Coin tossing: Fair coin tossed n times $\to$ ind / id. dist

prob (H) = prob $(T) = 1/2$

2^n possible outcomes $\to$ each has prob 2^{-n}

What is prob of exactly k heads? ${}^nC_k / 2^n$

What is the prob of n heads in $2n$ tosses?

$$\text{Ans} = \frac{(2n!)}{(n!)^2 2^{2n}} \approx \frac{1}{\sqrt{\pi n}}(1+\delta_n) \quad \text{where } \delta_n \to 0 \text{ as } n \to \infty$$

As $2n \to \infty$ Prob of n heads $\to 0$.

So clearly convergence to $1/2$ is not happening in this sense

Weak Law of Large numbers:

Let Ω_n be the set of $\{0,1\}$ sequences of length n

an individual sequence is written ω.

Prop Let $X : \Omega_n \to \mathbb{R}$ be any function & let $\varepsilon > 0$

$$P(\{\omega \mid |X(\omega)| \geq \varepsilon\}) \leq \frac{1}{\varepsilon^2} \mathbb{E}(X^2)$$

Proof $$P(\{\omega \mid |X(\omega)| \geq \varepsilon\}) = P(\{\omega \mid X^2(\omega) \geq \varepsilon^2\})$$

$$= \sum_{\omega : X^2(\omega) \geq \varepsilon^2} \frac{1}{2^n} \leq \sum_{\omega : X^2(\omega) \geq \varepsilon^2} \frac{X^2(\omega)}{\varepsilon^2} \frac{1}{2^n} \leq \frac{1}{\varepsilon^2} \sum_{\omega \in \Omega_n} \frac{X^2(\omega)}{2^n}$$

$$= \frac{1}{\varepsilon^2} \mathbb{E}(X^2). \quad \blacksquare$$

Now define $X_j(\omega) = \begin{cases} 1 & \text{if } \omega[j] = H \\ 0 & \text{if } \omega[j] = T \end{cases}$

$$S_n(\omega) = \sum_{i=1}^n X_i(\omega)$$

e.g. $\mathbb{E}(X_i) = 1/2 \qquad \mathbb{E}(X_i X_j) = 1/4$ for $i \neq j$

$1/4 - 1/4 - 1/4 + 1/4$ $\qquad \mathbb{E}((X_i - 1/2)(X_j - 1/2)) = 0$ for $i \neq j \qquad \mathbb{E}(X_i - 1/2)^2 = 1/4$

$$S_n - \frac{n}{2} = \sum_{i=1}^n (X_i - 1/2)$$

$$\cancel{\mathbb{E}(S_n - \tfrac{n}{2})^2} = \mathbb{E}\left(\frac{S_n}{n} - 1/2\right)^2 = \frac{1}{n^2} \mathbb{E}\left(\sum_{i,j} (X_i - 1/2)(X_j - 1/2)\right)$$

$$= \frac{1}{4n} \quad \text{(cross-terms vanish)}$$

Now. $$P\left(\left\{\omega \,\middle|\, \left|\frac{S_n(\omega)}{n} - \frac{1}{2}\right| \geq \varepsilon\right\}\right) \leq \mathbb{E}\left(\frac{S_n}{n} - 1/2\right)^2 / \varepsilon^2 \quad \text{(Chebyshev)}$$

$$P(\ldots) \leq \frac{1}{4n\varepsilon^2}$$

$$\lim_{n \to \infty} P\left(\left\{\omega \,\middle|\, \left|\frac{S_n(\omega)}{n} - \frac{1}{2}\right| \geq \varepsilon\right\}\right) = 0$$

Strong law of large numbers
Let Ω be infinite sequences of 0's & 1's $\omega \in \Omega$.
Define $X_j : \Omega \to \mathbb{R}$ as before.

$$S_n(\omega) = \sum_{j=1}^{n} X_j(\omega)$$

$$\lim_{n\to\infty} \frac{S_n(\omega)}{n} = \frac{1}{2} \quad ?? \quad \text{Absolutely false!!}$$

$$Pr\left(\left\{\omega \,\middle|\, \lim_{n\to\infty} \frac{S_n(\omega)}{n} = \frac{1}{2}\right\}\right) = 1 \qquad \text{Limit is inside.}$$

Do we know how to make sense of this?
How do we assign probabilities to subsets of Ω?

