\( (\Omega, F, P) \)

**Sequence of events** \( A_1, A_2, A_3, \ldots \)  
\[ S_m = \bigcup_{i=m}^{\infty} A_i \Rightarrow S_1 \supset S_2 \supset S_3 \supset S_4 \ldots \Rightarrow S_\infty = \bigcap_{m=1}^{\infty} S_m = \bigcap_{m=1}^{\infty} \bigcup_{i=m}^{\infty} A_i = \{A_n \text{ infinitely often}\} \]

(some book uses \( \sup_{i \geq m} A_i \))  
(Some book calls \( S_\infty = \lim_{n} \sup A_n \))

\[ T_m = \bigcap_{i=m}^{\infty} A_i \Rightarrow T_1 \subset T_2 \subset T_3 \subset T_4 \ldots \Rightarrow T_\infty = \bigcup_{m=1}^{\infty} T_m = \bigcup_{m=1}^{\infty} \bigcap_{i=m}^{\infty} A_i = \{A_n \text{ almost always}\} \]

(upward ladder)  
(Some book calls \( T_\infty = \lim_{n} \inf A_n \))

* \( \{A_n \text{ almost always}\}^c = \{A_n^c \text{ infinitely often}\} \)

(ex) Infinite coin tosses \( A_n = \text{n^{th} toss is head} \)

\[ \{A_n \text{ i.o.}\} = \{w : \text{outcome with infinitely many heads}\} \]

\[ \{A_n \text{ a.a.}\} = \{w : \text{outcome with finitely many tails}\} \]

* \( P(\{A_n \text{ a.a.}\}) = \lim_{m \to \infty} P(T_m) \leq P(\{A_n \text{ i.o.}\}) = \lim_{m \to \infty} P(S_m) \)

And in the case when they are equal (i.e. \( P(\{A_n \text{, a.a.}\}) = P(\{A_n \text{, i.o.}\}) \Rightarrow P) \)

\[ A_1, A_2, \ldots \Rightarrow S_\infty \text{ or } T_\infty \]

\[ \lim_{n \to \infty} P(A_i) = p \quad \text{ (Homework)} \]
ex/ Percolation Theory

\[ A_{ij} = \{ \text{atom } i \text{ and atom } j \text{ are connected} \} \]

Based on doping control \( P(A_{ij}) \)

\[ A = \{ \text{a connected path from top to bottom} \} \]

\[ = \bigcap_{i,j \text{ from top to bottom}} A_{ij} \]

ex/ Social network

\[ A_{ij} = \{ \text{individual } i \text{ and individual } j \text{ are connected through facebook} \} \]

\[ A = \{ \text{entire world is connected through facebook} \} \]

- random graph theory

ex/ Law of large numbers

- infinite coin toss with \( P(H) = \frac{1}{2} \)

- limit event that the probability of having half of the tosses being head is 1
  (or outcomes with less than (or more than) half head have probability 0)
Tail events - \( \bigwedge \{ A_n \} \) and \( \bigvee \{ A_n \} \) or any other that involve infinitely many \( A_n \)’s.

1. **Kolmogorov Zero-One Law.**
   If \( A_1, A_2, \ldots \) are independent, then for any tail event \( A \)
   \[ P(A) = 0 \quad \text{or} \quad P(A) = 1 \]

2. **Borel-Cantelli’s Lemma.**
   (1) If \( \sum_{n=1}^{\infty} P(A_n) \) converges (probability of each \( A_n \) is small enough)
   \[ P(\{ A_n \cap \Omega \}) = 0 \]

   (2) If \( \sum_{n=1}^{\infty} P(A_n) \) diverges AND \( A_n \)’s are independent
   \[ P(\{ A_n \cap \Omega \}) = 1 \]

---

**Random Variables**

ex \( \{ H, T \} \) instead we use \( X = \begin{cases} 0 & \text{if outcome is head} \\ 1 & \text{otherwise} \end{cases} \)

\( \Rightarrow \) talk about average.

ex/ toss 3 dice. \( X = \text{sum of the 3 dice} \)
\( \in \{3, \ldots, 18\} \)

\( w_1 = 3, w_2 = 3, w_6 = 1 \)

\( X = 7 \quad \overset{w_1 = 2, w_2 = 3, w_3 = 2}{\text{\checkmark}} \)
Random noise - uniform in $[0,1]$

$X = 8$-bit quantization of the noise

each $X = x$ represent $(a, b]$ of outcome.

Define a random variable $X$ as a function

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

Q: How do we map $(\Omega, \mathcal{F}, P)$ to $\mathbb{R}$?
A: transform the probability to the corresponding measurable subsets in $\mathbb{R}$.

$$X^{-1}(I) = \{ \omega \in \Omega \text{ such that } x(\omega) \in I \}$$

Instead of using $\text{length}(I)$, we define a new measure

$$\text{NEW length}(I) = P( X^{-1}(I) )$$

You need to have $X^{-1}(I) \in \mathcal{F}$

Formal definition: $X : \mathcal{S} \to \mathbb{R}$ is a random variable if for any Borel set $I \subset \mathbb{R}$, $X^{-1}(I) \in \mathcal{F}$
New length $(I) = P_F(X \in I) = P(X^{-1}(I))$

$(\Omega, \mathcal{F}, P) \xrightarrow{r.v. X} (\mathbb{R}, \mathcal{B}, P_F)$

Borel algebra = all measurable subsets in $\mathbb{R}$

ex/ $A \in \mathcal{F}$ Define indicator variable.

$$1_A(w) = \begin{cases} 1 & \text{w} \in A \\ 0 & \text{w} \notin A \end{cases}$$

$(\Omega, \mathcal{F}, P)$ Induced probability space in $\mathbb{R}$

$(\mathbb{R}, \mathcal{B}, P_F)$

More precisely: $\{0,1\}, \{\emptyset, \{0,1\}, \{0\}, \{1\}\}$

Why it is useful? More later

$P_F(\emptyset) = P(1_A^{-1}(\emptyset)) = P(\emptyset) = 0$

$P_F(\{0,1\}) = P(1_A^{-1}(\{0,1\})) = P(\Omega) = 1$

$P_F(\{0\}) = P(1_A^{-1}(\{0\})) = P(A^c) = 1 - P(A)$

$P_F(\{1\}) = P(1_A^{-1}(\{1\})) = P(A)$

Average $1_A$

$= 1 \cdot P(1_A(w) = 1) + 0 \cdot P(1_A(w) = 0)$

$= P(A)$
Two random variables $X : \Omega \rightarrow \mathbb{R}$
$Y : \Omega \rightarrow \mathbb{R}$

What does it mean by $X = Y$?

Strict sense $X(w) = Y(w)$ for all $w \in \Omega$

Broad sense $X = Y$ "almost surely"

if $P(\{w : X(w) \neq Y(w)\}) = 0$

ex/ $(\Omega, \mathcal{F}, P)$ is uniform distribution on $[0, 1]$

\[
X(w) = 1 \quad \quad Y(w) = \begin{cases} 0 & \text{if } w \in \mathbb{Q} \text{ (rational number)} \\ 1 & \text{otherwise} \end{cases}
\]

But $X = Y$ almost surely because

$P(\{w : X(w) \neq Y(w)\}) = P(w \in \mathbb{Q} \land [0, 1])$

$= P\left( \bigcup_{i=1}^{\infty} \{q_i\} \right)$

$= \sum_{i=1}^{\infty} P(\{q_i\}) = \sum_{i=1}^{\infty} 0 = 0$