Chapter 7

Probability.

7.1 Probability Spaces.

Definition 7.1.1. *A probability space is defined by a tuple (Ω, P), where:*

1. *Ω is a set, called the sample space. Any element $\omega \in \Omega$ is an atomic event. Conceptually, we think of atomic events as possible outcomes of our experiment. Any subset $A \subset \Omega$ is an event.*

2. *P, called the probability function, is a function that assigns a number in $[0, 1]$ to any event, denoted as $P : 2^{\Omega} \to [0, 1]$, and satisfies:*

   (a) For any event $A \subset \Omega$, $P(A) = \sum_{\omega \in A} P(\omega)$.

   (b) Normalization over the atomic events to 1, which means $\sum_{\omega \in \Omega} P(\omega)i = 1$.

Example 7.1.1. *Consider a dice rolling, where each of the faces is indexed by $1, 2, 3, 4, 5, 6$ and has an equal chance of being rolled. Therefore, our atomic events are associated with the rolling result, and P is defined as $P(\omega) = \frac{1}{6}$ for any such atomic event. An example of an event can be $A =$ "the dice falls on an even number". The probability of this outcome is:*

$$P(A) = \sum_{\omega \in A} P(\omega) = P(\{2\}) + P(\{4\}) + P(\{6\}) = 3 \cdot \frac{1}{6} = \frac{1}{2}$$

Claim 7.1.1. *The probability function satisfies the following properties:*

1. *$P(\emptyset) = 0$.*

2. *Monotonicity: If $A \subset B \subset \Omega$, then $P(A) \leq P(B)$.*

3. *Union Bound: $P(A \cup B) \leq P(A) + P(B)$.*

4. *Additivity for disjoint events: If $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$.*

5. *Complementarity: Denote by $\bar{A}$ the complementary event of A, which means $A \cup \bar{A} = \Omega$. Then, $P(\bar{A}) = 1 - P(A)$.*

Example 7.1.2. *Let's proof the additivity of disjointness property. Let A, B disjointness events, so $A \cap B = \emptyset$ then*

$$\begin{aligned} P(A \cup B) &= \sum_{w \in A \cup B} P(w) \\ &= \overbrace{\sum_{w \in A, w \notin B} P(w)}^{P(A)} + \overbrace{\sum_{w \in B, w \notin A} P(w)}^{P(B)} + \overbrace{\sum_{w \in A, w \in B} P(w)}^{0} \\ &= P(A) + P(B) \end{aligned}$$

Definition 7.1.2. *Let (Ω, P) be a probability space. A random variable X on (Ω, P) is a function $X : \Omega \to \mathbb{R}$. An indicator, is a random variable defined by an event $A \subset \Omega$ as follows*

$$X(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases}$$

Sometimes, we will use the notation $\{X = x\}$ to denote the event A such:

$$A = \{\omega : X(\omega) = x\} := \{X = x\}$$

Example 7.1.3. *Consider rolling a pair of dice. Denote by $X : [6] \times [6] \to [6]$ the random variable that is set to be the result of the first roll. Let Y be defined in almost the same way, but setting the result of the second die. Namely, if we denote by $\{(i, j)\}$ the atomic event associated with sample i on the first die and j on the second die, then:*

$$\begin{aligned} X(\{i, j\}) = i \\ Y(\{i, j\}) = j \end{aligned}$$

In addition, one can define the random variable z as the sum, $Z = X + Y$. Since the sum is also a function from Ω to $\mathbb{R}$, Z is also a random variable. An example of an indicator could be W, which gets 1 if $Z \in \{2, 7, 8\}$.

Definition 7.1.3. *We will say that two random variable $X, Y : \Omega \to \mathbb{R}$ are independent if for any $x \in \operatorname{Im} X$ and $y \in \operatorname{Im} Y$:*

$$P(X = x \cap Y = y) = P(X = x) \cdot P(Y = y)$$

Example 7.1.4. *X, Y defined in Example 7.1.3 are independent.*

$$\begin{aligned} P(\{X = i\} \cap \{Y = j\}) &= \sum_{i' = i \text{ and } j' = j} P(\{(i', j')\}) = P(\{(i, j)\}) \\ &= \frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6} = P(X = i) P(Y = j) \end{aligned}$$

7.2 Throwing Keys to Cells.

Imagines that following experiment, we have m cells and n keys (balls, numbers, your favorite object type). We throw each of the keys independently to cells. We would like to analyze how capacity of the cells is distributed.

Definition 7.2.1. *Let $X : \Omega \to \mathbb{R}$ be a random variable, the expectation of X is*

$$\mathbf{E}[X] = \sum_{\omega \in \Omega} X(\omega)P(\omega) = \sum_{x \in \operatorname{Im} X} xP(X = x)$$

Observes that if P is distributed uniformly, then the expectation of X is just the arithmetic mean:

$$\mathbf{E}[X] = \sum_{\omega \in \Omega} X(\omega)P(\omega) = \frac{1}{|\Omega|} \sum_{\omega \in \Omega} X(\omega)$$

Claim 7.2.1. *The expectation satisfies the following properties:*

1. *Monotonic, If $X \le Y$ (for any $\omega \in \Omega$) then $\mathbf{E}[X] \le \mathbf{E}[Y]$.*
2. *Linearity, for $a, b \in \mathbb{R}$ it holds that $\mathbf{E}[aX + by] = a\mathbf{E}[X] + b\mathbf{E}[Y]$.*
3. *Independently, if X, Y are independent, then $\mathbf{E}[X \cdot Y] = \mathbf{E}[X] \cdot \mathbf{E}[Y]$.*
4. *For any constant $a \in \mathbb{R}$ we have that $\mathbf{E}[a] = a$.*

Proof.

