# 3 Discrete Random Variables

**We need good data and good statistical thinking in order to make good predictions.**

## 3.1 Random variables

It is better to define a random variable as a *function* mapping the sample space to the real line.

### Definition of Random Variable:

Given an experiment with sample space $S$, a *random variable* is a function from the sample space $S$ to the real numbers $\mathbb{R}$.

A random variable $X$ assigns a numerical value $X(s)$ to each possible outcome $s$ of the experiment. The randomness comes from the fact that we have **a random experiment** the mapping itself is **deterministic**.

### Example: coin tosses

We toss a fair coin twice. The sample space $S = HH, HT, TH, TT$.

Some random variables on this space:

- $X$ is the number of Heads. So $X(HH) = 2, X(TH) = X(HT) = 1, X(TT) = 0$;
- $Y$ is the number of Tails. $Y = 2 - X$ or $Y(s) = 2 - X(s) \forall s$;
- $I$ is 1 is the first toss lands Heads and 0 otherwise. This is an example of *indicator random variable*.

We also can encode sample space as $(1,1), (1,0), (0,1), (0,0)$. Then we can give explicit formulas for $X, Y, I$ :

$$X(s_1, s_2) = s_1 + s_2, Y(s_1, s_2) = 2 - s_1 - s_2, I(s_1, s_2) = s_1$$

After we perform the experiment and the outcome $s$ has been realized, the random variable *crystallizes* into the numerical value $X(s)$.

# 3.2 Distributions and probability mass functions

$X$ is said to be **discrete** if there is a finite list of values $a_1, a_2, \ldots$ such that $P(X = a_j \ for \ some \ j) = 1$. If X is a discrete, then set of values $x$ such that $P(X = x) > 0$ is called the **support** of $X$.

In contrast, a *continuous* RV can take on any real value in an interval (possibly even the entire real line);

It is also possible to have an RV that is a *hybrid* of discrete and continuous.

The *distribution* of a RV specifies (for example) the probabilities of all events associated with the random value, such as the probability of it equaling 3 and the probability of it being at least 110.

For a discrete RV, the most natural way to do so is with a **probability mass function (PMF)**:

PMF of discrete RV $X$ is the function $p_X$ given by $p_X(x) = P(X = x)$. If $x$ not in support of $X$, $p_X(x) = 0$. $X = x$ denotes as an *event*, all outcomes $s$ to which $X$ assigns to the number $x$.

## PMF example:

Two fair coin tosses. RV: $X$, the number of Heads:

$$p_X(0) = P(X = 0) = 1/4$$

$$p_X(1) = P(X = 1) = 1/2$$

$$p_X(2) = P(X = 2) = 1/4$$

## Properties of a valid PMF:

$X$ is a discrete RV with support $x_1, x_2, \ldots$

The PMF $p_X$ of $X$ must satisfy following criteria:

- Nonnegative $p_X(x) > 0$ if $x = x_j$ for some $j$ and $p_X(x) = 0$ otherwise;
- Sums to 1: $\sum_{j=1}^{\infty} p_X(x_j) = 1$

## Example: Poisson Distribution

An RV $X$ has the Poisson distribution with parameter $\lambda$ where $\lambda > 0$ if PMF of $X$:

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}, k = 0, 1, 2...$$

The Poisson also arises through the Poisson process, a model that is used in a wide variety of problems in which events occur at random points in time.

# 3.3 Bernoulli and Binomial

Named distributions!

RV $X$ has *Bernoulli distribution* with parameter $p$ if $P(X = 1) = p$ and $P(X = 0) = 1 - p$ where $0 < p < 1$. It can be written as $X \sim Bern(p)$.

Any event has a Bernoulli RV that is naturally associated with it! This is called the *indicator RV*.

Indicator RV of event $A$ is $I_A$ or $I(A)$.

*Bernoulli trial*: an experiment which can result in “success” or “failure” (not both).

$n$ independent Bernoulli trials each with success probability $p$. RV $X$ is the number of success. The distribution of $X$ is called *Binomial distribution* with parameters $n$ and $p$. $X \sim Bin(n, p)$.

*Binomial PMF*: If $X \sim Bin(n, p)$ then the PMF of $X$ is:

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

for $k = 0, 1, \ldots, n$ and $P(X = k) = 0$ otherwise.

PMF of the $Bin(10, 1/2)$ is symmetric about 5. If $p \neq 1/2$, the PMF is *skewed*.

If $X \sim Bin(n, p)$ and $q = 1 - p$ (so $q$ is failure probability of Bernoulli trial), $n - X \sim Bin(n, q)$.

Corollary:

Distribution of $X$ is symmetric about $n/2$. Why?

$$P(X = k) = P(n - X = k) = P(X = k - k) \forall k > 0$$

# 3.4 Hypergeometric