- $\operatorname{Var}(X) = E\left[X^2\right] - E[X]^2$.
- $\operatorname{Cov}(X,Y) \coloneqq E[XY] - E[X]E[Y] = E\big[(X - E[X])(Y - E[Y])\big]$
- **Markov's Inequality**: if $X \geq 0$, then $P(X \geq a) \leq \frac{E[X]}{a}$ for all $a > 0$.
- **Chebyshev's Inequality**: for finite variance, $P\big(g(X) > a^2 \operatorname{Var}(X)\big) \leq \frac{E[g(X)]}{a^2 \operatorname{Var}(X)}$.
- Law of total probability: if $\{B_i : i \in I\}$ is a partition of the sample space, then

$$P(A) = \sum_{i \in I} P(A \mid B_i) P(B_i).$$

- Law of total expectation: $E[X] = E\big[E[X \mid Y]\big]$.
- Law of total variance: $\operatorname{Var}(X) = E[\operatorname{Var}(X \mid Y)] + \operatorname{Var}(E[X \mid Y])$.
- Bayes's Theorem: $P(A \mid B) = \frac{P(B|A)P(A)}{P(B)}$.
- Markov chain: future is independent of the past, i.e., X_{n+1} is at most dependent on X_n.
- Transition probability: $p_{ij}^{n,m} = P(X_m = j \mid X_n = i)$.
- Transition probability matrix: if π_i is the distribution of X_i, then $\pi_t = \pi_0 \prod_{i=0}^{t-1} \boldsymbol{P}^{i,i+1}$.
- Stationary Markov chain: transition probability matrix is independent of the timestamp n.
- **Stochastic matrix**: non-negative matrix such that row sums are 1.
- **Chapman-Kolmogorov Equations**: $\boldsymbol{P}^{(m)} = \boldsymbol{P}\boldsymbol{P}^{(m-1)} = \boldsymbol{P}^{(m-1)}\boldsymbol{P}$. If X is stationary, then for all $m, n \in \mathbb{N}$,

$$P_{ij}^{0,m+n} = \sum_{k \in S} P_{ik}^{0,m} P_{kj}^{0,n},$$

and $P(X_n = j \mid X_0 = i) = (\boldsymbol{P}^n)_{ij}$.

- if we have a column vector

$$\mu \coloneqq \begin{bmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_s) \end{bmatrix}$$

for some function f on the state space, then

$$\begin{aligned}(\boldsymbol{P}^n \mu)_i &= \sum_{j=1}^{s} \boldsymbol{P}_{ij}^n \mu_j \\ &= \sum_{j=1}^{s} P(X_n = x_j \mid X_0 = x_i) f(x_j) \\ &= E[f(X_n) \mid X_0 = x_i].\end{aligned}$$

Suppose $X_0 \sim \lambda$, then clearly

$$\begin{aligned}E[f(X_n)] &= \sum_{i=1}^{s} E[f(X_n) \mid X_0 = x_i] \lambda_i \\ &= \sum_{i=1}^{s} \lambda_i (\boldsymbol{P}^n \mu)_i \\ &= \lambda \boldsymbol{P}^n \mu.\end{aligned}$$

- **Stationary distribution**: $\pi = \pi \boldsymbol{P}$.
- If λ is an eigenvalue of $\boldsymbol{P}$, then $|\lambda| \leq 1$.
- **Absorbing state**: for all $j \neq i$, we have $P_{ij} = 0$.
- **Intercommunicating states**: $\exists m, n \in \mathbb{N}$ such that $P_{xy}^{(m)}, P_{yx}^{(n)} > 0$.
- **Irreducible chain**: all states intercommunicate, i.e., only one class.
- **Return probability**: $P_{ii}^{(n)} = P(X_n = i \mid X_0 = i)$.
- **First return probability**: $f_{ii}^{(n)} = P(X_1 \neq i, \cdots X_{n-1} \neq i, X_n = i \mid X_0 = i)$. $f_{ii}^{(0)} = 0$ and $f_{ii}^{(n)} \leq P_{ii}^{(n)}$.
- $P_{ii}^{(n)} = \sum_{k=0}^{n} f_{ii}^{(k)} P_{ii}^{(n-k)}$.
- $f_{ii} = \sum_{n=0}^{\infty} f_{ii}^{(n)}$ is the probability of returning to i in finite time. i is **recurrent** if $f_{ii} = 1$ and **transient** if $f_{ii} < 1$.
- For any recurrent state i, $P\left(\sum_{n=1}^{\infty} I\{X_n = i\} = \infty \mid X_0 = i\right) = 1$ and so

$$E\left[\sum_{n=0}^{\infty} I\{X_n = x\} \mid X_0 = x\right] = \sum_{n=0}^{\infty} \boldsymbol{P}^n(x,x) = \infty,$$

but $\boldsymbol{P}^n(x,x)$ may converge to 0.

