Expectation & Variance

- **Law of Total Probability:** $P(A) = \sum_n P(A \mid B_n)P(B_n)$
- **Tower Property (Expectation):** $E[X] = E[E[X \mid Y]]$
- **Conditional Variance Formula:**

$$\mathrm{Var}(X) = E[\mathrm{Var}(X \mid Y)] + \mathrm{Var}(E[X \mid Y])$$

- **Indicator Variables:** I_A. $E[I_A] = P(A)$.
- $\mathrm{Var}(X) = E[X^2] - (E[X])^2$

Probability Generating Functions (PGF)

For discrete RV $X \in \{0, 1, \dots\}$:

$$G_X(s) = E[s^X] = \sum_{k=0}^{\infty} p_k s^k$$

Key Properties:

- $G(1) = 1$ and $G(0) = P(X = 0) = p_0$.
- **Mean:** $\mu = E[X] = G'(1)$.
- **Variance:** $\sigma^2 = G''(1) + G'(1) - [G'(1)]^2$.
- **Sums:** If $Z = X + Y$ (indep), $G_Z(s) = G_X(s)G_Y(s)$.
- **Random Sums:** If $S = \sum_{i=1}^{N} X_i$ $(N, X_i$ indep),

$$G_S(s) = G_N(G_X(s))$$

Differentiation rule: $\dfrac{d}{ds}G_S(1) = E[N]E[X]$.

Definitions

- **Markov Property:** Past is irrelevant given the present.
- **CK Equations:** $P^{(n+m)} = P^{(n)}P^{(m)} = P^{n+m}$.

State Classification

- **Communication:** $i \leftrightarrow j$ if $\exists n, m$ s.t. $P_{ij}^{(n)} > 0$ and $P_{ji}^{(m)} > 0$.
- **Class:** A set of states that all communicate.
- **Irreducible:** The entire state space is one class.
- **Periodicity** $d(i)$**:** gcd of all n such that $P_{ii}^{(n)} > 0$.
  - If $d(i) = 1 \Rightarrow$ **Aperiodic**.
  - If $P_{ii} > 0 \Rightarrow$ Aperiodic.
- **Recurrence:** State i is recurrent if $\sum_{n=1}^{\infty} P_{ii}^{(n)} = \infty$.
  - **Positive Recurrent:** expected return time $\mu_{ii} < \infty$.
  - **Null Recurrent:** return prob $= 1$, but $\mu_{ii} = \infty$.
  - Finite, irreducible MC $\Rightarrow$ positive recurrent.
- **Transient:** returns with prob < 1. $\sum P_{ii}^{(n)} < \infty$.

Existence & Uniqueness

π is stationary if $\pi = \pi P$ and $\sum_i \pi_i = 1$.

- **Global Balance Equations:** $\pi_j = \sum_i \pi_i P_{ij}$.
- **Ergodic Theorem:** If MC is irreducible + aperiodic + positive recurrent, then a unique π exists and:

$$\lim_{n\to\infty} P_{ij}^{(n)} = \pi_j$$

- **Interpretation:** π_j is the long-run proportion of time in state j.
- **Relation to Return Time:** $\pi_j = \dfrac{1}{E[T_j \mid X_0 = j]}$.

Shortcuts & Special Cases

- **2-State Chain:** $P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$.

$$\pi = \left(\frac{\beta}{\alpha+\beta}, \frac{\alpha}{\alpha+\beta}\right)$$

- **Doubly Stochastic:** columns sum to 1 (and rows sum to 1), finite/irreducible: $\pi_j = \dfrac{1}{N}$ (uniform).
- **Birth-Death Process:** only $i \to i+1$ (p_i) and $i \to i-1$ (q_i).

$$\pi_k = \pi_0 \prod_{i=0}^{k-1} \frac{p_i}{q_{i+1}}$$

  Use $\sum_k \pi_k = 1$ to find π_0.

Reversibility (Detailed Balance)

Reversible if $\exists \pi$ s.t.

$$\pi_i P_{ij} = \pi_j P_{ji} \quad \forall i, j$$

- If this holds, π is the stationary distribution.
- **Hint:** Try this first for random walks on graphs or birth-death processes.

Graph $G = (V, E)$. $d_i =$ degree of i.

- **Unweighted RW:** $P_{ij} = 1/d_i$ if $j \sim i$.
- **Stationary Dist:** $\pi_i = \dfrac{d_i}{\sum_j d_j} = \dfrac{d_i}{2|E|}$.
- **Weighted RW:** weight w_{ij}. $d_i = \sum_k w_{ik}$.

$$P_{ij} = \frac{w_{ij}}{d_i}, \qquad \pi_i = \frac{d_i}{\sum_k d_k}$$

Used to find absorption probabilities and mean time to absorption.

Technique

Condition on first step $X_0 \to X_1$.

$$u_i = \sum_{j\in S} P_{ij} u_j$$

u_i is the value (prob or time) starting at i.

Type 1: Hitting Probability

Let A be a set of target states. Let $h_i = P(\text{chain hits } A \mid X_0 = i)$.

