following expression

$$P[X \in E]$$
Chain Rule for Events

- Let \( X \) be a random variable over the sample space \( \Omega \)
- Let \( E_1, E_2, \ldots, E_k \) be events
- The probability that the random variable simultaneously satisfies all the events is represented by
  \[
  P[X \in E_1, \ldots, X \in E_k]
  \]
- The chain rule states that, the probability is equal to the following expression
  \[
  P[X \in E_1] \\
  \times P[X \in E_2 | X \in E_1] \\
  \times P[X \in E_3 | X \in E_1, X \in E_2] \\
  \vdots \\
  \times P[X \in E_k | X \in E_1, \ldots, X \in E_{k-1}]
  \]
Consider the following experiment described in English

**Experiment**

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

- The bins are numbered \( \{1, \ldots, n\} \), represented by \([n]\)
- Let \( X = (X_1, \ldots, X_m) \) be a joint distribution over the sample space \([n]^{\otimes m}\), where \( X_i \) represents the bin that the \( i \)-th ball falls into
- The random variable \( X_i \) is independent of the random variable \((X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_m)\)
- The random variable \( X_i \) is a uniform random variable over \([n]\), i.e., for any \( x \in [n] \), we have
  \[
  P[X_i = x] = \frac{1}{n}
  \]
Let $E$ be the set of all $(x_1, \ldots, x_n) \in [n]^{\otimes m}$ such that all $x_i$s are distinct.

We are interested in computing

$$\mathbb{P} [X \in E]$$
Example Problem IV

Technique 1.

- Observe that $|\Omega| = n^m$
- For any $x \in \Omega$, we have $P[X = x] = 1/n^m$
- So, $P[X \in E] = |E| / n^m$
- 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

$$P[X \in E] = \frac{n(n-1) \cdots (n-m+1)}{n^m}$$
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
  \[
P[X \in E] = P[X \in E_1, \ldots, X \in E_m]\]
- The main observation is the following
  \[
P[X \in E_i | X \in E_1, \ldots, X \in E_{i-1}] = \left(1 - \frac{i-1}{n}\right)\]
- By chain rule, we have
  \[
P[X \in E] = 1 \cdot \left(1 - \frac{1}{n}\right) \cdot \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{m-1}{n}\right)\]
1. Let $X$ be a random variable over the sample space $\Omega$.
2. Let $E_1$ and $E_2$ be two events.

**Lemma**

If $E_1 \subseteq E_2$, then

$$P[X \in E_1] \leq P[X \in E_2]$$

**Proof.**

$$P[X \in E_2] = \sum_{x \in \Omega: x \in E_2} P[X = x]$$

$$= \sum_{x \in \Omega: x \in E_1} P[X = x] + \sum_{x \in \Omega: x \in E_2 \setminus E_1} P[X = x]$$

$$\geq P[X \in E_1]$$
Probability Inequalities II

Lemma

\[ P[X \in E_1, X \in E_2] \leq P[X \in E_1] \]

Proof.

\[
P[X \in E_1, X \in E_2] = P[X \in E_1] \cdot P[X \in E_1 | X \in E_2] \\
\leq P[X \in E_1] \cdot 1 \\
= P[X \in E_1]
\]
Lemma (Union Bound)

\[ P[X \in E_1 \cup E_2] \leq P[X \in E_1] + P[X \in E_2] \]

Proof.

\[ P[X \in E_1 \cup E_2] = \sum_{x \in \Omega: x \in E_1 \cup E_2} P[X = x] \]

\[ = \sum_{x \in \Omega: x \in E_1} P[X = x] + \sum_{x \in \Omega: x \in E_2 \setminus E_1} P[X = x] \]

\[ = P[X \in E_1] + P[X \in E_2 \setminus E_1] \]

\[ \leq P[X \in E_1] + P[X \in E_2] \]
Expected Outcome

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

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

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

$$E[X] = \sum_{x \in \Omega} x \cdot P[X = x] = \sum_{x=1}^{\infty} x \cdot \frac{1}{2^x} = 2$$
Theorem (Linearity of Expectation)

Suppose $(X_1, X_2)$ is a joint distribution over the sample space $\Omega_1 \times \Omega_2$. The following holds.

$$\mathbb{E}[X_1 + X_2] = \mathbb{E}[X_1] + \mathbb{E}[X_2]$$

We emphasize that this result holds even if $X_1$ and $X_2$ are correlated!

Think: The proof is left as an exercise.
Let $X$ be a random variable over the sample space $\Omega$
Let $E \subseteq \Omega$ be an event
Let $f : \Omega \rightarrow \{0, 1\}$ be the following function

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

The random variable $f(X)$ is referred to as the “indicator variable for the event $E$”
It is represented by $1\{E\}$

**Lemma**

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

The proof is left as an exercise.