Probability 2 - Notes

Dom Hutchinson

May 1, 2019

Contents

1 Introduction

1.1 The Probability Triple

Definition 1.1 - *Sample Space,* Ω
A *Sample Space*, Ω, is the set of all possible outcomes.

Definition 1.2 - *Sigma Field,* $\sigma - Field$
A *Sigma Field*, $\mathcal{F}$, of subsets of a sample space Ω satisfies the following conditions

i) $\emptyset \in \mathcal{F}$;

ii) If $A_1, A_2, \dots \in \mathcal{F}$ then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$; And,

iii) If $A \in \mathcal{F}$ then $A^c \in \mathcal{F}$ where $A^c := \Omega \setminus A$.

Definition 1.3 - *Probability Space*
A *Probability Space* is a triple $(\Omega, \mathcal{F}, \mathbb{P})$.

1.2 The Sigma Field

Definition 1.4 - *$\mathcal{F}$-measurable*
Events in $\mathcal{F}$ are said to be $\mathcal{F}$-measurable.
If an event A is $\mathcal{F}$-measurable then the information in $\mathcal{F}$ is enough to determine whether, or not, A has occurred.
If a function f is $\mathcal{F}$-measurable then then the information in $\mathcal{F}$ is enough to determine to value of f.
N.B. Occasionally this is referred to simply as *measurable.*

Remark 1.1 - *Sigma Fields from Collection of Events*
The σ-field generated by a collection of events $\mathcal{C}$, $\sigma(\mathcal{C})$, is the smallest σ-field that contains $\mathcal{C}$.
N.B. This is the intersection of all σ-fields containing events of $\mathcal{C}$.

Definition 1.5 - *Power Set*
The *Power Set* of set S, 2^S, is a set that consist of all subsets of S.

Remark 1.2 - *Binary Representation of Power Set*
A *Power Set* can be represented by a binary table where there is a unique column for each element and then each row reads as a different binary value.
If the value in $A_{ij} = 1$ then a_i is in subset j.
Else, if the value in $A_{ij} = 0$ then a_i is *not* in subset j.

Example 1.1 - *Binary Representation of Power Set*
Here is a binary representation of the power set of $\Omega = \{\omega_1, \omega_2, \omega_3\}$.

ω_1	ω_2	ω_1	
0	0	0	$\emptyset$
0	0	1	$\{\omega_3\}$
0	1	0	$\{\omega_2\}$
0	1	1	$\{\omega_2, \omega_3\}$
1	0	0	$\{\omega_1\}$
1	0	1	$\{\omega_1, \omega_3\}$
1	1	0	$\{\omega_1, \omega_2\}$
1	1	1	$\{\omega_1, \omega_2, \omega_3\}$

Remark 1.3 - *Individual Events in $\mathcal{F}$*
Let $\omega_1, \omega_2 \in \Omega$ be different events & $\mathcal{F}$ be a σ-field on Ω.
We can only distinguish between ω_1 & ω_2 in $\mathcal{F}$ if they are in distinct elements of $\mathcal{F}$.
N.B. The converse does not hold.

Remark 1.4 - *All σ-Fields have a disjoint subset that form the Population*
If Ω is a finite set, then given any σ-field $\mathcal{F}$ on Ω there exists a finest partition $\mathcal{P}$ of Ω under $\mathcal{F}$.
N.B. $\mathcal{P} = \{A_1, \dots, A_n\}$ st $\bigcup_{i=1}^n A_i = \Omega$ & $A_i \cap A_j = \emptyset \ \forall i, j \in \mathbb{N}$.
Example 1.2 - *ω-Field*
Consider the scenario in which two coins are tossed and each value is recorded.
We have that $\Omega = \{HH, HT, TH, TT\}$.
Here are some possible ω-fields.

$$\begin{array}{rcll} \mathcal{F} & = & \sigma(\{HH, HT\}) & \\ & = & \{\emptyset, \{HH, HT\}, \{TT, TH\}, \Omega\} & \text{This encodes information about the first toss.} \\ \mathcal{F}_2 & = & \sigma(\{HH, TH\}) & \\ & = & \{\emptyset, \{HH, TH\}, \{TT, HT\}, \Omega\} & \text{This encodes information about the second toss.} \\ \mathcal{F}_{1,2} & = & 2^\Omega & \text{This encodes information about both tosses.} \end{array}$$

Definition 1.6 - *Borel σ-Field*
The *Borel σ-Field* is used for the uncountable set $\Omega = [0, 1]$.
It is generated by all possible open subintervals of form $(a, b) \subset [0, 1],\ 0 \leq a < b \leq 1$.

Theorem 1.1 - *Subsets of σ-Fields*
Here are three similar theorems about the subset of a σ-field.

i) Arbitrary intersections of a σ-field are σ-field.

ii) For any $\mathcal{C}$ consisting of subsets of Ω, $\sigma(\mathcal{C})$ is a σ-field.

iii) The power set of Ω is a σ-field.

Proof 1.1 - *Theorem 1.1*
Let $\mathcal{F}_n$ be a collection of σ-fields for n in some indexing set.
We need to verify the three defining axioms of a σ-field for $\bigcap\limits_n \mathcal{F}_n$.

$$\begin{array}{rl} i) & \text{Since } \mathcal{F}_n \text{ is a } \sigma-\text{field, } \emptyset \in \mathcal{F}_n \\ & \text{Hence } \emptyset \in \bigcap\limits_n \mathcal{F}_n \\ ii) & \text{Take } A_1, A_2, \dots \in \bigcap\limits_n \mathcal{F}_n \\ & \text{Then } \forall\ i, n \text{ we have } A_i \in \mathcal{F}_n \\ & \text{By second axiom } \bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{F}_n \\ & \text{Hence } \bigcup\limits_{i=1}^{\infty} A_i \in \bigcap\limits_n \mathcal{F}_n \\ iii) & \text{Take } A \in \bigcap\limits_n \mathcal{F}_n \\ & \text{Then } \forall\ n \text{ we have } A \in \mathcal{F}_n \\ & \text{Since} \mathcal{F}_n \text{ is a } \sigma-\text{field then } A^c \in \mathcal{F}_n \\ & \text{Hence } A^c \in \bigcap\limits_n \mathcal{F}_n \end{array}$$

Since all axioms hold then $\bigcap\limits_n \mathcal{F}_n$ is a σ-field.
So theorem *i)* holds.
Theorem *ii)* is a direct consequence of *i)* so holds.

For Theorem *iii)* we need to check all axioms hold.

$$\begin{array}{rl} i) & \emptyset \in 2^\Omega \text{ since } \emptyset \in \Omega \\ ii) & \text{Take } A_1, A_2 \in 2^\Omega \text{ then } A_i \subset \Omega\ \forall\ i. \\ & \text{Since } \bigcup\limits_{i=1}^{\infty} A_i \subset \Omega \implies \bigcup\limits_{i=1}^{\infty} A_i \in 2^\Omega. \\ iii) & \text{Take } A \in 2^\Omega \text{ then } A \subset \Omega. \\ & \text{But } A^C = \Omega \backslash A \subset \Omega \implies A^c \in 2^\Omega \end{array}$$

Since all three axioms hold we conclude that 2^Ω is a σ-field.

Definition 1.7 - *Probability Measure*
A *Probability Measure* $\mathbb{P}$ is a function $\mathbb{P} : \mathcal{F} \to [0, 1]$ which satisfies the following axioms

i) $\mathbb{P}(\emptyset) = 0$;

ii) $\mathbb{P}(\Omega) = 1$; And,

iii) If $A_1, A_2, \dots \in \mathcal{F}$ are pair-wise disjoint then $\mathbb{P}\left(\bigcup\limits_{i=1}^{\infty} A_i\right) = \sum\limits_{i=1}^{\infty} \mathbb{P}(A_i)$.
This is called σ-additivity.

1.3 Definitions of Stochastic Processes

Definition 1.8 - *Filtration*
A *Filtration* is a family of σ-fields, $\{\mathcal{F}_t : t \geq 0\}$, such that $\mathcal{F}_{t_1} \subset \mathcal{F}_{t_2}$.

Definition 1.9 - *Stochastic Process*
For any set $\Delta \subseteq \mathbb{R}$, a collection $\{X_t\}_{t\in\Delta}$ of random variables is called a *Stochastic Process.*
N.B. This indexing set may be continuous or discrete. Typically X_n denotes a discrete time process & X_t a continuous time process.

Definition 1.10 - *Adapted Stochastic Processes*
An *Adapted Stochastic Process* is a *Stochastic Process* that cannot see into the future.
Each *Stochastic process*, X, is associated with a filtration $\mathcal{F}_n$ such that X_n is $\mathcal{F}_n$-measurable $\forall\ n \in \mathbb{N}$ (or X_t is $\mathcal{F}_t$-measurable $\forall t \in \mathbb{R}^+$ if continuous).
The process X is said to be *adapted* to the filtration $\mathcal{F}_n$ or $\mathcal{F}_t$.

Definition 1.11 - *State Space*
The *State Space*, S, of a stochastic process is the set of all values that a quantity can take at a specific time.

Remark 1.5 - *Sample Space of Stochastic Processes*
For a *discrete-time stochastic process* Ω is typically taken to be $S^\mathbb{N} = \{(x_0, x_1, \dots) : x_i \in S\}$.
For a *continuous-time stochastic process* Ω is typically taken to be the space of all functions $f : [0\infty) \to \mathbb{R}$ and/or right continuous functions with left limits.

Example 1.3 - *Discrete-Time Stochastic Process*
Consider a guy who tosses a coin infinitely main times. Let the sequence of random variables $\{X_n\}_{n\in\mathbb{N}}$ encode the outcomes by mapping heads to 0 and tail to 1.

$$\begin{array}{rcl} S & = & \{0, 1\} \\ \Omega & = & \{H, T\} \\ \Delta & = & \mathbb{N} \end{array}$$

Example 1.4 - *Continuous-Time Stochastic Process*
Packets arrive at a router and need to be stored until they can be pass on. The router has a finite capacity buffer (with capacity C), and packets arrive and depart in continuous time.

$$\begin{array}{rcl} S & = & \{0,\dots,C\} \text{ Discrete} \\ \Omega & = & \text{All right} - \text{continuous paths taking values in } [0,C] \\ \Delta & = & \mathbb{R}^+ \end{array}$$

1.4 Markov Property

Definition 1.12 - *Markov Property*
A *Stochastic Process* has the *Markov Property* if future values only depend upon the present value, and no previous values.
i.e. $X_{n+1}=f(X_n)$.