1. Monotonic, if $X \le Y$ then :

$$\mathbf{E}[X] = \sum_{\omega \in \Omega} X(\omega)P(\omega) \le \sum_{\omega \in \Omega} Y(\omega)P(\omega) = \mathbf{E}[Y]$$

2. Linearity,

$$\begin{aligned}\mathbf{E}[aX + bY] &= \sum_{\omega \in \Omega} \left(aX(\omega) + bY(\omega)\right) P(\omega) \\ &= a \sum_{\omega \in \Omega} X(\omega)P(\omega) + b \sum_{\omega \in \Omega} Y(\omega)P(\omega)\end{aligned}$$

3. Independently,

$$\begin{aligned}\mathbf{E}[XY] &= \sum_{x,y \in \operatorname{Im} X \times \operatorname{Im} Y} xyP(X = x \cap Y = y) \\ &= \sum_{x,y \in \operatorname{Im} X \times \operatorname{Im} Y} xyP(X = x)P(Y = y) \\ &= \sum_{x \in \operatorname{Im} X} \sum_{y \in \operatorname{Im} Y} xyP(X = x)P(Y = y) \\ &= \sum_{x \in \operatorname{Im} X} xP(X = x) \sum_{y \in \operatorname{Im} Y} yP(Y = y) \\ &= \sum_{x \in \operatorname{Im} X} xP(X = x)\mathbf{E}[Y] \\ &= \mathbf{E}[X]\,\mathbf{E}[Y]\end{aligned}$$

4. Let X be the random variable which is also the constant function $X(\omega) = a$ for any $\omega \in \Omega$. Then we have that

$$\begin{aligned}\mathbf{E}[X] &= \sum_{\omega \in \Omega} X(\omega)P(\omega) \\ &= \sum_{\omega \in \Omega} aP(\omega) = a \cdot 1 = a\end{aligned}$$

□

Example 7.2.1. *[COMMENT] Expectation of indicators and their multiplication. Let X be an indicator of event A, what are $\mathbf{E}[X]$ and $\mathbf{E}\left[X^2\right]$? Recall that $X(\omega) = 1$ only if $\omega \in A$ and 0 otherwise, thus:*

$$X^k(\omega) = \begin{cases} 1^k = 1 & \omega \in A \\ 0^k = 0 & \textit{else} \end{cases}$$

Therefore,

$$\mathbf{E}\left[X^k\right] = \sum_{\omega\in\Omega} X^k(\omega)P(\omega) = \sum_{\omega\in\Omega} X^k(\omega)P(\omega)$$

Example 7.2.2. *[COMMENT] How many keys trowed into the same cell as the first key thrown to?*

```
let B[0 : n − 1] be a new array
for i ← [0, n−1] do
    make B_i an empty list
end
for i ← [1, n] do
    insert A_i into list B_⌊nA_i⌋
end
for i ← [0, n−1] do
    sort list B_i
end
concatenate the lists B_0, B_1, .., B_{n−1} together and
return the concatenated lists
```

Algorithm 1: bucket-sort(A, n)

Denote by $X_i : [n] \to [n]$ then random variable that counts the number of elements fallen in the ith bucket. The Expectation of the sorting running time is:

$$\begin{aligned}
\mathbf{E}[T] &= \mathbf{E}\left[\text{Inserting into buckets } + \sum_i \text{Sorting } i\text{th bucket}\right] \\
&= \mathbf{E}\left[\Theta(n) + \sum_i X_i^2\right] = \Theta(n) + \sum_i \mathbf{E}\left[X_i^2\right] \\
\mathbf{E}\left[X_i^2\right] &= \mathbf{E}\left[\left(\sum X_i^j\right)^2\right] = \mathbf{E}\left[\sum_{j,j'} X_i^j X_i^{j'}\right] = \sum_{j,j'} \mathbf{E}\left[X_i^j X_i^{j'}\right] \\
&= \sum_{j\neq j'} \mathbf{E}\left[X_i^j X_i^{j'}\right] + \sum_j \mathbf{E}\left[X_i^j X_i^j\right] \\
&= \sum_{j\neq j'} \mathbf{E}\left[X_i^j X_i^{j'}\right] + \sum_j \mathbf{E}\left[X_i^j\right] \\
&= 2\binom{n}{2}\left(\frac{1}{n}\right)^2 + n \cdot \frac{1}{n} \\
&= \frac{n-1}{n} + 1 = 2 - \frac{1}{n} \Rightarrow \mathbf{E}[T] = \Theta(n) + n\left(2 - \frac{1}{n}\right) = \Theta(n)
\end{aligned}$$