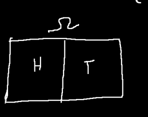
Consider Ω = {H, T}. And 3 flips.




$\omega_1 = H$
$\omega_2 = T$
$\omega_3 = H$

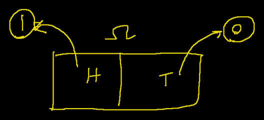
* What is the average of the three flips? Don't know...
* How "randomly spread out" are the three flips? Don't know...

The reason why we can't answer, is because we didn't define "average" and "spread" on sets with outcomes. We need to link outcomes to numerical values. Consider the indicator function:

$$\mathbb{1}_{\omega=H} = \begin{cases} 1 & \text{if } \omega = H \\ 0 & \text{o/t} \end{cases}$$

sample average

$$n=3 \quad \mathbb{1}_{\omega_1=H} = 1,\ \mathbb{1}_{\omega_2=H} = 0,\ \mathbb{1}_{\omega_3=H} = 1 \Rightarrow \bar{x} = \frac{2}{3} = 0.67$$




This mapping from ω's to x's (numerical values) is a function, name it X(ω):

X : Ω → ℝ

This function is called a random variable (rv). It sometimes also has a unit e.g. $, inches, meters, etc.

In our setting, X(H) = 1, X(T) = 0

What is P(X = 1)? But this is technically a nonsense expression. Why? Because P : 2^Ω → [0, 1] and "X=1" is not an ω.

It is still an expression we will use but it is shorthand for the following:

P(X = 1) := P({ω : X(ω) = 1}) = P({H}) = |{H}| / |Ω| = 1/2.

The "support" of rv X denoted Supp[X] is the range of its possible values. In our coin case, the possible values X can return are zero or one. Thus we say, Supp[X] = {0, 1}. Generally speaking,

Supp[X] = {x : P(X = x) > 0} ⊆ ℝ

Definition of a "discrete rv" X: |Supp[X]| ≤ |ℕ| i.e. it is finite or at most "countably infinite". Note the following "Humpty Dumpty" rule:

$$\sum_{x \in Supp[X]} P(X = x) = 1$$

Proof: not on the midterm or final

$$\Omega = \bigcup_{x \in Supp[X]} \{\omega : X(\omega) = x\} \quad \text{If } \nexists\, \omega \text{ s.t. } X(\omega) \notin Supp[X] \Rightarrow P(\{\omega\}) = 0$$

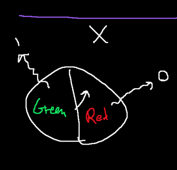
ω's are collectively exhaustive. Consider x_1 ≠ x_2, then:

$$\{\omega : X(\omega) = x_1\} \cap \{\omega : X(\omega) = x_2\} = \phi$$ otherwise X is not a function

[A1] assume discrete rv [A3]

$$1 = P(\Omega) = P\left(\{\omega : X(\omega) = x_1\} \cup \{\omega : X(\omega) = x_2\} \cup \dots\right) = P(\{\omega : X(\omega) = x_1\})$$
$$+ P(\{\omega : X(\omega) = x_2\}) + \dots \sum_{x \in Supp[X]} P(\{\omega : X(\omega) = x\}) = \sum_{x \in Supp[X]} P(X = x)$$

---




$X(\text{Green}) = 1,\ X(\text{Red}) = 0$

$P(X=1) = \frac{1}{2},\ P(X=0) = \frac{1}{2}$

Note this is the "same" rv as the rv derived from the coin flip experiment if we look at the x values and their probabilities only i.e.

$$X \sim \begin{cases} 1 & \text{wp } \frac{1}{2} \\ 0 & \text{wp } \frac{1}{2} \end{cases}$$

distributed as — with probability

There are many random experiments with many different Ω's that can produce this rv that is "distributed" the same. From now on, we won't care at all about the random experiment nor the Ω, we only care about the values x ∈ Supp[X] and their probabilities. "Little x" is called the realized value. Rv's get "realized" and emit little x's. X --> x when X is realized.