Definition 1.13 - *Markov Chain*
A *Markov Chain* in discrete time is a discrete state space process with the *Markov Property.*
Formally.
Let $X=\{X_n\}_{n\in\mathbb{N}}$ be a discrete time, discrete space stochastic process & $\mathcal{F}_n=\sigma(X_0,\dots,X_n)$ be the filtration generate by the process.
X is a *Markov Chain* if for each fixed n and each $i_0,\dots,i_{n+1}\in S$ the following holds

$$\begin{array}{lrcl} & \mathbb{P}(X_{n+1}=i_{n+1}|X_n=i_n,\dots,X_0=i_0) & = & \mathbb{P}(X_{n+1}=i_{n+1}|X_n=i_n) \\ \text{Equivalently} & \mathbb{P}(X_{n+1}=i_{n+1}|\mathcal{F}_n) & = & \mathbb{P}(X_{n+1}=i_{n+1}|Xn) \end{array}$$

Definition 1.14 - *Time-Homogeneous Markov Chain*
A *Markov Chain* $X=\{X_n\}_{n\in\mathbb{N}}$ is *Time-Homogeneous* if

$$\forall\ i,j\in S, \mathbb{P}(X_{m_1+1}=j|X_{m_1}=i)=\mathbb{P}(X_{m_2+1}=j|X_{m_2}=i)\ \forall\ m_1,m_2\in[0,n-1]$$

Definition 1.15 - *Markov Process*
A *Markov Process* is a continuous-time stochastic process X with filtration $\mathcal{F}_t$ where $\forall\ 0\leq s<t$ and $A\subset S$

$$\mathbb{P}(X_t\in A|\mathcal{F}_s)=\mathbb{P}(X_t\in A|X_s)$$

Example 1.5 - *Not Markov Process*
Consider a particle moving on a line, that is constantly bombarded by other particles that change its velocity.
Let X_n be its position & U_n be its velocity at time $n\in\mathbb{N}$.
We simplify its motion to

$$\begin{array}{rcl} X_{n+1} & = & X_n+U_n \\ U_{n+1} & = & U_n+\eta_n \end{array}$$

where $\eta_n=Bin(2,\frac{1}{2})-1$.
Consider $\mathbb{P}(X_{n+1}=x|X_n=x,X_{n-1}=x-1)$.
Then $U_{n-1}=1$ & $U_n=0\implies\eta_n=-1$.
Since $\mathbb{P}(\eta_n=-1)=\frac{1}{4}\implies\mathbb{P}(X_{n+1}=x|X_n=x,X_{n-1}=x-1)=\frac{1}{4}$.
Now consider $\mathbb{P}(X_{n+1}=x|X_n=x,X_{n-1}=x)$.
Then $U_{n-1}=0$ & $U_n=0\implies\nu_n=0$.
Since $\mathbb{P}(\eta_n=0)=\frac{1}{2}\implies\mathbb{P}(X_{n+1}=x|X_n=x,X_{n-1}=x)=\frac{1}{2}$.
Hence $\{X_n\}$ is not a Markov Process.

Example 1.6 - *Fixed Time v Fixed Realisation*
Consider tossing a coin 100 times.

Let $|_n(\omega)$ encode the outcome of the n^{th} toss with 0 for heads. & 1 for tails.
Let $X_0 = 0$ and for $x \in \mathbb{N}^{\leq 100}$ $X_n(\omega) = \sum_{i=1}^{n} |_n(\omega)$. Then $\{|_n\}_{n=1,\dots,100}$ and $\{X_n\}_{n=1,\dots,100}$ are stochastic processes.
Take $\Omega = \{(\omega_1, \dots, \omega_{100} : \omega_i \in \{H, T\},\ i \in [1, 100]\}$.
There are two views of the stochastic process X

i) *With Fixed n*
We have a random variables $X_n(\cdot)$ that depends on ω.
If the coin is fair then X_n $Bin(n, \frac{1}{2})$.

ii) *With Fixed ω*
We have a function $X.(\omega)$, which is deterministic, called a *sample path* or *realisation of the process X*.

In the case of continuous-time process

i) *With Fixed t.*
For each t, $X_t(\cdot)$ is a random variable. X_s & $X)t$ are usually not independent for $s \neq t$.
For a finite collection of times $\{t_1, t_2, \dots, t_n\}$, the joint distribution of the random vector $(X_{t_1}, X_{t_2}, \dots, X_{t_m})$ is called a *finite-dimensional distribution.*
The collection of all fdd's contain all information about the process X.

ii) *With fixed ω.*
Each $X.(\omega)$ is a function that maps $[0, \infty) \to \mathbb{R}$.
For any fixed $\omega \in \Omega$ there is a corresponding path $\{X_t(\omega) : t \geq 0\}$. This is called a *sample path* or *realisation* of X at ω.

1.5 Increasing & Decreasing Sequences of Events

Definition 1.16 - *Increasing Sequence*
Let $A_1, A_2, \dots$ be a sequence of events.
This sequence is said to be *increasing* if $A_n \subseteq A_{n+1}\ \forall\ n \in \mathbb{N}$.

Proposition 1.1 - *Union of Increasing Sequence*
Let $A = \bigcup\limits_{n=1}^{\infty} A_n$ be the event *at least one of the A_n occurred.*
Then $A_n = \bigcup i = 1^n A_i$ and we can think if A as a limit of the A_i.

Example 1.7 - *Increasing Sequence*
Let $A_n = \{\text{From } n^{th} \text{ toss onwards, all tosses yield heads}\}$. Then

$$\begin{aligned}
\forall\ \omega \in A_n \quad & \Leftrightarrow \quad \text{toss } n, n+1, \dots \text{ are heads.} \\
& \implies \quad \text{toss } n+1, n+2, \dots \text{ are heads} = A_{n+1} \\
& \Leftrightarrow \quad \omega \in A_{n+1}
\end{aligned}$$

Hence $A_n \subset A_{n+1}$.
Let $A = \bigcup\limits_{n=1}^{\infty} A_n$ be the event that at least one of A_n occurs.
Then *exists* N *st* $\forall\ n \geq N$ A_n occurs.

Theorem 1.2 - *Continuity of Probability, Increasing*
Suppose $A_1, A_2, \dots$ is an increasing sequence of events and let $A = \bigcup\limits_{n=1}^{\infty} A_n$. Then

$$\mathbb{P}(A) = \lim_{n \to \infty} \mathbb{P}(A_n)$$

Proof 1.2 - *Continuity of Probability, Increasing*
Let $D_1 = A_1$ & $D_n = A_n \backslash A_{n-1}$ for $n \geq 2$.
Then $D_i \subset D_j = \emptyset \; \forall i \neq j$ and $A_n = \bigcup_{i=1}^{n} D_i$.
Thus $A = \bigcup_{i=1}^{\infty} D_i$.
By σ-additivity of probability measures we have

$$\mathbb{P}(A_n) = \sum_{i=1}^{n} \mathbb{P}(D_i) \; \& \; \mathbb{P}(A) = \sum_{i=1}^{\infty} \mathbb{P}(D_i) = \lim_{n\to\infty} \mathbb{P}(A_n)$$

The result follows by the definition of infinite sums.

Definition 1.17 - *Decreasing Sequence*
Let $B_1, B_2, \dots$ be a sequence of events.
This sequence is said to be *increasing* if $B_{n+1} \subseteq B_n \; \forall \; n \in \mathbb{N}$.

Theorem 1.3 - *Continuity of Probability, Decreasing*
Suppose $B_1, B_2, \dots$ is a decreasing sequence of events and let $B = \bigcap_{n=1}^{\infty} B_n$. Then

$$\mathbb{P}(B) = \lim_{n\to\infty} \mathbb{P}(B_n)$$

Proof 1.3 - *Continuity of Probability, Decreasing*
Let $A_n = B_n^c$ be the complement of B_n.
Then the A_n's are an increasing sequence. Also

$$A := \bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} = \left(\bigcap_{n=1}^{\infty} B_n \right)^c = B^c$$

By the previous theorem

$$\mathbb{P}(B^c) = \mathbb{P})A_= \lim)n \to \infty \mathbb{P}(A_n) = \lim_{n\to\infty} \mathbb{P}(B_n^c)$$

Since $\mathbb{P}(B^c) = 1 - \mathbb{P}(B)$ & $\mathbb{P}(B_n^c) = 1 - \mathbb{P}(B_n)$, then

$$1 - \mathbb{P}(B) = 1 - \lim_{n\to\infty} \mathbb{P}(B_n)$$

Then the result follows.

Example 1.8 - *Continuity of Probability*
Initially in an urn there is one white and one red ball. A ball is chosen at random and returned to the urn alongside an extra red ball. Thus when the n^{th} ball is chosen there are n red balls and 1 white ball. Hence

$$\mathbb{P}(n^{th} \text{ ball chosen is red}) = \frac{n}{n+1}$$

What is the probability that a white ball is never chosen?
Let $B_n\{$The first n balls are all red $\}$
This is a decreasing sequence $\implies B_{n+1} \subseteq B_n$.
Then $B = \bigcap_{i=1}^{\infty} B_i = \{$All balls chosen are red$\}$.

$$\begin{array}{rcl} \mathbb{P}(B) & = & \lim_{n\to\infty} \mathbb{P}(B_n) \\ & = & \lim_{n\to\infty} \frac{1}{2} \times \frac{2}{3} \times \cdots \times \frac{n}{n+1} \\ & = & \lim_{n\to\infty} \frac{1}{n+1} \\ & = & 0 \end{array}$$

Example 1.9 - *Continuity of Probability*
Let $X : \Omega \to \mathbb{R}$ be a continuous random variable and F_X be its cumulative density function.

i) *Show* $\lim_{n\to\infty} F_X(x + \frac{1}{n}) = F_X(x) \ \forall \ x \in \mathbb{R}$.
We have $F_X(x) = \mathbb{P}(X \leq x)$ & $F_X(x + \frac{1}{n}) = \mathbb{P}(X \leq x + \frac{1}{n})$.
Let $B_n = \{X \leq x + \frac{1}{n}\}$ a decreasing sequence.
Let $B = \bigcap\limits_{n=1}^{\infty} B_n = \{X \leq x\}$.
Hence $F_X(x) = \mathbb{P}(B) = \lim_{n\to\infty} \mathbb{P}(B_n) = \lim_{n\to\infty} F_X(x + \frac{1}{n})$.

ii) *Show* $\lim_{n\to\infty} F_X(-n) = 0$.
Let $B_n = \{X \leq -n\}$ a decreasing sequence.
Then $\lim_{n\to\infty} F_X(-n_= \lim_{n\to\infty} \mathbb{P}(B_n) = \mathbb{P}(\bigcap_{n=1}^{\infty} B_n) = \mathbb{P}(\emptyset) = 0$.

iii) *Show* $\lim_{n\to\infty} F_X(n) = 1$.
Let $A_n = \{X \leq n\}$ an increasing sequence.
Then $\lim_{n\to\infty} F_X(n) = \lim_{n\to\infty}(A_n) = \mathbb{P}(\bigcup_{n=1}^{\infty} A_n) = \mathbb{P}(\Omega) = 1$.

2 Random Walks

Definition 2.1 - *Random Walk*
A *Random Walk* is a process which at each discrete time step the value either increases or decrease by 1, only.

2.1 Absorbing Barriers

Definition 2.2 - *Absorbing Barriers*
Absorbing Barriers are values which if a random walk reaches it never leaves.

Theorem 2.1 - *One-Step Conditioning Argument*
Let X, Y, A be events where A is dependent of X, Y. Then

$$\mathbb{P}(A) = \mathbb{P}(A|X)\mathbb{P}(X) + \mathbb{P}(A|Y)\mathbb{P}(Y)$$

Example 2.1 - *Gambler's Ruin*
A gambler has £k. Her opponent has £(N-k).
Each time a game is player a £is placed. The gambler wins with probability p & her opponent with probability $q = 1 - p$.
Successive players of the game are independent. The game ends when one player has no money left.
What is the probability the gambler is ruined?

Let X_n be the gambler's capital in sterling after n bets.
There are *absorbing barriers* at 0 and N, the gambler is ruined if $X_n = 0$.
The process $X = \{X_n\}_{n\in\mathbb{N}_0}$ is a *Markov Chain* with the following transitions

i) Interior Points, for $k \in [1, N-1]$.
   (a) $p_{k,k+1} = \mathbb{P}(X_{n+1} = k+1 | X_n = k) = p$.
   (b) $p_{k,k-1} = \mathbb{P}(X_{n+1} = k-1 | X_n = k) = q = 1 - p$.

ii) Boundary points, for all other values of k.

(a) $p_{0,0} = \mathbb{P}(X_{n+1} = 0 | X_n = 0) = 1$.

(b) $p_{N,N} = \mathbb{P}(X_{n+1} = N | X_n = N) = 1$.

By the *one-step conditioning argument* we see can derive that $p_k = p_{k+1}p + p_{k-1}q$ for $k \in [1, N-1]$.
We have boundary conditions $p_0 = 1$ & $p_N = 0$.
We now solve these as difference equations.
Let $p_k = \theta^k$ for $\theta \in \mathbb{R}$.
Then $\theta^k = \theta^{k=1}p + \theta^{k=1}q$.
Set $k = 1 \implies \theta = \theta^2 p + q \implies 0 = p\theta^2 - \theta + q = (p\theta - q)(\theta - 1)$.

