A primer on finite probability spaces

Finite Probability Space: A set Ω equipped with a function $\mathbb{P} : \Omega \to \mathbb{R}^+$ with the property that $\sum_{x \in \Omega} \mathbb{P}(x) = 1$. We consider $\mathbb{P}(x)$ to be the probability that $x \in \Omega$ occurs.

Event: An *event* is a subset $A \subseteq \Omega$, and the probability that A occurs is $\mathbb{P}(A) = \sum_{x \in A} \mathbb{P}(x)$.

Random Variable: A *random variable* is simply a function $X : \Omega \to \mathbb{R}$.

Expected Value: If X is a random variable, then the *expected value* of X is defined to be $\mathbb{E}(X) = \sum_{x \in \Omega} \mathbb{P}(x) X(x)$.

Example: Consider a random experiment where a fair coin is tossed two times. To model this experiment, we let $\Omega = \{HH, HT, TH, TT\}$ be the set of possible outcomes and we set $\mathbb{P}(x) = \frac{1}{4}$ for each $x \in \Omega$ since each of these possibilities occurs with probability $\frac{1}{4}$. If A is the event "At least one Heads appears", then $A = \{HH, HT, TH\}$ and we have $\mathbb{P}(A) = \frac{3}{4}$. If we set X to be the random variable "number of Heads which appear", then $X(HH) = 2$, $X(HT) = X(TH) = 1$ and $X(TT) = 0$ so $\mathbb{E}(X) = 1$.

Observation 1 *If X is a random variable with expected value t, then there exists $y \in \Omega$ with $X(y) \geq t$ and $z \in \Omega$ with $X(z) \leq t$.*

Proof: Choose $y \in \Omega$ with $X(y)$ maximum and $z \in \Omega$ with $Y(z)$ minimum. Then we have

$$\begin{aligned} t &= \mathbb{E}(X) = \sum_{x \in X} \mathbb{P}(x)X(x) \leq \sum_{x \in X} \mathbb{P}(x)X(y) = X(y) \\ t &= \mathbb{E}(X) = \sum_{x \in X} \mathbb{P}(x)X(x) \geq \sum_{x \in X} \mathbb{P}(x)X(z) = X(z). \end{aligned}$$

Observation 2 *If X and Y are random variables, then $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$.*

Proof:

$$\begin{aligned} \mathbb{E}(X+Y) &= \sum_{x \in \Omega} \mathbb{P}(x)(X(x) + Y(x)) \\ &= \sum_{x \in \Omega} \mathbb{P}(x)X(x) + \sum_{x \in \Omega} \mathbb{P}(x)Y(x) \\ &= \mathbb{E}(X) + \mathbb{E}(Y). \end{aligned}$$

Indicator random variable: A random variable X is called an *indicator* random variable if it takes the value 1 on a subset $A \subseteq \Omega$ and the value 0 on $\Omega \setminus A$. Note that in this case $\mathbb{E}(X) = \mathbb{P}(A)$.

Proposition 3 *If σ is a permutation of $\{1, 2, \ldots, n\}$ chosen uniformly at random, then the expected number of fixed points of σ is 1 ($i \in \{1, 2, \ldots, n\}$ is a fixed point if $\sigma(i) = i$).*

Proof: Let X be the number of fixed points in σ, and for every $i \in \{1, 2, \ldots, n\}$ let X_i be the indicator random variable which has the value 1 if $\sigma(i) = i$ and 0 otherwise. It is immediate that $\mathbb{E}(X_i) = \mathbb{P}(\sigma(i) = i) = \frac{1}{n}$ so by linearity of expectation we have

$$\mathbb{E}(X) = \sum_{i=1}^{n} \mathbb{E}(X_i) = n\frac{1}{n} = 1.$$

□

Theorem 4 (Markov) *If X is a random variable which is always nonnegative, then $\mathbb{P}(X \geq t) \leq \frac{1}{t}\mathbb{E}(X)$.*

Proof: Using Ω for our probability space we have

$$\begin{aligned} \mathbb{E}(X) &= \sum_{x \in \Omega} \mathbb{P}(x) X(x) \\ &\leq \sum_{x \in \Omega : X(x) \geq t} \mathbb{P}(x) X(x) \\ &\leq t \sum_{x \in \Omega : X(x) \geq t} \mathbb{P}(x) \\ &\leq t\,\mathbb{P}(X \geq t) \end{aligned}$$

as desired. □