The rv we've been discussion is (IMHO) the most important, most fundamental rv called the "standard Bernoulli". It is the first of the "brand name" rv's and denoted:

$$X \sim \text{Bern}\left(\tfrac{1}{2}\right) := \begin{cases} 1 & \text{wp } \frac{1}{2} \\ 0 & \text{wp } \frac{1}{2} \end{cases}, \quad Supp[X] = \{0, 1\}$$

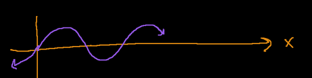
More generally, the "Bernoulli rv" is defined as:

$$X \sim \text{Bern}(p) := \begin{cases} 1 & \text{wp } p \\ 0 & \text{wp } 1-p \end{cases}, \quad Supp[X] = \{0, 1\}$$

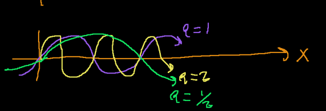
X is distributed as a Bernoulli with "parameter" p. If p = 1/2, it is called the "standard Bernoulli". A parameter is a choice that defines the specifics of a model.

We've seen this in precalc. For example,

$f(x) = \sin(x)$

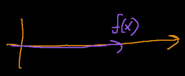



$f(x) = \sin(a x)$




"a" is a parameter that controls "frequency" of the wave. Yes if a changes, it's a different function. But it's still in the family of sine curves.

$\text{If } a = 0 \Rightarrow f(x) = \sin(0x) = 0\ \forall x$




Is a=0 a valid value for the parameter? Yes/no. I will say no. Because no longer is it a sine curve. We'll call this case "degenerate" because it degenerates the function to something trivial.

Thus, the "parameter space", the set of valid values of a would be (0, ∞). Note: 0 ∉ the parameter space because it yields a degeneracy.

What can the values of p be? What is the parameter space? We defined the states "with probabilities" so p can only be in [0,1].

What if p = 1? The rv realizes to... 1 always => X ~ {1 wp 1
What if p = 0? The rv realizes to... 0 always => X ~ {0 wp 1

These cases of p=1 and p=0 yield a rv that is trivial, degenerate. Thus, we omit these cases from the parameter space. Thus,

p ∈ (0, 1) i.e. all values between 0 and 1 not including 0 and 1.

Technically, the "degenerate rv" is a rv and it's our second brand-name rv:

X ~ Deg(c) := {c wp 1,  Supp[X] = {c} and param space c ∈ ℝ.

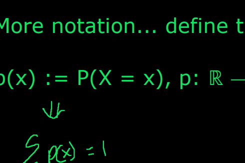
Again, this is "oxymoronic" rv since it aint random!! Thus it is trivial an not interesting. It just spits out the same thing every time.

More notation... define the "probability mass function" (PMF) as

p(x) := P(X = x), p: ℝ → [0, 1] where p(x) = 0 if x ∉ Supp[X] and
p(x) = 1 if X ~ Deg

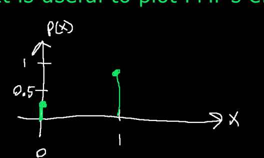
$$\Downarrow$$
$$\sum_{x \in Supp[X]} p(x) = 1$$

It is useful to plot PMF's e.g. X ~ Bern(0.75)

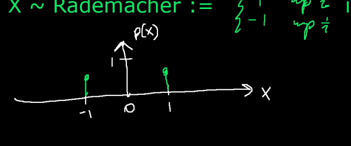



Is P(x) a continuous function? No.
How many discontinuities does it have? |Supp[X]|

X ~ Rademacher := $\begin{cases} 1 & \text{wp } \frac{1}{2} \\ -1 & \text{wp } \frac{1}{2} \end{cases}$ i.e. the "random walk in one dim."




Consider the "uniform discrete" rv: X ~ U({1, 10, 100}) := $\begin{cases} 1 & \text{wp } 1/3 \\ 10 & \text{wp } 1/3 \\ 100 & \text{wp } 1/3 \end{cases}$




Generally,
X ~ U(A) := {1 / |A| for all x ∈ A,  Supp[X] = A,
param space A ⊂ ℝ s.t. |A| ≤ ℕ