Random Variables: Motivation

"Experiments" | Outcomes $\in \Omega$ | Measurements

Eg. Toss a fair coin N times.

H T H H

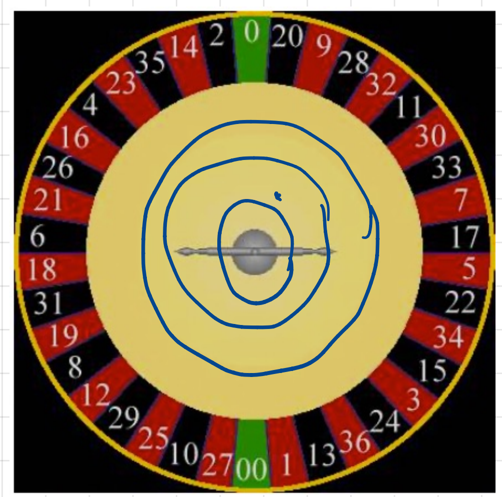
HTHH

$X = \#\text{heads}$

$X = 3$

$X \leq 3, \; X \geq 2$

Eg. Throw a dart at a board of radius R.



$(r, \theta) = (0.15, \frac{\pi}{6})$

$R = \text{dist. from center}$

$R = 0.15$

$R \leq 0.25$

- Each experiment has an **outcome**.
- The set of all possible outcomes is the **sample space** Ω.
- "Probability" is a measure of the likelihood of a **set of outcomes** = an **event** $E \subseteq \Omega$.

$X : \Omega \to \mathbb{R}$

"random variables"

$(\Omega, \mathcal{F}, \mathbb{P})$ probability space. "inaccessible"

Outcomes are elements $\omega \in \Omega$.

Events are subsets of Ω, in $\mathcal{F}$.

A random variable is a function $X : \Omega \to S$

"state space" usually $\mathbb{R}$; could be $\mathbb{C}$ (could be $\mathbb{R}^d$; then usually call X a "random vector")

(Probably should call them "random functions" but the very old "variable" terminology has stuck since used by Laplace in the early 19^{th} Century.)

Need to be able to calculate probabilities of events like

$$\{X \leq 1\} = \{\omega \in \Omega : X(\omega) \leq 1\}$$

Def: A function $X : \Omega \to \mathbb{R}$ is a random variable if $\{X \leq t\} \in \mathcal{F}$ for all $t \in \mathbb{R}$.

CDFs (Again)

$X : \Omega \to \mathbb{R}$ random variable on $(\Omega, \mathcal{F}, P)$.

Define $F_X : \mathbb{R} \to \mathbb{R} : \quad F_X(t) = \mathbb{P}(X \leq t) = \mathbb{P}\{\omega \in \Omega : X(\omega) \leq t\}$

Proposition: F_X is non-decreasing, right-continuous, and

$$\lim_{t \to -\infty} F_X(t) = 0, \quad \lim_{t \to +\infty} F_X(t) = 1.$$

Pf $\cdot$ If $s \leq t$ $\quad \{X \leq s\} \subseteq \{X \leq t\}$

$$F_X(s) = \mathbb{P}(X \leq s) \leq \mathbb{P}(X \leq t) = F_X(t)$$

$\cdot$ If $t_n \downarrow t$, $\quad \{X \leq t_n\} \downarrow \{X \leq t\} = \bigcap_{n=1}^{\infty} \{X \leq t_n\}$

$$\therefore \underbrace{\mathbb{P}(X \leq t_n)}_{F_X(t_n)} \downarrow \mathbb{P}(X \leq t) = F_X(t)$$

$\cdot$ If $t_n \downarrow -\infty$, $\quad \{X \leq t_n\} \downarrow \emptyset \quad F_X(t_n) = \mathbb{P}(X \leq t_n) \downarrow 0$

If $t_n \uparrow \infty$, $\quad \{X \leq t_n\} \uparrow \Omega \quad F_X(t_n) = \mathbb{P}(X \leq t_n) \uparrow \mathbb{P}(\Omega) = 1$.

$\therefore$ F_X is the CDF of a unique Borel probability measure μ_X on $\mathbb{R}$

$\uparrow$

The **probability distribution** of X.

Often μ_X is all we'll really know about X.

And more often, we won't even know μ_X, but will only have some limited clues about it.

$$X \rightsquigarrow F_X \rightsquigarrow \mu_X$$

(Great) Expectations

Eg. Finite sample space $\Omega = \{\omega_1, \omega_2, \dots, \omega_N\}$

(May as well have $\mathcal{F} = 2^\Omega$.)

Then $\quad \mathbb{P}(E) = \sum_{\omega \in E} \mathbb{P}(\{\omega\}) = \sum_{\omega \in \Omega} \mathbb{1}_{\omega \in E} \, \mathbb{P}(\{\omega\})$ $\qquad \mathbb{1}_A(\omega) = \begin{cases} 1, & \text{if } \omega \in A \\ 0, & \text{if } \omega \notin A \end{cases}$

If $X : \Omega \to \mathbb{R}$ is a random variable,

$$F_X(t) = \mathbb{P}(X \le t) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \le t\}) = \sum_{\omega \in \Omega} \mathbb{1}_{X(\omega) \le t} \, \mathbb{P}(\{\omega\})$$

Can we get a "snapshot" number that tells us <u>something</u> about this distribution?

↳ weighted average: $\quad \mathbb{E}(X) := \sum_{\omega \in \Omega} X(\omega) \, \mathbb{P}(\{\omega\})$

Eg. Toss a fair coin 3 times; $X = \#$ Heads.

$\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$

$\mathbb{P} = \quad \overbrace{\qquad\qquad\qquad}^{} \; 1/8 \; \overbrace{\qquad\qquad\qquad}^{}$

$X = \quad 3 \quad 2 \quad 2 \quad 2 \quad 1 \quad 1 \quad 1 \quad 0$

$$\mathbb{E}(X) = \frac{1}{8}(3 + 3 \cdot 2 + 3 \cdot 1 + 0)$$

$$= \frac{3}{2} = 1.5$$

$$E(X) = \sum_{\omega\in\Omega} X(\omega)\,\mathbb{P}(\{\omega\}) \qquad (\star)$$

Makes perfect sense if Ω is finite. Also okay if Ω is countable.
But won't help us if $\mathbb{P}(\{\omega\}) = 0$ for all $\omega\in\Omega$.

Undergraduate Probability Approach:

$$\sum_{\omega} X(\omega)\,\mathbb{P}(\{\omega\}) = \sum_{t}\ \sum_{\omega : X(\omega)=t} X(\omega)\,\mathbb{P}(\{\omega\}) = \sum_{t} t \sum_{X(\omega)=t} \mathbb{P}(\{\omega\}) = \underline{\underline{\sum_{t} t\,\mathbb{P}(X=t)}}$$

Find an $\int$ analog of $\sum$ in the "continuous" setting.

Problems: Many.

$$E(X+Y) = E(X) + E(Y).$$

We will develop the right generalization of $(\star)$ to work in any probability space: the Lebesgue Integral

$$E(X) = \int_{\Omega} X\, d\mathbb{P}.$$