IMPORTANT PROBABILITY FACTS AND IDEAS:

A **probability space** is a triple $(\Omega, \mathcal{F}, \mathbb{P})$ where Ω is the sample space, $\mathcal{F}$ is a set of subsets of Ω (forming a σ-algebra) and $\mathbb{P} : \mathcal{F} \to \mathbb{R}$ is a **probability measure** satisfying:

- $\mathbb{P}[\Omega] = 1$
- $\mathbb{P}[A^c] = 1 - \mathbb{P}[A]$ for $A \in \mathcal{F}$.
- If $A_1, \ldots$ is countable sequence of disjoints element of $\mathcal{F}$ then

$$\mathbb{P}[\bigcup_i A_i] = \sum_i \mathbb{P}[A_i].$$

A **random variable** is a function $X : \Omega \to \mathbb{R}$ (that is $\mathcal{F}$ measurable). Two random variables X, Y are **equal in distribution** or **identically distributed** if for all $A \subset \mathbb{R}$ (measurable)

$$\mathbb{P}[X \in A] = \mathbb{P}[Y \in A].$$

The **expectation** of X is denoted $\mathbb{E}[X]$ and satisfies

- Linear $\mathbb{E}[cX + dY] = c\mathbb{E}[X] + d\mathbb{E}[Y]$ for $c, d \in \mathbb{R}$.
- If $0 \leq X \leq M$ then $0 \leq \mathbb{E}[X] \leq M$.
- For $c \in \mathbb{R}$, $\mathbb{E}[c] = c$.
- If $X \leq Y$ then $\mathbb{E}[X] \leq \mathbb{E}[Y]$.
- $\mathbb{E}[|X|] \geq |\mathbb{E}[X]|$.
- If ϕ is convex then $\mathbb{E}[\phi(X)] \geq \phi(\mathbb{E}[x])$.
- If X and Y are independent then $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$.

A sequence of random variables X_n **converges in probability** to X if for all $\epsilon > 0$

$$\mathbb{P}[|X_n - X| > \epsilon] \to 0.$$

It **converges almost surely** to X if

$$\mathbb{P}[\{\omega : X_n(\omega) \to X(\omega)\}] = 1.$$

We say X_n **converges weakly** or **converges in distribution** to X if for all bounded continuous functions $f(x)$,

$$\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)].$$

Equivalently (in dimension 1)

$$\mathbb{P}[X_n \leq x] \to \mathbb{P}[X \leq x]$$

for all continuity points of $F(x) = \mathbb{P}[X \leq x]$, that x with $\mathbb{P}[X = x] = 0$.

The **Weak Law of Large Numbers** says that if X_n are **independent and identically distributed** (IID) with $\mathbb{E}[|X_i|] < \infty$ and $\mathbb{E}[X_i] = \mu$ then $\frac{1}{n}\sum_{i=1}^n X_i \to \mu$ in probability. The **Strong Law of Large Numbers** is the same but for almost sure convergence.

The Central Limit Theorem says that if X_i are IID with mean μ and variance $\mathbb{E}[X_i^2] = \sigma^2 < \infty$ then

$$\frac{\sum_{i=1}^n X_i - \mu n}{\sigma\sqrt{n}} \to N(0,1)$$

where the convergence is in distribution.

For a random variable X the **conditional expectation with respect to an event** A is

$$\mathbb{E}[X|A] := \frac{\mathbb{E}[XI(A)]}{\mathbb{P}[A]}$$

The **conditional expectation with respect to a random variable** Y (if Y is a discrete random variable) is

$$\mathbb{E}[X|Y] := \psi(Y), \qquad \psi(y) = \mathbb{E}[X|Y = y].$$

The **conditional expectation with respect to a σ-algebra** $\mathcal{G}$ is the $\mathcal{G}$-measurable random variable such that for all $A \in \mathcal{G}$,

$$\mathbb{E}[XI(B)] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}]I(B)].$$

It has the following properties

- Linear $\mathbb{E}[cX + dY] = c\mathbb{E}[X] + d\mathbb{E}[Y]$ for $c, d \in \mathbb{R}$.
- **Tower Property:** If $\mathcal{G} \subset \mathcal{H}$ then

$$\mathbb{E}[\mathbb{E}[X \mid \mathcal{H}] \mid \mathcal{G}] = \mathbb{E}[X \mid \mathcal{H}].$$

- When $\mathcal{G}$ is the trivial σ-algebra we have that

$$\mathbb{E}[\mathbb{E}[X \mid \mathcal{H}]] = \mathbb{E}[X].$$

- If X is $\mathcal{G}$ measurable then $\mathbb{E}[X \mid \mathcal{G}] = X$ and $\mathbb{E}[XY \mid \mathcal{G}] = X\mathbb{E}[Y \mid \mathcal{G}]$.
- If X is independent of $\mathcal{G}$ then $\mathbb{E}[X \mid \mathcal{G}] = \mathbb{E}[X]$.

A **filtration** is a sequence of increasing σ-algebras $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \ldots$. This represents an increasing amount of information. A sequence X_n is a **martingale** with respect to $\mathcal{F}_n$ if each X_n is $\mathcal{F}_n$ measurable and for all n,

$$\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] = X_n.$$

Often the filtration will be the generated by $X_1, \ldots, X_n$ and then the definition is

$$\mathbb{E}[X_{n+1} \mid X_1, \ldots, X_n] = X_n.$$

For all $n > m$, $\mathbb{E}[X_n \mid \mathcal{F}_m] = X_m$ and $\mathbb{E}[X_n] = \mathbb{E}[X_0]$. For a continuous family X_t it is a martingale if $\mathbb{E}[X_t \mid \mathcal{F}_s] = X_s$ for all $0 \le s < t$.