If $p \neq q$ then there are two distinct solutions $\theta = \frac{p}{q}$ & $\theta = 1$.
Hence the general solution is $p_k = A(\frac{q}{p})^k + B(1)^k = A(\frac{q}{p})^k + B$ for $k \in [0, n]$.
Plugging in the boundary conditions we get

$$\begin{array}{lll} & p_0 = 1, p_0 = A + B & p_N = 0, p_N = A(\frac{q}{p})^n + B \\ \implies & A = \dfrac{1}{1 - (\frac{q}{p})^N} & B = \dfrac{-(\frac{q}{p})^N}{1 - (\frac{q}{p})^N} \end{array}$$

Hence

$$p_k = \frac{(\frac{q}{p})^k - (\frac{q}{p})^N}{1 - (\frac{q}{p})^N}$$

If $p = q = \frac{1}{2}$ then the only solution to the equation $p\theta^2 - \theta + q = 0$ is $\theta = 1$.
In this case we try $p_k = (A + Bk)\theta^k = A + Bk$.
Plugging in boundary conditions we get

$$p_0 = 1 \text{ \& } p_0 = A \implies A = 1 p_n = 0 \text{ \& } p_1 + NB \implies B = \frac{1}{N}$$

Hence the probability of ruin is $p_k = 1 - \frac{k}{N}$.

Theorem 2.2 - *Probability of Ruin with Absorbing Barrier at 0*
Let $k \geq 1$ be fixed.
Let p_k be the probability of ruin in the random walk with absorbing barrier at 0, and $p_k^{(N)}$ be the probability of ruin in the gambler's ruin problem with upper barrier at N, in both cases starting at $X_0 = k$. Then

$$\lim_{N \to \infty} p_k^{(N)} = p_k$$

Proof 2.1 - *Probability of Ruin with Absorbing Barrier at 0*
Let $A_n = \{\omega \in \Omega : \exists n \geq 1 \ st \ X_n(\omega) = 0; X_m(\omega) \leq N_1 \ \forall \ m \in [0, n-1]\}$.
This is the event where X gets absorbed at 0 and never reaches n.
Then $\mathbb{P}(A_n) = p_K^{(N)}$. Now

$$\begin{array}{rcl} \omega \in A_n & \iff & \exists \ n \geq 1 \ st \ X_n(\omega) = 0 \ \& \ X_m(\omega) \leq N - 1 \ \forall \ m \in [0, n-1] \\ & \implies & \exists \ n \geq 1 \ st \ X_n(\omega) = 0 \ \& \ X_m(\omega) \leq N \forall \ m \in [0, n-1] \end{array}$$

. So $A_n \subset A_{n+1}$, an increasing sequence of events.
Take $A = \bigcup_{N=1}^{\infty} A_N$ then $A = \{\omega \in \Omega : X_n(\omega) = 0; n \geq 1\}$.
By the continuity of probability

$$p_k = \mathbb{P}(A) = \lim_{N \to \infty} \mathbb{P}(A_N) = \lim_{N \to \infty} p_k^{(N)}$$

Theorem 2.3 -
For the random walk with an absorbing barrier at 0 but no upper barrier

$$\mathbb{P}(ruin|X_0 = k) = \mathbb{P}(random\ walk\ hits\ eventually|X_0 = k) = \begin{cases} (\frac{q}{p})^k & if\ q < p \\ 1 & if\ q \geq p \end{cases}$$

Proof 2.2 -
Here k is fixed.
In the cases $p \neq q$ we have

$$p_k^{(N)} = \frac{(\frac{q}{p})^k - (\frac{q}{p})^N}{1 - (\frac{q}{p})^N} \xrightarrow{N \to \infty} \begin{cases} 1 & q > p \\ 0 & q < p \end{cases}$$

For $p = q = \frac{1}{2}$ we have

$$p_k^{(N)} = 1 - \frac{k}{N} \xrightarrow{N \to \infty} 1$$

Proposition 2.1 - *No Absorbing Barriers*
Suppose there are no absorbing barriers. We get that the probability of a process reaching 0 is given as

$$\mathbb{P}(unrestricty\ random\ walk\ hits\ 0\ eventually|X_0 = k) = \mathbb{P}(random\ walk\ with\ absorption\ at\ 0\ gets\ absorbed|X_0$$

The solution to this can be seen in the previous theorem.
We get that if $q < p$ then there is a positive probability of $1 - (\frac{p}{q})^k$ that a random walk starting at k will never reach 0.
If $p = q$ then the random walk will always, eventually reach 0.

2.2 Transience and Recurrence

Notation 2.1 -
Consider a general time-homogeneous Markov chain $X = \{X_n\}_{n \in \mathbb{N}}$ starting in state $X_0 = i \in S$.
We denote the following questions as follows

i) Will X ever return to i? f_{ii}.

ii) Will X ever visit a given state j? f_{ij}.

iii) If so, how long will it take? m_{ij}.

iv) And how often will it happen?

Theorem 2.4 - *n-step Transition Probability*
The probability of transitioning from initial state i to state j in $n \in \mathbb{N}$ steps is given by

$$p_{ij}(n) = \mathbb{P}(X_n = j|X_0 = i)$$

We also define

$$p_{ij}(0) = \begin{cases} 1 & if\ i = j \\ 0 & otherwise \end{cases}$$

Theorem 2.5 - *Probability of Transition*
For $n \geq 1$

$$p_{ij}(n) = \sum_{m=1}^{n} p_{jj}(n-m) f_{ij}(m)$$

Proof 2.3 - *Probability of Transition*
Let $A=\{X_n=j\}$ & B_m={First visit to state j at step m}=$\{X_1\neq j,\dots,X_{m-1}\neq j,X_m=j\}$.
Then $B_1,B_2,\dots$ are pairwise disjoint & $A\subset(B_1\bigcup\dots\bigcup B_n)$.
Then $A=A\bigcap(B_1\bigcup\dots\bigcup B_n)=(A\bigcap B_1)\bigcup\dots\bigcup(A\bigcap B_n)$.
Hence

$$\begin{aligned}p_{ij}(n) &= \mathbb{P}(X_n=j|X_0=i)\\ &= \mathbb{P}(A|X_0=i)\\ &= \textstyle\sum_{m=1}^n \mathbb{P}(A\bigcap B_m|X_0=i)\\ &= \textstyle\sum_{m=1}^n \mathbb{P}(A|B_m,X_0=i)\mathbb{P}(B_m|X_0=i)\\ &=\end{aligned}$$

Definition 2.3 - *Transient & Recurrent*
A state $j\in S$ is *transient* if $f_{jj}<1$.
A state $j\in S$ is *recurrent* if $f_{jj}=1$. This means the chain will definitely return to its origin in the future.

Proposition 2.2 - *First Passage Probabilities, f_{ij}*
We define T_{ij} be the time at which a chain reaches j for the first time, after starting at i.
T_{ij} is a random variable with for $n\geq1$

$$\begin{aligned}\mathbb{P}(T_{ij}=n) &= \mathbb{P}(First\ visit\ to\ j\ is\ after\ n\ steps|X_0=i)\\ f_{ij}(n) &= \mathbb{P}(First\ visit\ to\ j\ is\ after\ n\ steps|X_0=i)\\ &= \mathbb{P}(X_1\neq j,\dots X_{n-1}\neq j,X_n=j|X_0=i)\\ &\in [0,1]\\ f_{ij} &= \sum_{n=1}^{\infty} f_{ij}(n)\\ &= \mathbb{P}(X\ ever\ visits\ j|X_0=i)\\ &\in [0,1]\end{aligned}$$

Definition 2.4 - *Expected First Passage*
We define m_{ij} be the expected time for first passage from i to j.

$$m_{ij}=\mathbb{E}(\text{Time of first return of } i|X_0=i)=\mathbb{E}(T_{ij})=\sum_{n=1}^{\infty}nf_{ij}(n)$$

N.B. If $f_ij<1$ then $m_{ij}=\infty$.

Theorem 2.6 - *Number of Visits*

$$\begin{aligned}\sum_{n=1}^{\infty}p_{ij}(n) &= \sum_{n=0}^{\infty}\mathbb{P}(X_n=j|X_0=i)\\ &= \sum_{n=0}^{\infty}\mathbb{E}(1_{\{X_n=j\}}|X_0=i)\\ &= \mathbb{E}\left(\sum_{n=0}^{\infty}1_{\{X_n=j\}}|X_0=i\right)\\ &= \mathbb{E}(Number\ Visits\ to\ j|X_0=i)\end{aligned}$$

Proposition 2.3 - *Generating Functions*
We can define the following generating functions for *first passage probabilities* & *n-step transition probabilities*

$$P_{ij}(s)=\sum_{n=0}^{\infty}p_{ij}(n)s^n \quad F_{ij}(s)=\sum_{n=1}^{\infty}f_{ij}(n)s^n$$

with the conventions $p_{ij}=1_{i=j}$ and $f_{ij}(0)=0\ \forall i,j$.

Remark 2.1 - *Generating Functions*
The generating functions defined in **Proposition 2.3** are well-defined for $|s|<1$.
If we take $s=1$ then

$$F_{ij}(1)=\sum_{n=1}^{\infty}f_{ij}(n)=f_{ij}$$

Now consider

$$\begin{array}{rcl} F'_{ij}(s) & = & \frac{d}{ds}\sum\limits_{n=1}^{\infty}f_{ij}(n)s^n \\ & = & \sum\limits_{n=1}^{\infty}\frac{d}{df}(f_{ij}(n)s^n) \\ & = & \sum\limits_{n=1}^{\infty}f_{ij}(n)ns^{n-1} \\ F'_{ij}(1) & = & \sum\limits_{n=1}^{\infty}f_{ij}(n)n \\ & = & m_{ij} \\ & = & \mathbb{E}(T_{ij}) \end{array}$$

Similarly

$$P_{ij}(1)=\sum_{n=1}^{\infty}p_{ij}(n)=\mathbb{E}(Number\ Visits\ to\ j|X_0=i)$$

Theorem 2.7 -
For $n\geq 1$

$$p_{ij}(n)=\sum_{m=1}^{n}p_{jj}(n-m)f_{ij}(m)$$

Proof 2.4 -
Let $A=\{X_n=j\}$ the events that at step n we are at state j
And $B_m=\{\text{First visit to state } j \text{ at step } m\}=X_1\neq j,\dots,X_{m-1}\neq j,X_m=j\}$.
Then $B_1,\dots$ are pairwise disjoint and $A\subset(B_1\bigcup\cdots\bigcup B_n)$.
So $A=(A\bigcap B_1)\bigcup\cdots\bigcup(A\bigcap B_n)=A\bigcap(B_1\bigcup\cdots\bigcup B_n)$.
Hence

$$\begin{array}{rcl} p_{ij}(n) & = & \mathbb{P}(X_n=j|X_0=i) \\ & = & \mathbb{P}(A|X_0=i) \\ & = & \sum_{m=1}^{n}\mathbb{P}(A\bigcap B_m|X_0=i) \\ & = & \sum_{m=1}^{n}\mathbb{P}(A|B_m,X_0=i)\mathbb{P}(B_m|X_0=i) \\ \text{By Markov Property} & = & \sum_{m=1}^{n}\mathbb{P}(A|X_m=j)f_{ij}=m \\ & = & \sum_{m=1}^{n}\mathbb{P}(X_n=j|X_m=j)f_{ij}(m) \\ \text{By time homegenity} & = & \sum_{m=1}^{n}\mathbb{P}(X_{n-m}=j|X_0=j)f_{ij}(m) \\ & = & \sum_{m=1}^{n}p_{jj}(n-m)f_{ij}(m) \end{array}$$

Theorem 2.8 -

$$P_{ij}(s)=\mathbf{1}_{i=j}+F_{ij}(s)P_{jj}(s)$$

Proof 2.5 -
By definition

$$\begin{array}{rcl} P_{ij}(s) & = & \sum_{n=0}^{\infty}p_{ij}(n)s^n \\ & = & p_{ij}(0)+\sum_{n=1}^{\infty}p_{ij}(n)s^n \\ & = & p_{ij}(0)+\sum_{n=1}^{\infty}\sum_{m=1}^{n}p_{ij}(n-m)f_{ij}(m)s^n \\ & = & \mathbf{1}_{i=j}+\sum_{n=1}^{\infty}\sum_{m=1}^{n}p_{ij}(n-m)f_{ij}(m)s^n \\ & = & \mathbf{1}_{i=j}+\sum_{n=1}^{\infty}\sum_{m=m}^{\infty}p_{ij}(n-m)f_{ij}(m)s^n \\ & = & \mathbf{1}_{i=j}+\sum_{m=1}^{\infty}f_{ij}(m)s^m\sum_{n'=0}^{\infty}p_{jj}(n')s^{n'} \\ & = & \mathbf{1}_{i=j}+F_{ij}(s)P_{jj}(s) \end{array}$$

Theorem 2.9 -
For arbitrary state $i, j \in S$ j is recurrent iff $P_{ij}(1) = \sum_{n=0}^{\infty} p_{ij}(n) = \infty$.
N.B. $\sum_{n=1}^{\infty} p_{ij}(n)$ is the expected number of visits to j if the chain starts at i.

