Topic 2  Single Random Variables

What is a random variable?

Def A random variable is a function that maps an outcome to a number on the real line.

Note that this real number does not need to be bounded in [0,1] because it is not a probability.

For example, if we flip a coin,
\$\xi_1 = "H" \$, \$\xi_0 = "T" \$
\[ X(\xi_1) = 1 , \quad X(\xi_0) = 0 \]
\[ \text{this 1 and 0 is arbitrary.} \]

Notations:
\[
X \quad \text{capital letter} ; \ a \ \text{random variable} \\
\xi \quad \text{an outcome} \\
X(\xi) \quad \text{the value of} \ X \ \text{at a particular outcome} \ \xi \\
x \quad \text{small letter} ; \ a \ \text{generic real number}.
\]

\[
\{ X \leq x \} = \{ \xi \mid X(\xi) \leq x \}
\]
Example
Flip a coin 3 times.
Sample space is
\[ \{ (HHH), (HHT), (HTH), (THH), (TTH), (THT), (HTT), (TTT) \} \]
Let \( X \) = number of heads.

Then:
(1) There are 8 \( \Omega \), each corresponds to one outcome;
(2) There are only 4 values of \( X(\Omega) \):
\[
X(\Omega) = 0, 1, 2, 3.
\]
Note that \( X \) is a function. E.g.
\[
X((HTH)) = 2, \quad X((TTH)) = 1.
\]
When we ask "What is the probability that \( X = 2 \)?"
we essentially mean that:
How many \( \Omega \) would give \( X(\Omega) = 2 \)?

Discrete Random Variables

Def A random variable \( X \) is discrete if
the range of \( X \) contains a countable number of elements \( \{x_1, x_2, \ldots\} \).

Def A probability mass function (PMF) is
\[
P_X(x) = \text{IP}(X = x).
\]
Example: Flip 3 coins. \(X=\# \text{ of head.}\)

The PMF of \(X\) is

\[
P_X(0) = \frac{1}{8} \quad \text{so} \quad P(X=0) = P(\text{TTT}) = \frac{3}{8}.
\]

\[
P_X(1) = \frac{3}{8}
\]

\[
P_X(2) = \frac{3}{8}
\]

\[
P_X(3) = \frac{1}{8}
\]

---

**Basic Properties of \(P_X(x)\):**

1. \(\sum_k P_X(x_k) = 1\)

2. \(P_X(x) = \left\{
\begin{array}{ll}
0 & \text{if } x \notin \{x_1, x_2, \ldots \} \\
\geq 0 & \text{if } x \in \{x_1, x_2, \ldots \}
\end{array}
\right.\)
Expected Value (Mean)

**Def** The expected value of a discrete random variable $X$ is

$$E[X] = \mu_X = \sum_{k} x_k p_X(x_k)$$

**Example**

\[
\begin{align*}
\mu_X &= (0)(\frac{1}{8}) + (1)(\frac{3}{8}) + (2)(\frac{3}{8}) + (3)(\frac{1}{8}) \\
&= \frac{12}{8} = 1.5
\end{align*}
\]

That is, if we randomly flip the coin 3 times, on average we will have 1.5 heads.

**Example**

Flip a coin 3 times. Reward $1 if 2 Heads
$8 if 3 Heads
$0 if 0 or 1 Head

To enter the game, need to pay $1.5 first.

**Question:** to bet or not to bet?

Let $X=$ amount of reward.

\[
\begin{align*}
\mu_X &= (0)(\frac{4}{8}) + (1)(\frac{3}{8}) + (8)(\frac{1}{8}) \\
&= \frac{11}{8}
\end{align*}
\]

Entrance fee : $\frac{12}{8}$.

So $\mu_X - \text{Entrance Fee} = \frac{-1}{8}$

On average, you will lose.
Sample Mean VS Mean

- $\mathbb{E}[X]$ is the "long term" average of the r.v. $X$.
- If we conduct an experiment for $n$ times, and suppose we make an empirical observation that $X_k$ is observed for $N_k$ times, then the "average" would be

$$\bar{X} = \frac{x_1 N_1 + x_2 N_2 + \ldots}{n}, \text{ where } \sum_k N_k = n.$$ 

- In general, $\bar{X} \to \mathbb{E}[X]$, but $\bar{X} \neq \mathbb{E}[X]$ for finite $n$.
- Result of law of large number. Will study later.

Expected Value of Functions of R.V.

For any function $g$,

$$\mathbb{E}[g(X)] = \sum_k g(x_k) \mathbb{P}(X = x_k).$$

Example

Let $Z = X^2$. Find $\mathbb{E}[Z] = \mathbb{E}(X^2)$.

Approach 1:

$$\mathbb{E}[Z] = (1)(\frac{1}{2}) + (9)(\frac{1}{2}) = \frac{10}{2} = 5.$$ 

Approach 2:

$$\mathbb{E}[Z] = \mathbb{E}(X^2) = (3^2)(\frac{1}{4}) + (-1)^2(\frac{1}{4}) + (4)^2(\frac{1}{4}) = \frac{20}{4} = 5.$$
Some Properties of $\mathbb{E}[X]$.

1. Linearity

\[
\mathbb{E}[g(X) + h(X)] = \mathbb{E}[g(X)] + \mathbb{E}[h(X)]
\]

Proof \(\textbf{pf}\):

\[
\mathbb{E}[g(X) + h(X)] = \sum_{k} (g(x_k) + h(x_k)) P_X(x_k)
\]
\[
= \sum_{k} g(x_k) P_X(x_k) + \sum_{k} h(x_k) P_X(x_k)
\]
\[
= \mathbb{E}[g(X)] + \mathbb{E}[h(X)].
\]

2. Scale

\[
\mathbb{E}(cX) = c \mathbb{E}(X).
\]

Proof \(\textbf{pf}\):

\[
\mathbb{E}(cX) = \sum_{k} (cX_k) P_X(x_k) \quad \overset{g(x) = cx}{=} \quad g(X) = cX
\]
\[
= c \sum_{k} X_k P_X(x_k) = c \mathbb{E}(X).
\]

3. DC shift

\[
\mathbb{E}(X + c) = \mathbb{E}(X) + c
\]

Proof \(\textbf{pf}\):

\[
\mathbb{E}(X + c) = \sum_{k} (X_k + c) P_X(x_k)
\]
\[
= \sum_{k} X_k P_X(x_k) + \sum_{k} c P_X(x_k)
\]
\[
= \mathbb{E}(X) + c \sum_{k} P_X(x_k)
\]
\[
= \mathbb{E}(X) + c 1 = \mathbb{E}(X) + c.
\]

Example

\[
\mathbb{E}[(2X+10)^2]
\]
\[
\mathbb{E}\left[4X^2 + 40X + 100\right]
\]
\[
= 4 \mathbb{E}[X^2] + 40 \mathbb{E}[X] + 100
\]
\[
= (4)(5) + (40)(0) + 100 = 120.
\]