What is the probability of the sequence
HTHTHT····· ? Ans → 0.
So any element of Ω has prob. 0. We cannot get probabilities of all sets by looking at probabilities of singletons or <u>atoms</u>.
What is the prob of $Pr(HT^\infty) = 0$

$Pr(THT^\infty) = 0$

$Pr(T^2HT^\infty) = 0$

$\vdots$

Prob exactly one $H = 0$

Prob exactly 2 $H = 0$

Prob exactly n, $H = 0$

Prob at least 1 H is $= 0 + 0 + \cdots + 0 + \cdots = 0$ Huh?!
Prob all $T = 0$ so Prob at least 1 H is 1 not 0!!
We forgot $(HT)^\infty$ and similar sequences. What we proved is Prob (finite # of H) $= 0$.
For what sequences can we assign probs?
Notation $s \in \Omega_n$ for some n $s \le \omega$; s <u>prefix</u> of $\omega \in \Omega$.

$s\uparrow := \{\omega \mid s \leq \omega\}$  a wineglass

$P_r(s\uparrow) = 2^{-|s|}$ $|s|$: length of s.

But $\{\omega \mid \lim_{n\to\infty} \frac{S_n}{n} = 1/2\}$ is not a wine glass!
How do we assign measures to such sets?

What are the basic axioms we want?

$(\Omega, \mathcal{F}, \mathbb{P}) \to$ probability triple

$\Omega \to$ sample space $\mathcal{F} \to$ mbl sets $\mathbb{P} \to$ prob measure

$\mathbb{P} : \mathcal{F} \to [0,1]$

(i) Want $\mathbb{P}(\phi) = 0$ & $\mathbb{P}(\Omega) = 1$ so $\phi, \Omega \in \mathcal{F}$

(ii) $\mathbb{P}(A) + \mathbb{P}(A^c) = 1$ so $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$

(iii) $A, B \in \mathcal{F}$ & $A \cap B = \phi \Rightarrow \mathbb{P}(A) + \mathbb{P}(B) = \mathbb{P}(A \cup B)$

in fact $\{A_i \mid i \in \mathbb{N}\}$ $A_i \in \mathcal{F} \Rightarrow \bigcup_i A_i \in \mathcal{F}$.

$\forall i \neq j$ $A_i \cap A_j = \phi \Rightarrow \mathbb{P}(\bigcup_i A_i) = \sum_i \mathbb{P}(A_i)$.

Uncountable sums don't make sense as we saw with coin tossing.

def A σ-algebra $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ is a family of sets s.t.
(i) $\phi \in \mathcal{F}$ (ii) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$
(iii) $A_i \in \mathcal{F}$ for $i \in \mathbb{N} \Rightarrow \bigcup_i A_i \in \mathcal{F}$.

Examples (i) $\mathcal{P}(\Omega)$ (ii) $\{\phi, A, A^c, \Omega\}$
Need more interesting examples

Prop Let $\{\mathcal{F}_\alpha \mid \alpha \in \mathcal{A}\}$ be any family of σ-algebras on Ω.
Then $\bigcap_{\alpha \in \mathcal{A}} \mathcal{F}_\alpha$ is a σ-algebra.

Proof. If $A \in \bigcap_\alpha \mathcal{F}_\alpha$ then $\forall \alpha$ $A \in \mathcal{F}_\alpha$ so $A^c \in \mathcal{F}_\alpha$ so $A^c \in \bigcap_\alpha \mathcal{F}_\alpha$. Similarly for $\bigcup_i A_i$. ∎

Thus Given any family $\mathcal{H}$ of subsets of Ω there is a least σ-algebra containing $\mathcal{H}$ = $\bigcap$ all σ-alg containing $\mathcal{H}$.

(iii) Let $(a,b) = \{ x \mid a < x < b \}$ be the family of open intervals of $\mathbb{R}$. The σ-algebra generated is called the <u>Borel</u> algebra.

(iv) Let (M, d) be a metric space, we define $r \in \mathbb{R}^{>0}, x \in M$ $B_r(x) := \{ u \in M \mid d(x,u) < r \}$. These are called open balls. The σ-algebra generated by the open balls is the Borel algebra of M.

(v) Coin tossing space Ω: the σ-algebra used is the one generated by the wine glasses. The same σ-algebra is generated by <u>cylinders</u>

~~Measures~~

$$C(n_1, \ldots n_k; b_1 \ldots b_k) := \{ \omega \mid \forall 1 \le i \le k \quad \omega_i = b_i \} .$$

<u>Measures</u> Given $(\Omega, \mathcal{F})$ a <u>measure</u> is a map $\mu : \mathcal{F} \to \mathbb{R}^{\geq 0}$ s.t.