Proof 2.6 -
Recall

- j is recurrent iff $f_{jj} = 1$.
- $f_{jj} = \sum_{n=1}^{\infty} f_{jj}(n) = F_{jj}(1)$.

Proof

i) Suppose $i = j$.
By **Theorem 2.7** $F_{jj}(s) = \dfrac{P_{jj}(s) - 1}{P_{jj}(s)}$.
So $F_{jj}(1) = 1 \implies P_{jj}(1) = \infty$.
This meets the requirement and hence the result holds for $i = j$.

ii) Suppose $i \neq j$.
Since $F_{ij}(1) = f_{ij} = \mathbb{P}(\text{The random variable ever visits } j | X_0 = i) > 0$ for random walks.
And $P_{ij}(1) = F_{ij}(1) P_{jj}(1)$.
We conclude $P_{ij} = 0$ iff $P_{jj} = \infty$.

Theorem 2.10 -
If j is transient then $p_{ij}(n) \leftarrow 0$ as $n \leftarrow \infty\ \forall\ i$.

Proof 2.7 -
Since j is transient then $\sum_{n=0}^{\infty} p_{ij}(n) < \infty$ be **Theorem 2.8** .
Hence $p_{ij}(n) \leftarrow 0$ as $n \leftarrow \infty$.

2.3 Applications of Random Walks

Proposition 2.4 - *Spatial Homogeneity*
The *Spatial Homogeneity* of a random walk means that whatever we say about the recurrence and return times for state 0 also holds for all any state i.

Theorem 2.11 - P_{00}
Note that if n is odd then $X_n \neq 0$ since an even number of movements is required to return to the origin.
Let $n = 2m$, if $X_m = 0$ then there were exactly m upward movements & m downward movements.
The number of upwards movements is modelled by $Binomial(2m, p)$, so

$$\begin{array}{rcl} p_{00}(2m) & = & \mathbb{P}(X_{2m} = 0 | X_0 = 0) \\ & = & \mathbb{P}(m\text{upwards stepsin a total of } 2m \text{ steps} \\ & = & p^m q^m \binom{2m}{m} \\ P_{00}(s) & = & \sum_{i=0}^{\infty} p_{00}(i) s^i \\ & = & \sum_{i=0}^{\infty} p_{00}(2i) s^{2i} \\ & = & \sum_{i=0} 6\infty \binom{2i}{i} p^i q^i s^{2i} \\ & = & \sum_{i=0} 6\infty \binom{2i}{i} \left(\frac{4pqs^2}{4}\right) \\ & = & (1 - 4pqs^2)^{-1/2} \end{array}$$

N.B. See Page 25 of Booklet 2 for the identity used at the end here.

Theorem 2.12 -
Consider an unrestricted random walk starting at 0.

i) The probability that the walk returns to 0 eventually is $1 - |p - q|$.

ii) If $p = q = \frac{1}{2}$ then return is certain, but the expected time till first return is ∞

Proof 2.8 -

i) The probability of eventual return is $f_{00} = F_{00}(1)$.
Theorem 2.7 implies

$$\begin{array}{rrcl} & P_{00}(s) & = & 1 + F_{00}(s) + P_{00}(s) \\ \implies & F_{00}(s) & = & 1 - \frac{1}{P_{00}(s)} \\ & & = & 1 - \sqrt{1 - 4pqs^2} \\ \implies & F_{00}(1) & = & 1 - \sqrt{1 - 4pq} \\ & & = & 10\sqrt{(p+q)^2 - 4pq} \text{ Since } p + q = 1 \\ & & = & 1 - \sqrt{p^2 - 2pq + q^2} \\ & & = & 1 - \sqrt{(p-q)^2} \\ & & = & 1 - |p - q| \end{array}$$

ii) If $p = q = \frac{1}{2} \implies f_{00} = F_{00}(1) = 1$.
recall T_{00} us the time of first return to 0 if the walk starts at 0.
Then T_{00} is almost surely finite.

$$\begin{array}{rcl} m_{00} & = & \mathbb{E}(T_{00}) \\ & = & \sum_{n=1}^{\infty} n f_{00}(n) \\ & = & \lim_{s\uparrow 1} \sum_{n=1}^{\infty} n s^{n-1} f_{00}(n) \\ & = & \lim_{s\uparrow 1} F'_{00}(s) \end{array}$$

Since $p = \frac{1}{2} = q$ we have $F_{00} = 1 - \sqrt{1 - s^2}$
$\implies F'_{00}(s) = \dfrac{s}{\sqrt{1 - s^2}}$
$\implies \lim_{s\uparrow 1} \dfrac{s}{\sqrt{1 - s^2}} = \infty$

Definition 2.5 - *Null Recurrent*
A recurrent state i is called *Null Recurrent* if $m_{ii} = \infty$.

By the previous theorem all states in a simple random walk are null recurrent.

Definition 2.6 - *Positive Recurrent*
A recurrent state i is called *Positive Recurrent* if $m_{ii} < \infty$.

2.4 Stopping Time & Wald's Lemma

Definition 2.7 -
Let $X = \{X_n\}_{n\in\mathbb{N}}$ be a stochastic process & T be a non-negative inter-valued random variable. T is said to be a *Stopping Time* for X if $\forall\ n$ the event $\{T \leq n\}$ is completely determined by the values of $X_0, X_1, \dots, X_n$.

Example 2.2 - *Stopping Time*
Consider a simple unconstrained random walk starting at 0.
Let $T = min\{n : X_n = 5\}$ (i.e. The first time $X_n = 5$.
By looking at all the values of X_M for $m \leq n$ you can see if X_m has been equal to 5 and hence if $T \leq n$.
T is a stopping time.

Example 2.3 - *Not Stopping Time*
Consider the sample simple unconstrained random walk starting at 0.
Let $T = max\{n : X_n = 5\}$ then we cannot know whether $T \leq n$ without knowing all future values of X as well.
T is not a stopping time.

Theorem 2.13 - *Wald's Lemma*
Let $Z_1, Z_2, \dots$ be a sequence of independent, identically distributed random variables with $\mathbb{E}(|Z_n|) < \infty$ and $X_n = \sum_{m=1}^{n} Z_m$.
Let T be a stopping time for the process $X = \{X_n\}_{n\in\mathbb{N}}$ with $\mathbb{E}(T) < \infty$. Then

$$\mathbb{E}(X_T) = \mathbb{E}(Z_1)\mathbb{E}(T)$$

N.B. The fact that $\{T \leq\}$ depends only on $X_0, X_1, \dots, X_n$ is equivalent to $\{T \leq n\}$ depending only on $Z_1, \dots, Z_n$ so we could say instead that T is a stopping time for $\{Z_n\}$.

Proof 2.9 - *Wald's Lemma*
Since $X_T = \sum_{m=1}^{T} Z_m = \sum_{m=1}^{\infty} Z_m \mathbf{1}_{m\leq T}$ then

$$\begin{array}{lrcl} & \mathbb{E}(X_T) & = & \mathbb{E}\left(\sum_{m=1}^{\infty} Z_m \mathbf{1}_{m\leq T}\right) \\ & & = & \sum_{m=1}^{\infty} \mathbb{E}(Z_m \mathbf{1}_{m\leq T}) \\ \text{By Smooting Property} & & = & \sum_{m=1}^{\infty} \mathbb{E}(\mathbb{E}(Z_m \mathbf{1}_{m\leq T}|\mathcal{F}_{m-1})) \end{array}$$

Notice that $\{M \leq T\}^C = \{T \geq m\}^C = \{T \leq m-1\}$ is $\mathcal{F}_{m-1}$-measurable. Hence by *'take out what is know'*

$$\begin{array}{rcl} \mathbb{E}(X_T) & = & \sum_{m=1}^{\infty} \mathbb{E}(\mathbf{1}_{m\leq T}\mathbb{E}(|_m|\mathcal{F}_{m-1})) \\ & = & \sum_{m=1}^{\infty} \mathbb{E}(\mathbf{1}_{m\leq T}\mathbb{E}(Z_1)) \\ & = & \mathbb{E}(Z_1)\sum_{m=1}^{\infty} \mathbb{E}(\mathbf{1}_{m\leq T}) \\ & = & \mathbb{E}(Z_1)\sum_{m=1}^{\infty} \mathbb{P}(T \geq m) \\ & = & \mathbb{E}(Z_1)\mathbb{E}(T) \end{array}$$

Example 2.4 - *Simple Random Walk*
Let $\mathbb{P}(Z_n = 1) = 2/3$ and $\mathbb{P}(Z_n = -1) = 1/3$, so that we have a simple random walk.
Assume $X_0 = 0$, and let $T = min\{n : X_n = 5\}$.
It can be shown that $\mathbb{E}(T) < \infty$ and also it is clear that $\mathbb{E}(|_n) = 1/3$.
Wald's Lemma tells us that

$$\mathbb{E}(X_T) = \mathbb{E}(Z_1)\mathbb{E}(T) = \frac{1}{3}\mathbb{E}(T)$$

But we know that $X_T = 5$, by the definition of T. So we see that

$$\mathbb{E}(T) = 15$$

Remark 2.2 - *Alternative Method of Establishing Transience/Recurrence of Random Walks*
We look at random walks in $\mathbb{Z}^d$ with $d = 1, 2$.

In the case when $d=1$ we have $p_{00}(2m)=\begin{pmatrix}2m\\m\end{pmatrix}p^mq^m=\frac{(2m)!}{m!m!}p^mq^m$.
By *Stirling's Formula*

$$\begin{array}{rrcl}
 & (2m)! & \sim & \sqrt{4\pi m}\left(\frac{2m}{e}\right)^{2m} \text{ as } m\to\infty \\
 & (m!)^2 & \sim & \sqrt{2\pi m}\left(\frac{m}{e}\right)^{2m} \text{ as } m\to\infty \\
\implies & p_{00}(2m) & \sim & \frac{1}{\sqrt{\pi m}}2^{2m}p^mq^m \\
 & & = & \frac{1}{\sqrt{\pi m}}(4pq)^m \\
\text{If} & p & = & q=\frac{1}{2} \\
\implies & p_{00}(2m) & \sim & \frac{1}{\sqrt{\pi m}} \text{ as} m\to\infty \\
\text{Then} & \sum_{m=M}^\infty p_{00}(2m) & > & \sum_{m=M}^\infty 0.99\frac{1}{\sqrt{\pi m}} \\
 & & = & \frac{0.99}{\sqrt{\pi}}\sum_{m=M}^\infty\frac{1}{\sqrt{m}} \\
 & & = & \infty \\
 & \textit{State } 0 \text{ is recurrent.} & & \\
\text{If} & p & \neq & q \\
\implies & 4pq & < & 1 \\
\implies & p_{00}(2m) & \sim & \frac{(4pq)^m}{\sqrt{\pi m}} \\
\implies & \sum_{m=M}^\infty p_{00}(2m) & < & \sum_{m=M}^\infty 1.01\frac{(4pq)^m}{\sqrt{\pi m}} \\
 & & < & \sum_{m=M}^\infty(4pq)^m \\
\text{Since } r<1 & & < & \infty \\
 & \textit{State } 0 \text{ is transcient.} & &
\end{array}$$

In the case when $d=2$ we have $p_{ij}=\begin{cases}\frac{1}{4} & \text{if } |i-j|=1\\ 0 & \text{otherwise}\end{cases}$.
Let $X_n=(X_n^{(1)},X_n^{(2)})$ where each coordinate is a random walk.
These random walks are not simple since steps of size 0 are allowed.
Let X_n^+ be the projection of X_n onto $y=x$ & X_m^- be the projection of X_n onto $y=-x$.
Then X_n^+ & X_n^- are simple symmetric random walks on $\frac{\mathbb{Z}}{\sqrt{2}}$ since whenever X_n moves both X_n^+ & X_n^- also move, but with size $\sqrt{2}$.