- Number of revisits to i: $N_i \sim \operatorname{Geo}(1 - f_{ii})$. $E[N_i \mid X_0 = i] = \frac{f_{ii}}{1-f_{ii}}$ and expected number of visits including the initial one is $\frac{1}{1-f_{ii}}$.

- i is a transient state iff $\sum_{n=1}^{\infty} P_{ii}^{(n)}$ is finite.
- If i is transient, then $\lim_{m\to\infty} \sum_{n=m}^{\infty} P_{ii}^{(n)} = 0$ by monotone convergence theorem.
- All finite-state irreducible chains are recurrent.
- Reducible chains will enter one of the recurrent classes in the long-run.
- **Period**: $d(i) := \gcd \left\{ n \in \mathbb{N}^+ : P_{ii}^{(n)} > 0 \right\}$. i is **aperiodic** iff $d(i) = 1$.
- If $i \leftrightarrow j$, then $d(i) = d(j)$.
- $\forall i \in S$, $\exists N \in \mathbb{N}$ such that $\forall n \geq N$, $P_{ii}^{(n \cdot d(i))} > 0$ and $P_{ji}^{(m+n \cdot d(i))} > 0$ whenever $P_{ji}^{(m)} > 0$.
- If $\boldsymbol{P}$ is the TPM for a finite-state irreducible aperiodic chain, then $\exists N \in \mathbb{N}^+$ such that $\boldsymbol{P}^{(N)}$ has all positive entries (definition for regular chain).
- Irreducible, aperiodic, finite-state $\implies$ regular chain.
- Regular $\implies$ irreducible.
- $\boldsymbol{P}^{(k)}$ is regular $\implies$ $\boldsymbol{P}^{(n)}$ is regular $\forall n \geq k$.
- Let $\boldsymbol{P}$ be a regular transition probability matrix for some regular Markov chain with state space $S := \{1, 2, \cdots, N\}$, then

  1. the limit $\pi_j := \lim_{n\to\infty} P_{ij}^{(n)}$ exists and is independent of i;
  2. $\sum_{i=1}^{N} \pi_i = 1$ and $\pi := \left(\pi_1, \pi_2, \cdots, \pi_N\right)$ satisfies $\pi \boldsymbol{P} = \pi$;
  3. π is unique.

  π_j is the marginal probability $P(X_n = j)$ in the long-run
- **Stopping time**: $T_A := \min \{n \in \mathbb{N} : X_n \in A\}$ is the first time X enters A.
- If $f(x) = P(T_A < T_B \mid X_0 = x)$, then for all $x \notin A \cup B$,

$$
\begin{aligned}
f(x) &= \sum_{y \in S} P(T_A < T_B \mid X_1 = y, X_0 = x) P(X_1 = y \mid X_0 = x) \\
&= \sum_{y \in S} P(T_A < T_B \mid X_0 = y) P(X_1 = y \mid X_0 = x) \\
&= \sum_{y \in S} P_{xy} f(y).
\end{aligned}
$$

- **First-step analysis**:

  1. Identify quantity of interest $a_i(T) = h(i, X_1, \cdots, X_T \mid X_0 = i)$.
  2. Consider $a_i(T) = \sum_{k \in S} h(\cdot \mid X_1 = k, X_0 = i) P(X_1 = k \mid X_0 = i)$.
  3. Consider $Y_n = X_{n+1}$ and establish $h(\cdot \mid X_1 = k, X_0 = i) = g_i\big(a_k(T)\big)$.
  4. Solve the system.

- Gambler's ruin: $X_0 = k$ for $0 < k < N$ with winning probability p.
  - Fair game: $P(X_T = 0 \mid X_0 = k) = 1 - \frac{k}{N}$ and $E[T \mid X_0 = k] = k(N-k)$.
  - Otherwise:

$$
\begin{aligned}
P(X_T = 0 \mid X_0 = k) &= 1 - \frac{1 - (q/p)^k}{1 - (q/p)^N}, \\
E[T \mid X_0 = k] &= \frac{1}{p-q} \left[\frac{N\left(1 - (q/p)^k\right)}{1 - (q/p)^N} - k \right]
\end{aligned}
$$

- Random walk: $\frac{\xi_i + 1}{2} \sim \text{Bernoulli}(p)$, $\frac{X_n + n}{2} \mid X_0 = 0 \sim \text{Bin}(n, p)$.