(i) $\mu(\emptyset) = 0$ (ii) $A \subseteq B \Rightarrow \mu(A) \le \mu(B)$

(iii) $A_i \in \mathcal{F}$, pwd $\Rightarrow \mu(\bigcup_i A_i) = \sum_i \mu(A_i)$.

If $\mu(\Omega) < \infty$ we say μ is a finite measure.
If $\mu(\Omega) = 1$ we say μ is a probability measure
If $\mu(\Omega) \le 1$ we say μ is a subprobability measure
If $\mu(\Omega) = \infty$ but $\exists$ a countable family A_i s.t. $\bigcup_i A_i = \Omega$ & $\forall i \; \mu(A_i) < \infty$ we say μ is σ-finite.

<u>Examples</u> (1) ~~Pr~~ Coin tossing space
$Pr(s\uparrow) = 2^{-|s|}$ somehow extended.

(2) $(\Omega, \mathcal{P}(\Omega))$ & $\delta_{\omega_0}(A) = \begin{cases} 1 & \text{if } \omega_0 \in A \\ 0 & \text{if } \omega_0 \notin A \end{cases}$ DIRAC measure or point mass.

(3) $(\mathbb{R}, \text{Borel})$ λ : Borel sets $\to \mathbb{R}^{>0}$ by "length"
Lebesgue measure But what is it?

How to define measure on $\mathbb{R}$?

$$\lambda((a,b)) = b-a$$
$$\lambda([a,b)) = \lambda((a,b]) = \lambda([a,b]) = b-a.$$
$$\lambda(\mathbb{Q}) = ?$$

One approach: Given $S \subseteq \mathbb{R}$ we define a <u>cover</u> of S to be a set of intervals $\mathcal{J} = \{(a_i, b_i)\}_{i \in I}$ s.t. $\bigcup_{i \in I} (a_i, b_i) \supseteq S$.

$$\lambda(S) \le \sum_{i \in I} \lambda((a_i, b_i)) \quad \text{so} \quad \lambda(S) \le \inf_{\mathcal{J} \text{ covers } S} \lambda(\mathcal{J})$$

Now ~~λ~~ Consider $\mathbb{Q}$: enumerate $\mathbb{Q}$ so

$q_1, q_2, q_3, \cdots q_n, \cdots$

Give me a "budgt" of ϵ to cover $\mathbb{Q}$:

$(q_1 - \epsilon/4, q_1 + \epsilon/4), (q_2 - \epsilon/8, q_2 + \epsilon/8), (q_3 - \epsilon/16, q_3 + \epsilon/16), \cdots$

Length of the cover is $\sum_{i=1}^{\infty} \epsilon \left(\frac{1}{2}\right)^i = \epsilon$.

~~Now~~ So $\inf = 0$ & $\lambda(\mathbb{Q}) = 0$.

Why not <u>define</u> $\lambda(S) = \inf_{\mathcal{J} \text{ covers}} \lambda(\mathcal{J})$?

Call this λ^*

This is not countably additive

<u>Prop</u> $\lambda^*(\bigcup_i A_i) \le \sum \lambda^*(A_i)$

<u>Proof</u> Choose any $\epsilon > 0$ Since $\lambda^*(A_1)$ is defined as the inf over covering intervals we can ~~define~~ find $\mathcal{J}_1$ s.t.

$$\lambda(\mathcal{J}_1) \le \lambda^*(A_1) + \epsilon/2$$
$$\cdots \quad \lambda(\mathcal{J}_n) \le \lambda^*(A_n) + \epsilon/2^n$$

so $\lambda(\mathcal{J}) \le \sum \lambda^*(A_n) + \epsilon$ where $\mathcal{J}$ = union of $\mathcal{J}_n$.

Now $\mathcal{J}$ covers $\bigcup_i A_i$ so $\lambda^*(\bigcup_i A_i) \le \lambda(\mathcal{J}) \le \sum \lambda^*(A_i) + \epsilon$

Thus $\lambda^*(\bigcup_i A_i) \le \sum \lambda^*(A_i)$.

The ϵ - elbow room.