$\mathbb{P}(X_{n+1}^+=\frac{1}{\sqrt{2}},X_{n+1}^-=\frac{1}{\sqrt{2}})=\mathbb{P}(\text{Moving Right})=\frac{1}{4}$.
$\mathbb{P}(X_{n+1}^+=\frac{1}{\sqrt{2}})\mathbb{P}(X_{n+1}^-=\frac{1}{\sqrt{2}})=\frac{1}{2}\frac{1}{2}=\frac{1}{4}$.
By considering the other 3 cardinal directions we can prove thatX_n^+ & X_n^- are independent.

$$\begin{array}{rrcl}
\text{Then} & p_{\mathbf{00}}(2m) & = & \mathbb{P}(X_{2m}^+=\mathbf{0},X_{2m}^-=\mathbf{0}|X_0^+=\mathbf{0},X_0^-=\mathbf{0}) \\
 & & = & \mathbb{P}(X_{2m}^+=\mathbf{0}|X_0^+=\mathbf{0})\mathbb{P}(X_{2m}^-=\mathbf{0}|X_0^-=\mathbf{0}) \\
 & & = & \left(\frac{(2m)!}{(m!)^2}\frac{1}{4^m}\right)^2 \\
\text{By Stirling's Formula} & & \sim & \frac{1}{\pi m} \text{ as } m\to\infty \\
\implies & \sum_{m=M}^\infty p_{\mathbf{00}}(2m) & > & \frac{0.99}{\pi}\sum_{m=M}^\infty\frac{1}{m} \text{ for sufficiently large } M \\
 & & = & \infty \\
\implies & \mathbf{0} \text{ is recurrent} & &
\end{array}$$

3 Markov Chains in Discrete Time

Definition 3.1 - *Transition Matrix*
The *Transition Matrix* of a *Markov Chain* is the matrix P where $(P)_{ij} = p_{ij}$.

$$P = \begin{pmatrix} p_{00} & p_{01} & \dots & p_{0n} \\ p_{10} & p_{11} & \dots & p_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n0} & p_{n1} & \dots & p_{nn} \end{pmatrix}$$

N.B. We can draw automata to represent *Transition Matrices.*

Proposition 3.1 - *Properties of Transition Matrix*
Since p_{ij} are probabilities and since the process must be in some state at time 1, P is a *Transition Matrix* iff

i) $1 \leq p_{ij} \leq 0 \ \forall\ i, j \in S$; and,

ii) $\sum_{j \in S} p_{ij} = 1 \ \forall\ i \in S$.

Example 3.1 - *Gambler's Ruin*
Let X_n be the capital of the gambler, so that $S = \{0, \dots, N\}$. Recall that

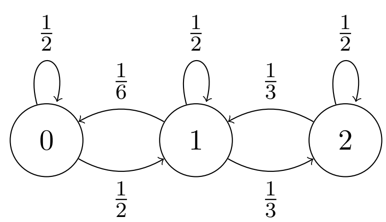
$$p_{ij} = \mathbb{P}(X_{n+1} = j | X_n = i) = \begin{cases} p & if\ i \notin \{0, N\}\ \&\ j = i + 1 \\ 1 & if\ i \notin \{0, N\}\ \&\ j = i + 1 \\ 1 & if\ i \in \{0, N\}\ \&\ j = i \\ 0 & otherwise \end{cases}$$

Therefore the transition matrix of the chain is

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ q & 0 & p & 0 & \dots & 0 & 0 \\ 0 & q & 0 & p & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & p \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}$$

Example 3.2 - *Transition Matrix Automata*
The following are a transition matrix P and its automata representation

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$



Theorem 3.1 - *Chapman-Kolmogorov Equations*
$\forall\ i, j \in S,\ n \in \mathbb{N},\ r \in [0, b]$

$$p_{ij}(n) = \sum_{k \in S} p_{ik}(r) p_{kj}(n - r)$$

Proof 3.1 - *Chapman-Kolmogorov Equations*

$$\begin{array}{rcll} p_{ij}(n) & = & \mathbb{P}(X_n = j | X_0 = i) & \\ & = & \sum_{k\in S} \mathbb{P}(X_n = j|X_r = k,\ X_0 = i)\mathbb{P}(X_r = k|X_0 = i) & \text{By Partition Theorem} \\ & = & \sum_{k\in S} \mathbb{P}(X_n = j|X_r = k)p_{ik}(r) & \text{Markov Property} \\ & = & \sum_{k\in S} \mathbb{P}(X_{n-r} = j|X_0 = k)p_{ik}(r) & \text{Time Homogeneity} \\ & = & \sum_{k\in S} p_{kj}(n-r)p_{ik}(r) & \end{array}$$

Proposition 3.2 - *Implication of Chapman-Kolmogorov Equations*
Let P_n be the matrix with $(P_n)_{ij} = p_{ij}(n)$.
The *Chapman-Kolmogorov Equations* says

$$(P_n)_{ij} = \sum_{k\in S}(P_r)_{ik}(P_{n-r})_{kj} = (P_rP_{n-r})_{ij} \implies P_n = P_rP_{n-r}$$

By considering $r = 1$ we see that

$$P_n = PP_{n-1} = \cdots = P^n$$

3.1 Analysis by Class

Definition 3.2 - *Communication*
Let $i, j \in S$. Then we can define the following relationships

i) *i communicates* with j if $\exists\ n \geq 0$ *st* $p_{ij}(n) > 0$. (Denoted $i \to j$).

ii) *i intercommunicates* with j if $i \to j$ & $j \to i$. (Denoted $i \leftrightarrow j$).

Proof 3.2 - *Intercommunication is an Equivalence Relation*
Reflexive
Since $P_{ii}(0) = 1$ then $i \to i \equiv i \leftrightarrow i$.
Symmetric
let $i \leftrightarrow j$. Then

$$\implies i \to j \ \&\ j \to i \implies j \leftrightarrow i$$

Transitive Let $i \to j$ & $j \to k$.
Then $\exists\ n, m, \in \mathbb{N}$ *st* $p_{ij}(n) > 0$ &$p_{jk}(m).0$.
Thus $p_{ik}(n + m) \geq p_{ij}(n)p_{jk}(m) > 0$ by *Chapman-Kolmogorov Equations*.
$\implies i \to k$.
Similarly $k \to i \implies i \leftrightarrow j$.

Definition 3.3 - *Communicating Classes*
Communicating Classes are partitions of the state set S $(E_1, E_2, \dots)$ st

$$\forall\ i, j \in E_r, i \leftrightarrow j$$

Proposition 3.3 - *States & Communicating Classes*
All states in the *same* communicating class intercommunicate with each other.
Any pair of states in *different* communicating classes do not intercommunicate with each other.

Theorem 3.2 - *Recurrency & Intercommunication*
Let $i \leftrightarrow j$.
Then i is recurrent iff j is recurrent.

Proof 3.3 - *Recurrency & Intercommunication*
Recall that state j is recurrent iff $\sum_{k=0}^{\infty} p_{jj}(k) = \infty$.

Assume j is recurrent & $i \leftrightarrow j$.
Then $\exists\ m, n \geq 0$ st $p_{ij}(m) > 0$ & $p_{ji}(n) > 0$.
By the *Chapman-Kolmogorov* equations $p_{ii}(m+r+n) \geq p_{ij}(m)p_{jj}(r)p_{ji}(n)$.
Thus

$$\sum_{m,r,n}^{\infty} p_{ij}(m)p_{jj}(r)p_{ji}(n) = \sum_{m=0}^{\infty} p_{ij}(m) \sum_{r=0}^{\infty} p_{jj} \sum_{n=0}^{\infty} p_{ij}(n) > \infty$$

Thus $\sum_{n=0}^{\infty} p_{ii}(n) = \infty \implies i$ is recurrent.

Remark 3.1 - *Recurrent & Transient Communicating Class*
We describe a communicating class as begin*recurrent* if the states in it are recurrent.
N.B. We define *Transient Communicating Classes* similarly.

Definition 3.4 - *Closed States*
A set of states C is *Closed* if $p_{ij} = 0\ \forall\ i \in C, j \notin C$.

Definition 3.5 - *Irreducible States*
A set of states C is *Irreducible* if $i \leftrightarrow j\ \forall\ i, j \in C$.

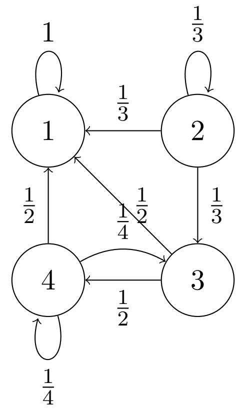
Definition 3.6 - *Absorbing State*
A state i for which the singleton set $\{i\}$ is a closed set is called an *Absorbing state.*
N.B. $p_{ii} = 1$.

Remark 3.2 - *Closed Communicating Class*
A set E is closed & irreducible iff E is a closed communicating class.

Remark 3.3 - *Closed State Space*
The state space S is always closed.

Remark 3.4 - *Irreducible Markov Chain*
If the state space S is irreducible then we say the *Markov Chain* is irreducible.

Example 3.3 - *Closed, Irreducible & Absorbing States*
In the following state space the set of inter-communication is $\{1 \leftrightarrow 1, 2 \leftrightarrow 2, 3 \leftrightarrow 3, 3 \leftrightarrow 4, 4 \leftrightarrow 4\}$.
Hence the communicating classes are $E_1 = \{1\},\ E_2 = \{2\}$ & $E_3 = \{3, 4\}$.
E_1 is closed & an absorbing state.
E_2 & E_3 are not closed, nor absorbing states, they are irreducible.



Theorem 3.3 - *Non-Closed Communicating Classes are Transient*
If C is a *Communicating Class* & C is not closed, then all the states in C are transient.

Proof 3.4 - *Non-Closed Communicating Classes are Transient*
Since C is not closed $\exists i \in C$ & $k \in C^c$ st $p_{ik} > 0$.
Furthermore, since C is a communicating classes and $k \notin C$ we have $i \not\leftrightarrow k$.
But $i \to k$ so $k \not\to i$. So, $f_{ki} = \mathbb{P}(\text{The chain } X \text{ is in state } i \text{ eventually}|X_0 = k) = 0$.
We compute f_{ii}

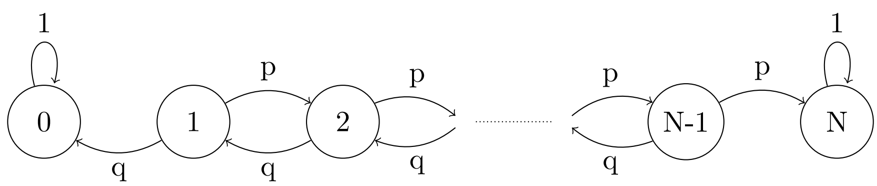
$$\begin{array}{rcl} f_{ii} & = & \mathbb{P}(X \text{ is in state } i \text{ eventually}|X_0 = i) \\ & = & \mathbb{P}(X \text{ is in state } i \text{ eventually}|X_0 = i, X_1 = j)\mathbb{P}(X_1 = j|X_0 = i) \\ & = & \sum_{j \in S} \mathbb{P}(X \text{ is in state } i \text{ eventually}|X_1 = j)p_{ij} \\ & = & \sum_{j \in S} f_{ji}p_{ij} \\ & = & f_{ki}p_{ik} + \sum_{j \in S/\{k\}} f_{ji}p_{ij} \\ \text{Since } f_{ki} = 0 & = & \sum_{k \in S/\{k\}} f_{ji}p_{ij} \\ \text{Since } f_{ji} \leq 1 & \leq & \sum_{j \in S/\{k\}} p_{ij} \\ & = & 1 - p_{ik} \\ & = & 1 \end{array}$$

Hence i is transient.
So all states in C are transient since C is a communicating class.