- **Eq 1 (Boundary):** If $i \in A$, $h_i = 1$.
- **Eq 2 (Boundary):** If i absorbing and $i \notin A$, $h_i = 0$.
- **Eq 3 (Recursive):** For transient i:

$$h_i = \sum_j P_{ij} h_j$$

Type 2: Mean Hitting Time

Let $T = \min\{n \geq 0 : X_n \in A\}$. Let $m_i = E[T \mid X_0 = i]$.

- **Eq 1 (Boundary):** If $i \in A$, $m_i = 0$.
- **Eq 2 (Recursive):** For $i \notin A$:

$$m_i = 1 + \sum_j P_{ij} m_j$$

Gambler's Ruin (1D Random Walk)

State space $\{0, \dots, N\}$, absorbing at $0, N$. $p = P(i \to i+1)$, $q = P(i \to i-1)$.

1. Probability of reaching N before 0 (h_i):

- **Fair** $(p = q = 0.5)$**:** $h_i = \dfrac{i}{N}$.
- **Biased** $(p \neq q)$**:** let $\rho = q/p$.

$$h_i = \frac{1-\rho^i}{1-\rho^N}$$

2. Mean time to absorption (m_i):

- **Fair:** $m_i = i(N - i)$.
- **Biased:**

$$m_i = \frac{1}{q-p}\left[i - N\left(\frac{1-\rho^i}{1-\rho^N}\right)\right]$$

If a chain has absorbing and transient states, reorder so transient (T) first, absorbing (A) last:

$$P = \begin{pmatrix} Q & R \\ 0 & I \end{pmatrix}$$

- Q: Trans $\to$ Trans transitions.
- R: Trans $\to$ Abs transitions.

Fundamental Matrix N

$$N = (I - Q)^{-1} = I + Q + Q^2 + \dots$$

Interpretation: N_{ij} is the expected number of visits to transient state j starting from transient state i before absorption.

Computations using N

- **Time to Absorption: t** = expected steps to absorption:

$$\mathbf{t} = N\mathbf{1} \quad \text{(row sums of } N\text{)}$$

- **Absorption Probabilities:** $B_{ij} = P(\text{absorbed in state } j \mid \text{start at } i)$:

$$B = NR$$

X_n = population size at gen n. $X_0 = 1$. Offspring distribution ξ has PGF $G(s)$, mean μ, var σ^2.

Standard Relations

- $X_{n+1} = \sum_{i=1}^{X_n} \xi_i^{(n)}$.
- $G_n(s) = G_{n-1}(G(s)) = G(G(\dots G(s)))$.
- $E[X_n] = \mu^n$.
- $P(X_n = 0) = G_n(0)$.

Variance of X_n

- **Case $\mu = 1$:** $\mathrm{Var}(X_n) = n\sigma^2$.
- **Case $\mu \neq 1$:**

$$\mathrm{Var}(X_n) = \frac{\sigma^2 \mu^{n-1}(\mu^n - 1)}{\mu - 1}$$

Extinction Probability (π_0)

Let $\pi_0 = P(\text{eventual extinction}) = \lim_{n\to\infty} P(X_n = 0)$.

- π_0 is the **smallest non-negative solution** to:

$$s = G(s)$$

- **Subcritical ($\mu < 1$):** $\pi_0 = 1$.
- **Critical ($\mu = 1$):** $\pi_0 = 1$ (unless $P(\xi = 1) = 1$).
- **Supercritical ($\mu > 1$):** $\pi_0 < 1$. Roots are 1 and π_0.
- **Total Progeny T:** $P(T = k) = \dfrac{1}{k} P(S_k = k - 1)$ (RW mapping).

Geometric Series

- $\sum_{k=0}^{\infty} r^k = \dfrac{1}{1-r}$ for $|r| < 1$.
- Finite: $\sum_{k=0}^{n} r^k = \dfrac{1 - r^{n+1}}{1 - r}$.
- Deriv: $\sum_{k=1}^{\infty} k r^{k-1} = \dfrac{1}{(1-r)^2}$.

Binomial

- $(x + y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$.
- $\binom{n}{k} = \dfrac{n!}{k!(n-k)!}$.

Linear Algebra

- Eigenvalues: $\det(P - \lambda I) = 0$.
- Left Eigenvector: $vP = \lambda v$. (Stationary π is left EV for $\lambda = 1$.)
- Matrix Inverse (2×2):

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

1. **Identify the State Space:** finite? infinite? graph?
2. **Classify:** reducible? periodic?
3. **Stationary Dist:**
   - If 2×2, use shortcut.
   - If graph/birth-death, use reversibility ($\pi_i P_{ij} = \pi_j P_{ji}$).
   - Otherwise, solve $\pi = \pi P$.
4. **Hitting Times/Probs:**
   - Define h_i or m_i.
   - Set boundary conditions (target = 0 or 1).
   - Write linear equations for transients.
   - If asking for return time to i, compute π_i then $\mu_{ii} = 1/\pi_i$.
5. **Pattern Recognition:**
   - "Population", "Offspring" $\to$ branching process.
   - "Betting", "Ruin" $\to$ gambler's ruin formulas.
   - "Long run proportion" $\to$ stationary distribution.