Def An outer measure is a function $\mu^*: \mathcal{P}(\Omega) \to \mathbb{R}^{\geq 0}$ s.t.
(i) $\mu^*(\phi) = 0$ (ii) $A \subseteq B \Rightarrow \mu^*(A) \leq \mu^*(B)$
(iii) $\mu^*(\bigcup_i A_i) \leq \sum_i \mu^*(A_i)$.

So we have constructed an outer measure which is defined on all sets but it is not a measure.

Thm Given Ω & $\mu^*: \mathcal{P}(\Omega) \to \mathbb{R}^{\geq 0}$ an outer measure we call a set E good if $\forall A \subseteq \Omega$ $\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c)$. The collection of good sets forms a σ-algebra & μ^* restricted to this σ-algebra is a measure.

Proof Omitted.

Vitali's example ~~Let us con~~ Thm There is no non-trivial translation-invariant measure that can be defined on all subsets of $\mathbb{R}$.

Proof ~~Supp~~ Let $x \sim y := x - y \in \mathbb{Q}$. Then $\sim$ is an eq. rel.^{n}
Define V by choosing one member from each equivalence class in the interval $[0,1]$.
$V \subseteq [0,1]$ so $\mu(V) \leq \mu([0,1]) = 1$.
$V + q := \{v + q \mid v \in V\}$ $q \in \mathbb{Q} \cap [0,1]$

Claim: $V + q_1 \cap V + q_2 = \phi$ if $q \neq q'$.
Suppose $x \in V + q_1 \cap V + q_2$ then $\exists v_1, v_2 \in V$ s.t.
$x = v_1 + q_1$ & $x = v_2 + q_2$ so
$v_1 = v_2 + q_2 - q_1$ so $v_1 \sim v_2$ $\otimes$.

Claim For any x in $(0,1)$ ~~$\exists v \in V$ s.t. $x \sim$~~ $\exists r \in (-1,1)$ s.t. $x \in V + r$: $\forall x \in (0,1)$ $\exists v \in V$ s.t. $x \sim v$ so. ~~so~~ $x - v \in \mathbb{Q}$, let $r = x - v$ so $x \in V + r$. Since both $x, v \in (0,1)$ $x - v \in (-1,1)$.

$S := \bigcup_{r \in (-1,1) \cap \mathbb{Q}} V + r$; $S \supseteq (0,1)$ & $S \subseteq (-1,2)$

Now $\mu(V+r) = \mu(V)$ since μ is tr.-inv.
$V+r$ are pw'd. Suppose $\mu(V) = \alpha > 0$

$$\mu(S) = \sum_{r\in(-1,1)} \alpha = \infty \quad \text{but } \mu(S) \le 3 \otimes.$$

Suppose $\mu(V) = 0$ then

$$\mu(S) = \sum 0 = 0 \quad \text{but } \mu(S) \ge 1 \otimes.$$

So $\mu(V)$ cannot be defined. ∎

~~Thus we are~~

This gives us the example we are looking for outer measure failing to be countably additive. So we have to choose well defined sets. The Borel sets can be measured. A much larger class of sets (the Lebesgue sets) can be measured as well but it is impossible to measure all sets.

—x—

Some properties of measures:

<u>Prop</u> (i) If $A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots \subseteq A_n \subseteq \cdots$ & $A = \bigcup_i A_i$ then $\lim_{i\to\infty} \mu(A_i) = \mu(A)$

(ii) If $A_1 \supseteq A_2 \supseteq \cdots A_n \supseteq \cdots$ & $A = \bigcap_i A_i$ then
$\lim_{i\to\infty} \mu(A_i) = \mu(A)$ if $\mu(A_1) < \infty$.

<u>Proof</u> Recall $A \subseteq B \Rightarrow \mu(A) \le \mu(B)$.
$B = A \cup (B\setminus A)$ so $\mu(B) = \mu(A) + \mu(B\setminus A) \ge \mu(A)$.

(i) We define a new family B_i by induction
$B_1 = A_1 \qquad B_{n+1} = A_{n+1} \setminus A_n \qquad B_i$ are pwd
& $\bigcup_i B_i = \bigcup_i A_i = A$ & $\bigcup_{i=1}^n B_n = A_n$ so $\mu(A) = \sum_{i=1}^{\infty} \mu(B_i)$

$$\mu(A_n) = \sum_{i=1}^n \mu(B_i)$$

$$\lim_{n\to\infty} \mu(A_n) = \lim_{n\to\infty} \sum_{i=1}^n \mu(B_i) = \sum_{i=1}^\infty \mu(B_i) = \mu(A)$$

(ii) Similarly for down continuity. But why $\mu(A_1) < \infty$?
Consider $A_i = (i, \infty)$ $\lambda(A_i) = \infty$ for all i so
$\lim_{i\to\infty} \lambda(A_i) = \infty$ but $\bigcap_i A_i = \phi = A$ so $\mu(A) = 0$.