Theorem 3.4 - *State Spaces can be Partitioned into Transient & Recurrent States*
The state space S can be uniquely partitioned into $S = T \bigcup C_1 \bigcup \dots$ where T is a set of transient states & each C_k is an irreducible closed set of recurrent states.

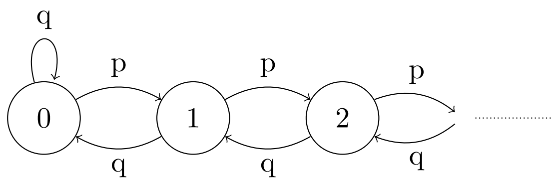
Proof 3.5 - *State Spaces can be Partitioned into Transient & Recurrent States*
We can partition S into communicating classes.
These classes are either recurrent or transient.
Let C_k for $k \in \mathbb{N}$ be the recurrent communicated classes & T be the union of all transient communicating classes.
By **Theorem 3.3** C_k must be closed.
Thus each C_k is irreducible since it is a communicating class.

Example 3.4 - *Gambler's Ruin*
Consider the following diagram for the Gambler's Ruin



Here we have communicating classes $\{0\}, \{1, \dots, N-1\}$ & $\{N\}$.
$\{0\}$ & $\{N\}$ are closed & are absorbing states. They are recurrent.
$\{1, \dots, N-1\}$ is not closed since $1 \to 0$ & $N - 1 \to N$, hence it is transient.

Example 3.5 - *Random Walk with Reflecting Barrier at 0*
Consider the following random walk which has a reflecting barrier at 0



Here the only communicating class is $\{0, 1, \dots\}$. It is closed.
All the states must be of the same recurrent type but which type they are depends on the values of p & q.
We shall consider computing $\mathbb{P}(A|X_0 = 0)$ where $A = \{X \text{ hits } 0 \text{ eventually}\}$. There are three cases

i) $p = q = \frac{1}{2}$.

$$\begin{aligned}\mathbb{P}(A|X_0 = 0) &= \mathbb{P}(A|X_0 = 0, X_1 = 0)\mathbb{P}(X_1 = 0|X_0 = 0) + \mathbb{P}(A|X_0 = 0, X_1 = 1)\mathbb{P}(X_1 = 1|X_0 = 0) \\ &= 1 \times q + 1 \times p \\ &= p + q \\ &= 1\end{aligned}$$

Here state 0 is recurrent.

ii) $p < q$.

$$\mathbb{P}(A|X_0 = 0) = 1 \times q + 1 \times p = 1$$

Here state 0 is recurrent.

iii) $q > p$.

$$\mathbb{P}(A|X_0 = 0) = 1 \times q + \mathbb{P}(A|X_1 = 1)p < 1$$

Here state 0 is transient.

3.2 Stationary Distributions

Definition 3.7 - *Stationary Distribution*
Let π denote a *horizontal vector* with a component for each state, $\pi = (\pi_1, \dots, \pi_j)$.
Let $X = \{X_n\}_{n\in\mathbb{N}}$ be a Markov chain with transition matrix P.
We say π is a *Stationary Distribution* of the chain if

i) $\pi_j \geq 0\ \forall\ j \in S$ & $\sum_{j\in S} \pi_j = 1$ (π is a mass function for S).

ii) $\pi = \pi P$, note that this is matrix multiplication.

N.B. Further $\pi = \pi P^n\ \forall\ n \in \mathbb{N}$.

Example 3.6 - *Stationary Distribution*

Consider a Markov chain with states $S = \{0, 1, 2\}$ & transition matrix $P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$.

Find a *Stationary Distribution* for this chain.

$$\begin{array}{rrcl}
\text{Set} & (\pi_0,\pi_1,\pi_2) & = & (\pi_0,\pi_1,\pi_2)\begin{pmatrix}\frac{1}{2} & \frac{1}{2} & 0\\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3}\\ 0 & \frac{1}{2} & \frac{1}{2}\end{pmatrix}\\
 & & = & (\frac{\pi_0}{2}+\frac{\pi_1}{6},\frac{\pi_0}{2}+\frac{\pi_1}{2}+\frac{\pi_2}{2},\frac{\pi_1}{3}+\frac{\pi_2}{2})\\
\implies & \pi_0 & = & \frac{\pi_0}{2}+\frac{\pi_1}{6}\\
\implies & \pi_0 & = & \frac{\pi_1}{3}\\
\& & \pi_2 & = & \frac{\pi_1}{3}+\frac{\pi_2}{2}\\
\implies & \pi_2 & = & \frac{2\pi_1}{3}\\
\text{Since} & \pi_0+\pi_1+\pi_2 & = & 1\\
\implies & \frac{\pi_1}{3}+\pi_1+\frac{2\pi_1}{3} & = & 1\\
\implies & \pi_1 & = & \frac{1}{2}\\
\implies & \pi & = & (\frac{1}{6},\frac{1}{2},\frac{1}{3})
\end{array}$$

Example 3.7 - *Random Walk with Reflecting Barrier at 0*
Consider the diagram in **Example 3.5**.
We are going to try and find a *Stationary Distribution* for this situation.

$$\begin{array}{rrcll}
\text{Let} & \pi & = & \pi P & \\
\implies & \pi_0 & = & q\pi_0+q\pi_1 & \\
\& & \pi_j & = & p\pi_j+q\pi_{j+1} & j\geq 1\\
\implies & (1-q)\pi_0 & = & q\pi_1 & \\
\implies & p\pi_0 & = & q\pi_1 & \\
\& & (p+q)\pi_j & = & p\pi_{j-1}+q\pi_{j+1} & j\geq 1\\
\implies & p(\pi_j-\pi_{j-1}) & = & q(\pi_{j+1}-\pi_j) & \\
\text{Summing the first } j \text{ equations} & p(\pi_0-\pi_0+\pi_1-\dots-\pi_{j-1}+\pi_j) & = & q(\pi_1-\pi_1+\dots+\pi_{j+1}) & \\
\implies & \pi_j p & = & q\pi_{j+1} & \\
\implies & \pi_{j+1}=\frac{p}{q}\pi_j & & & \\
\implies & \pi_j & = & \left(\frac{p}{q}\right)^j\pi_0 & \\
\text{Since} & \sum_{j=0}^{\infty}\pi_j & = & 1 & \\
\implies & \sum_{j=0}^{\infty}\left(\frac{p}{q}\right)^j\pi_0 & = & 1 & \\
\implies & \pi_0\sum_{j=0}^{\infty}\left(\frac{p}{q}\right)^j & = & 1 & \\
 & & = & \begin{cases}\infty & \text{If } p\geq q\\ \pi_0\dfrac{1}{1-\frac{p}{q}} & \text{If } p<q\end{cases} &
\end{array}$$

Thus π is a stationary distribution (finite) iff $p<q$.

3.3 Existence & Uniqueness of Stationary Distributions

Definition 3.8 - *Positive Recurrent*
A state i is positive recurrent if the mean time of first return is finite ($m_{ii}<\infty$).
N.B. This is a class property.

Theorem 3.5 - *Stationary Distributions in Irreducible Chains*
An *Irreducible Chain* has a *Stationary Distribution* π iff all of the states are positive recurrent.
Here π is unique and given by $\pi_i=\frac{1}{m_{ii}}$.

Theorem 3.6 -
Let $X = \{X_n\}_{n\in\mathbb{N}}$ be an irreducible aperiodic *Markov Chain* with a stationary distribution π. Then

$$\forall\ i, j \in S\ p_{ij}(n) \xrightarrow{n\to\infty} \pi_j$$

N.B. The existence of the π means all states are positive recurrent with $m_{ii} = \frac{1}{\pi_i}$

3.4 Periodicity

Definition 3.9 - *Period*
Let $j \in S$ be such that $p_{jj}(n) > 0$ for some integers $n \geq 1$.
Let $\mathcal{N}_j = \{n \geq 1 : p_{jj}(n) > 0\}$.
Then the *Period* of j is given by $d_j = gcd(\mathcal{N}_j)$.

Definition 3.10 - *Aperiodic*
If the period of a state, d_j, equals 1 then state j is aperiodic.

Remark 3.5 -
If $p_{jj} > 0$ then J is aperiodic since $1 \in \mathcal{N}_j$.

Remark 3.6 -
if $d_j \geq 2$ then $p_{jj}(n)$ cannot possibly converge, except occasionally to 0.

Theorem 3.7 - *Period within a Communicating Class*
Let $C \subseteq S$ be a communicating class of states.
Let $i, j \in C$.
Then $d_i = d_j$.

Proof 3.6 - *Period within a Communicating Class*
Suppose $d_i \neq d_j$, and without loss of generality assume $d_i < d_j$.
Since C is a communicating class, $i \leftrightarrow j$, and there exists $m \geq 0,\ n \geq 0$ st $p_{ij}(m) > 0$ and $p_{ji}(n) > 0$.
Suppose $p_{ii}(s) > 0$.
Then $p_{ii}(rs) \geq r.p_{ii}(s) > 0$ for $r \in \mathbb{N}$.
Therefore $p_{jj}(n + rs + m) \geq p_{ji}(n)p_{ii}(rs)p_{ij}(m) > 0$ so $(n + rs + m) \in \mathcal{N}_j$.
In particular, both $(n+s+m)$ and $(n+2s+m)$ are multiples of d_j, so $s = (n+2s+m)-(n+s+m) = s$ is a multiple of d_j.
We have shown that d_j is a divisor of all the values of s such that $p_{ii}(s) > 0$.
d_i is defined to be the *gcd* of the same set of s values, so $d_i \geq d_j$.
But we started by assuming $d_i < d_j$.
This is a contradiction.
So we must have $d_i = d_j$.

Example 3.8 - *Period*
Let $S = \{0, 1\}$ and with transition matrix $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.
So $p_{00}(n) = \begin{cases} 0 & \text{if } n \text{ is odd} \\ 1 & \text{if } n \text{ is even} \end{cases}$.
Since $p_{00}(n) > 0$ only when n is event, the period must be a multiple of 2.
And, since $p_{00}(2) > 0$, the period must be no more than 2.
Therefore $d_0 = 2$.

4 Markov Chains in Continuous Time

Definition 4.1 - *Lack of Memory Property*
Let X be a random variable.
X is said to have the *Lack of Memory Property* if

$$\mathbb{P}(X > t + h | T > h) = \mathbb{P}(T > h)$$

Remark 4.1 - *Distributions with Lack of Memory Property*
The *Exponential Distribution* has the *Lack of Memory Property.*

Proposition 4.1 - *Modelling Wait Times*
LET T be a random variable that models wait times between events.
Then T should have the following property

$$\begin{array}{rrcl} & \mathbb{P}(T \in (t, t+h] | T > t) & = & \lambda h + o(h) \\ \implies & \mathbb{P}(T > t + h | T > t) & = & 1 - \lambda h + o(h) \end{array}$$

N.B. $\lim\limits_{h \to 0} \frac{0(h)}{h} = 0$.