Gas about Choquet capacity.

For any countable family of sets
$$\mu(\cup_i B_i) \leq \sum_i \mu(B_i).$$

——x——

Integration: What are we integrating?

<u>def</u> A function $f:(X,\Sigma)\to(Y,\Omega)$ is <u>measurable</u> if $\forall$ $B\in\Omega$ $f^{-1}(B)\in\Sigma$.

Functions that are not measurable are too wild.

Unfortunately even null is not enough e.g.
$$\int_{-\infty}^{\infty} 1\, dx = \infty$$

$$\int_{-\infty}^{\infty} x\, dx = ?$$

Let's not be too ambitious.

(Step 1). $\mathbb{1}_A(x) = \begin{cases} 1 & \text{if } x\in A \\ 0 & \text{if } x\notin A \end{cases}$ χ_A same notation.

$$\int_X \chi_A\, d\mu := \mu(A)$$

<u>Next</u> $s:(X,\Sigma)\to\mathbb{R}$ is a <u>simple</u> function if it has a <u>finite</u> range so $s(X)=\{a_1,\cdots a_n\}$.

$$A_i = s^{-1}(a_i)$$
$$s = \sum_{i=1}^{n} a_i\cdot\chi_{A_i}$$ note $A_i\cap A_j=\phi$ if $i\neq j$.

$$\int s\, d\mu = \sum_{i=1}^{n} a_i\cdot\mu(A_i)$$

what if $a_1<0$, $\mu(A_1)=\infty$ & $a_2>0$ & $\mu(A_2)=\infty$?

<u>Def</u> We say s is <u>integrable</u> if whenever $a\in\text{range}(s)$ $a\neq 0 \Rightarrow \mu(s^{-1}(a))<\infty$.

Then $\int s\, d\mu$ formula is sensible & taken as the def.

If $\{f_i\}_{i\in\mathbb{N}}$ is a family of mbl functions then f defined as pt wise limit is also measurable

$$f(x) := \lim_{i\to\infty} f_i(x).$$

Given a non-negative mbl function $f: X \to \mathbb{R}$ there is a family of simple functions s_i such that $s_i \le s_{i+1} \le f$ & $\{s_i\}$ converges ptwise to f.

Suppose f is a non-negative real valued f^n mbl f^n. We say f is <u>integrable</u> if the everywhere non-neg simple functions below f are integrable & their integrals are bounded. In that case we define

$$\int_X f \, d\mu = \bigvee_{\substack{s \le f \\ s \text{ simple}}} \int_X s \, d\mu \quad \text{where}$$

<u>Ex</u> $(X, \Sigma, \delta_{x_0})$ $\qquad \int s(x)\, d\delta_{x_0} = s(x_0)\, \delta_{x_0}(s^{-1}(s(x_0)))$
$\qquad = s(x_0)$

$$\int_X f \, d\delta_{x_0}$$

Now let s be any simple function below f so $\forall x \quad s(x) \le f(x)$ so $\int s \, d\delta_x = s(x_0) \le f(x_0)$.
Consider the function $t(x) = \begin{cases} f(x) & \text{if } x = x_0 \\ 0 & \text{otherwise} \end{cases}$ this is simple $\le f$.

So $\int t \, d\delta_{x_0} = f(x_0)$ so

$$\sup_{s \le f} \int s \, d\delta_{x_0} = f(x_0).$$

$$\int f \, d\delta_{x_0} = f(x_0).$$

<u>Properties</u> $\int (\alpha f + g)\, d\mu = \alpha \int f \, d\mu + \int g \, d\mu \qquad f \le g \Rightarrow \int f d\mu \le \int g \, d\mu$

$\forall x \in X.\ 0 \le f_1(x) \le f_2(x) \le \cdots \le \infty$ & $\bigvee f_n(x) = f(x)$ Then f is mbl & $\int f \, d\mu = \bigvee_n \int f_n \, d\mu = \lim_{n\to\infty} \int f_n \, d\mu$