Proposition 4.2 - *Distribution for Modelling Wait Times*
Consider the T defined in **Proposition 4.1**.
Let $g(t) = \mathbb{P}(T > t)$. Then

$$\begin{array}{rrcl} & g(t+h) & = & \mathbb{P}(T > t+h) \\ & & = & \mathbb{P}(T > t+h | T > t)\mathbb{P}(T > t) + \mathbb{P}(T > t+h | T \leq t)\mathbb{P}(T \leq t) \\ & & = & (1 - \lambda h + o(h))g(t) + 0 \\ & & = & (1 - \lambda h)g(t) + o(h) \\ \implies & \dfrac{g(t+h) - g(t)}{h} & = & \dfrac{-\lambda h g(t)}{h} = \dfrac{o(h)}{h} \\ \implies & g'(t) & = & -\lambda g(t) + 0 \\ \implies & g(t) & = & -\frac{g'(t)}{\lambda} \\ & & = & ce^{-\lambda t} \end{array}$$

Since $g(0) = 1 \implies 1 = ce^0 \implies c = 1 \implies g(t) = e^{-lambdat}$.
So $T \sim Exp(\lambda)$.

Example 4.1 - *Who arrives first?*
Consider having arranged to meet with n friends.
Let T_i be the length of time you wait for the i^{th} first to arrived.
Assume the T_is are independent and identically distributed with $T_i \sim Exp(\lambda_i)$.

i) Derive the distribution of the length of time until the first friend turns up.

$$\begin{array}{rrcl} & \mathbb{P}(min(T_1, ..., T_n) > t) & = & \mathbb{P}(t_1 > t, \dots, T_n > T) \\ & & = & \prod_i \mathbb{P}(T_i > t) \text{ by indepdence} \\ & & = & \prod_i e^{-\lambda_i t} \\ & & = & e^{-t\sum_i \lambda_i} \\ \implies & min(T_1, \dots, T_n) & \sim & Exp\left(\sum_{i=1}^n \lambda_i\right) \end{array}$$

ii) Derive the probability that the first friend to turn up is friend i.

$$\begin{array}{rcl}\mathbb{P}(i^{th}\ friend\ is\ first) & = & \mathbb{P}(T_i = min(T_1, \dots, T_n)) \\ & = & \mathbb{P}(T_i < T_j, j \neq i) \\ & = & \int\limits_0^\infty \mathbb{P}(T_j > T_i, j \neq i | T_i = t)dt \\ & = & \int\limits_0^\infty \mathbb{P}(T_j > t, j \neq i)\lambda e^{-\lambda_i t}dt \\ & = & \int\limits_0^\infty \prod\limits_{j=1, j\neq i}^n \mathbb{P}(T_j > t)\lambda_i e^{-\lambda_i}dt \text{ by independce} \\ & = & \int\limits_0^\infty e^{-t\sum_{j\neq i}\lambda_j}\lambda_i e^{-\lambda_i t}dt \\ & = & \lambda_i \int\limits_0^\infty e^{-t\sum_{i=1}^n \lambda_i}dt \\ & = & \lambda_i \left[\frac{exp\left(-t\sum_{i=1}^n \lambda_i\right)}{-\lambda_{i=1}^n \lambda_i}\right]_0^\infty \\ & = & \frac{\lambda_i}{\sum_{i=1}^n \lambda_i}\end{array}$$

Proposition 4.3 - *Link Exponential & Binomial Distribution*
Consider the time till occurrence of an event is modelled by $Exp(\mu)$.
So the probability of the event occurring before time t is $1 - e^{-\mu t} \approx \mu t$ if μt is small.
If there are n independent events with the same distribution we can model the number of events that occur by time t with $Bi(n, \mu t)$

4.1 Poisson Process

Definition 4.2 - *Poisson Process*
Let $\{N_t\}_{t\geq 0}$ be a continuous time stochastic process where N_t counts the number of events to have occurred by time t.
$\{N_t\}_{t\geq 0}$ is a *Poisson Process* with rate λ if

i) $N_t \in \mathbb{N}\ \forall\ t \geq 0$;

ii) $N_0 = 0$;

iii) It has stationary increments *i.e.* $N_{t+s} - N_t$ depends only on s;

iv) It has independent increments *i.e.* $N_{t_2} - N_{t_1}, \dots, N_{t_n} - N_{t_{n-1}}$ are independent;

v) If $t > 0$ and $h > 0$ then

$$\begin{array}{rcl}\mathbb{P}(N_{t+h} - N_t < 0) & = & 0 \\ \mathbb{P}(N_{t+h} - N_t = 0) & = & 1 - \lambda h + o(h) \\ \mathbb{P}(N_{t+h} - N_t = 1) & = & \lambda h + o(h) \\ \mathbb{P}(N_{t+h} - N_t > 1) & = & o(h)\end{array}$$

as $h \to 0$.

N.B. $\lim\limits_{h\to 0}\frac{o(h)}{h} = 0$.

Theorem 4.1 - *Distribution of Poisson Process*
for $t \geq 0$

$$N_t \sim Po(\lambda t)$$

Proof 4.1 - *Theorem 4.1*

Proposition 4.4 - *Implications of Theorem 4.1*
From **Theorem 4.1** it follows that $\mathbb{E}(N_t) = Var(N_t) = \lambda t$.
Since $N_0 = 0$ we have stationary increments

$$N_{t+s} - N_t \sim N_s - N_0 = N_S - 0 \sim N_S \sim Po(\lambda s)$$

Thus the number of events in any interval of width s is distributed by $Po(\lambda s)$.

Proposition 4.5 - *Distribution of Initial Arrival Time*
Let S_1 be the time of the first arrival, so $S_1 = inf\{t \geq 0 : N_t > 0\}$.
Consider the event $\{S_1 > t\}$.
We have that $\{S_1 > t\} \equiv \{N_t = 0\}$.
Hence $\mathbb{P}(S_1 > t) = \mathbb{P}(N_t = 0) = e^{-\lambda t}$ since $N_t \sim Po(\lambda t)$.
Thus $S_1 \sim Exp(\lambda)$.

Proposition 4.6 - *Distribution of General Arrival Time*
Let S_n be the time of the n^{th} arrival, so $S_n = inf\{t \geq 0 : N_t = n\}$.
Consider the event $\{S_n > t\}$.
We have that $\{S_n > t\} \equiv \{N_t < n\}$.
Let T_i be the inter-arrival times st $T_1 = S_1,\ T_2 = S_2 - S_1, \dots, T_n = S_n - S_{n-1}$.
Note that $S_n = \prod_{i=1}^{n} T_i$.

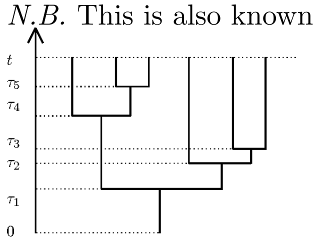
$$\begin{aligned}
\mathbb{P}(T_n > t | S_{n-1} = s_{n-1}, \dots, S_1 = s_1) &= \mathbb{P}(N_{t+s_{n-1}} - N_{s_{n-1}} = 0 | S_{n-1} = s_{n-1}, \dots, S_1 = s_1) \\
&= \mathbb{P}(N_{t+s_{n-1}} - N_{s_{n-1}} = 0) \text{ by indepdent increments} \\
&= \mathbb{P}(N_t - N_0 = 0) \text{ by stationary increments} \\
&= \mathbb{P}(N_t = 0) \\
&= e^{-\lambda t}
\end{aligned}$$

Since $e^{-\lambda t}$ is independent of $s_1, \dots, s_{n-1} \implies T_n$ is independent of $S_1, \dots, S_{n-1}$.
Hence T_n is independent of $T_1, \dots, T_{n-1}$ be definition.
So $T_i \sim Exp(\lambda)\ \forall\ i$.
Hence $S_n = \sum_{i=1}^{n} T_i \sim \Gamma(n, \lambda)$.

4.2 Birth Death Process

4.2.1 Linear Birth Processes

Definition 4.3 - *Linear Birth Process*
Consider a population of individuals.
In the *Linear Birth Process* each individual present at time t splits into 2 during the interval $(t, t+h)$ with probability $\lambda h + o(h)$.
We have that all individuals act independently of each other.
N.B. The birth rate is independent of the population size.
N.B. This is also known as a *Yule Process*.



Proposition 4.7 - *Distribution of Time between Arrivals*
Let N_t be the population size at time t, assume $N_0 = 1$.

For $n \geq 1$ let S_n be the time until the population reaches size n, so $S_n = inf\{t \geq 0 : N_t = n\}$.
Let T_j be the time to grow from size j to $j+1$, so $T_j = S_{j+1} - S_j$.
Consider a population of size j.
Let Y_i be hte waiting time until i^{th} individual splits.
then, by definition & **Proposition 4.2**, we have that $Y_1, \dots, Y_j$ are iid with distribution $Exp(\lambda)$.
We have that $T_j = min\{Y_1, \dots, Y_j\}$.
Since $T_j \sim Exp(k\lambda) \implies \mathbb{E}(T_j) = \frac{1}{j\lambda}$.
Using $S_n = T_1 + \dots + T_{n-1}$ we have that

$$\mathbb{E}(S_n) = \mathbb{E}(T_1 + \dots + T_{n-1}) = \frac{1}{\lambda}(1 + \frac{1}{2} + \dots + \frac{1}{n-1})$$

This is not convergent but can be approximated to $\frac{1}{\lambda}\int_1^n \frac{1}{x}dx = \frac{1}{\lambda}\log_n$.

Proposition 4.8 - *Distribution of Number of Births*
For a population of size n each has an independent probability of $\lambda h + o(h)$ for giving birth in interval $(t, t+h)$.
The distribution of number of births in time period $(t, t+h)$ is

$$Bi(n, \lambda h + o(h))$$

Proposition 4.9 - *Distribution of Population Size, N_t*
Let N_t denote the size of the population at time t with birth rate λ, then

$$N_T \sim Geo(e^{-\lambda t})$$

N.B. $\mathbb{E}(N_t) = e^{\lambda t}$.

4.2.2 Linear Birth & Death Processes

Definition 4.4 - *Linear Birth & Death Process*
Consider an individual in a population & the possible events that can occur to them in $(t, t+h)$.
In a *Linear Birth & Death Process* these are

i) Gives birth, splits in two, with probability $\lambda h + o(h)$;

ii) Dies with probability $\mu h + o(h)$; or,

iii) Neither event, with probability $1 - (\lambda + \mu)h + o(h)$.

N.B. We can consider the random variables $B \sim Exp(\lambda)$ & $D \sim Exp(\mu)$.
N.B. The birth & death rates are independent of the population size.

Proposition 4.10 - *Probability of Birth or Death*
Given an event has occurred:

i) The probability it was a birth is $\dfrac{\lambda}{\lambda + \mu}$

ii) The probability it was a death is $\dfrac{\mu}{\lambda + \mu}$

4.2.3 Generalised Birth & Death Processes

Definition 4.5 - *Generalised Birth & Death Process*
A continuous time stochastic process $\{N_t\}_{t \geq 0}$ is a *Generalised Birth & Death Process* if

- $N_t \in \mathbb{N}_0$;
- $\mathbb{P}(N_{t+h} - N_t = 1 | N_t = n) = \lambda_n h + o(h)$
- $\mathbb{P}(N_{t+h} - N_t = -1 | N_t = n) = \mu_n h + o(h)$
- $\mathbb{P}(N_{t+h} - N_t = 0 | N_t = n) = 1 - (\lambda_n + \mu_n)h + o(h)$
- $\lambda_n, \mu_n \geq 0 \ \forall \ n$;
- $\mu_0 = 0$.

N.B. The birth and death rate depends upon the size of the population.

Proposition 4.11 - *General Rate of Change*
Let $p_n(t) = \mathbb{P}(N_t = n)$.
Then

$$p_n'(t) = \lambda_{n-1}p_{n-1}(t) - (\lambda_n + \mu_n)p_n(t) + \mu_{n=1}p_{n+1}(t) \ \forall \ n \geq 1$$

and

$$p_0'(t) = -\lambda_0 p_0 * t) + \mu_1 p_1(t)$$

Example 4.2 - *Stationary Distribution - Constant Birth Rate & Increasing Death Rate*
Consider a situation where $\lambda_n = \lambda > 0 \ \forall \ n \in \mathbb{N}$ & $\mu_n = \mu n \forall \ n \in \mathbb{N}$.
Derive the stationary distribution for this process Let $p_n(t) = \mathbb{P}(N_t = n)$. From the general formula in **Proposition 4.11** we have

$$p_n'(t) = \lambda p_{n-1}(t) - (\lambda + \mu n)p_n(t) + \mu(n+1)p_{n+1}(t)$$

Consider the case that $n = 0$

$$\begin{array}{rcl} p_0(t+h) & = & \mathbb{P}(N_{t+h} = 0) \\ & = & \mathbb{P}(N_{t+h} = 0 | N_t = 0)\mathbb{P}(N_t = 0) + \mathbb{P}(N_{t+h} = 0 | N_t = 1)\mathbb{P}(N_t = 1) + o(h) \\ & = & (1 - \lambda h)p_0(t) + \mu h p_1(t) + o(h) \\ \implies \dfrac{p_0(t+h) - p_0(t)}{h} & = & -\lambda p_0(t) + \mu p_1(t) \end{array}$$

Let $h \to 0 \implies p_0'(t) = -\lambda p_0(t) + \mu p_1(t)$.
We assume the process reaches a steady space where it no longer changes, this occurs at large t.
Then means $\exists \hat{p}_n$ where

$$p_n(t) \xrightarrow{t \to \infty} \hat{p}_n \ \& \ \hat{p}_n'(t) \to 0$$

By the general formula & derivative of $\hat{p}_0$ we have that

$$\begin{array}{rcl} 0 & = & \lambda \hat{p}_{n-1} - (\lambda + \mu n)p_n + \mu(n+1)\hat{p}_{n+1} \\ 0 & = & -\lambda \hat{p}_0 + \mu \hat{p}_1 \end{array}$$

Rearranging gives

$$\begin{array}{rcl} \mu(n+1)\hat{p}_{n+1} - \lambda \hat{p}_n & = & \mu_n \hat{p}_n - \lambda \hat{p}_{n+2} \\ & \dots & \\ & = & \mu \hat{p}_1 - \lambda \hat{p}_0 = 0 \end{array}$$

Hence $\mu n \hat{p}_n - \lambda \hat{p}_{n-1} = 0 \ \forall \ n \in \mathbb{N}$.

$$\implies \hat{p}_n = \frac{\lambda}{\mu n}\hat{p}_{n-1} = \dots = \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n \hat{p}_0$$

Normalising

$$\begin{array}{rcl} 1 & = & \sum_{n=0}^{\infty} \hat{p}_n \\ & = & \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{\lambda}{\mu}\right) \hat{p}_0 \\ & = & \hat{p}_0 e^{\lambda/\mu} \\ \implies \hat{p}_0 & = & e^{-\lambda/\mu} \end{array}$$

Thus $\hat{p}_n = \frac{1}{n!}\left(\frac{\lambda}{\mu}\right) e^{-\lambda/\mu}\ \forall\ n \in \mathbb{N}^0$.

This stationary distribution is distributed $\sim Po\left(\frac{\lambda}{\mu}\right)$.

4.3 General Markov Chains in Continuous Time

Definition 4.6 - *Generator*
G is the *Generator* of a *Markov Chain* where G satisfies

i) $g_{ij} \geq 0$ for $j \neq i$;

ii) $g_{ii} \leq 0$; and

iii) $\sum_{j \in S} = 0$.

N.B. $g_{ii} = -\sum_{j \neq i} g_{ij}$.

Proposition 4.12 - *Chains & Their Generator*
For reasonable chains it can be shown that

$$\begin{array}{rcl} p_{ij}(h) & = & g_{ij}h + o(h) \text{ for } i \neq j \\ p_{ii}(h) & = & 1 - \sum_{j \neq i} p_{ij}(h) \\ & = & 1 - \sum_{j \neq i} g_{ij}h + o(h) \\ & = & 1 + g_{ii}h + o(h) \end{array}$$

Example 4.3 - *Linear Birth & Death Process*
We know that

$$\mathbb{P}(N_{t+h} = j | N_t = i) = p_{ij}(h) = \begin{cases} \mu i h + o(h) & \text{if } j = i - 1 \\ \lambda i h + o(h) & \text{if } j = i + 1 \\ 1 - (\lambda + \mu)hi + o(h) & \text{if } j = i \\ o(h) & \text{otherwise} \end{cases}$$

Using $p_{ij}(h) = g_{ij}h + o(h)$ & $p_{ii}(h) = g_{ii}h + 1 + o(h)$

$$g_{ij} = \begin{cases} \mu & \text{j} = \text{i} - 1 \\ \lambda & \text{j} = \text{i} + 1 \\ -(\lambda + \mu)i & \text{j} = \text{i} \\ 0 & \text{otherwise} \end{cases}$$

Definition 4.7 - *Forward Equations*
The *Forward Equations* is given as

$$(P_t')_{ij} = (P_t G)_{ij}$$

Since this holds $\forall\ i, j$ we have $P_t' = P_t G$ & $P_0 = I$.

Proposition 4.13 - *Deriving Forward Equations*
Consider the probability of being in state j at time $t + h$ if the chain started in i at time 0,

$p_{ij}(t+h)$.
Conditioning on the state of the chain at time t we get

$$\begin{array}{rcl} p_{ij}(t+h) & = & \mathbb{P}(X_{t+h}=j|X_0=i) \\ & = & \sum_{k\in S}\mathbb{P}(X_{t+h}=j|X_t=k,X_0=i)\mathbb{P}(X_t=k|X_0=i) \\ & = & \sum_{k\in S}\mathbb{P}(X_t=k,X_0=i)\mathbb{P}(X_t=k|X_0=i) \text{ by Markov Property} \\ & = & \sum_{k\in S}p_{kj}(h)p_{ik}(t) \\ & = & p_{jj}(h)p_{ij}(t)+\sum_{k\in S/\{s\}}p_{kj}(h)p_{ik}(t) \\ & = & (1+g_{jj}h+o(h))p_{ij}(t)+\sum_{k\in S/\{j\}}(g_{kj}h+o(h))p_{ik}(t) \\ & = & p_{ij}(t)+\sum_{k\in S}g_{kj}hp_{ik}(t)+o(h) \\ \implies \dfrac{p_{ij}(t+h)-p_{ij}(t)}{h} & = & \sum_{k\in S}g_{kj}p_{ik}(t)+\dfrac{o(h)}{h} \\ \implies p'_{ij}(t) & = & \sum_{k\in S}g_{kj}p_{ik}(t) \\ \implies (P'_t)_{ij} & = & (P_tG)_{ij} \end{array}$$

Proposition 4.14 - *Solving Forward Equations*
We can solve the *Forward Equations* as ordinary differential equations

$$\begin{array}{rrcl} & P'_t-P_tG & = & 0 \\ \implies & I(t) & = & e^{-\int Gdt}=e^{-Gt} \\ \implies & \frac{d}{dt}(e^{-Gt}P_t) & = & 0 \\ \implies & e^{-G}P_t & = & C \text{ some matrix} \\ \implies & e_0P_0 & = & C \\ \implies & 1I & = & C \\ \implies & C & = & I \\ \implies & P_t & = & e^{Gt} \end{array}$$

Theorem 4.2 -

$$\left.\frac{d^k}{dt^k}\right|_{t=0}P_t=G_k\ \forall\ k\geq 0$$

Proof 4.2 - *Theorem 4.2*

$$\begin{array}{rcl} \left.\frac{d^k}{dt^k}\right|_{t=0}P_t & = & \left.\frac{d^k}{dt^k}\right|_{t=0}\sum\limits_{n=0}^{\infty}\dfrac{t^n}{n!}G^n \\ & = & \sum\limits_{n=0}^{\infty}\left.\frac{d^k}{dt^k}\right|_{t=0}\dfrac{t^n}{n!}G^n \\ & = & \sum\limits_{n=0}^{\infty}\left.\frac{t^{n-k}}{(n-k)!}\right|_{t=0}G^n \\ & = & G^k \end{array}$$

Example 4.4 - *Computing P_t from G*

Consider the state space $S=\{1,2,3\}$ with $G=\begin{pmatrix}-2 & 1 & 1\\ 1 & -1 & 0\\ 2 & 1 & -3\end{pmatrix}$.
We partially diagonalise G.
Setting $|G-\lambda I|=0$ we get $\lambda_1=0,\lambda_2=-2\ \&\ \lambda_3=-4$.
Since the eigenvalues are distinct $\exists U\in M_3$ st

$$\begin{array}{rrcl} & G & = & U\begin{pmatrix}0 & 0 & 0\\ 0 & -2 & 0\\ 0 & 0 & -4\end{pmatrix}U^{-1} \\ \implies & G^n & = & U\begin{pmatrix}0^n & 0 & 0\\ 0 & (-2)^n & 0\\ 0 & 0 & (-4)^n\end{pmatrix}U^{-1} \end{array}$$

We know that $P_t = e^{tG}$ so

$$\begin{aligned}
P_t &= \sum_{n=0}^{\infty} \frac{t^n G^n}{n!} \\
&= \sum_{n=0}^{\infty} \frac{t^n}{n!} U \begin{pmatrix} 0^t & 0 & 0 \\ 0 & (-2)^t & 0 \\ 0 & 0 & (-4)^t \end{pmatrix} U^{-1} \\
&= U \begin{pmatrix} \sum_{n=0}^{\infty} \frac{0^n}{n!} & 0 & 0 \\ 0 & \sum_{n=0}^{\infty} \frac{(-2t)^n}{n!} & 0 \\ 0 & 0 & \sum_{n=0}^{\infty} \frac{(-4t)^n}{n!} \end{pmatrix} U^{-1} \\
&= U \begin{pmatrix} e^0 & 0 & 0 \\ 0 & e^{-2t} & 0 \\ 0 & 0 & e^{-4t} \end{pmatrix} U^{-1}
\end{aligned}$$

Hence $P_{ij}(t) = a_{ij} + b_{ij}e^{-2t} + c_{ij}e^{-4t}\ \forall\ i, j \in S$. Using **Theorem 4.2** & considering $i = j = 1$ we have

$$\begin{aligned}
P_{11}(t) &= a_{11} + b_{11}e^{-2t} + c_{11}e^{-4t}, \\
P'_{11}(t) &= -2b_{11}e^{-2t} - 4c_{11}e^{-4t} \\
\&\ P''_{11}(t) &= 4b_{11}e^{-2t} + 16c_{11}e^{-4t},
\end{aligned}$$

Setting $t = 0$

$$\begin{aligned}
P_{11}(0) &= (G^0)_{11} \\
&= a_{11} + b_{11} + c_{11} \\
P'_{11}(0) &= (G^1)_{11} = -2 \\
&= -2b_{11} - 4c_{11} \\
P''_{11}(0) &= (G^2)_{11} \\
&= (-2)^2 + (1 \times 1) + (2 \times 1) = 7 \\
&= 4b_{11} + 16c_{11}
\end{aligned}$$

Solving this series of simultaneous equations we get

$$a_{11} = \frac{3}{8},\ b_{11} = \frac{1}{4}\ \&\ c_{11} = \frac{3}{8}$$

Example 4.5 - *Exponential Holding Times*
Assume at time t the chain enters state i & then stays in i for some random time $U = inf\{s \geq 0 : X_{t+s} \neq i\}$. It the jumps to a new state.
Let h_{ij} be the probability it jumps from state i to state j, with $h_{ii} = 0$.
We have that

$$\begin{aligned}
\mathbb{P}(U > u + v | U > v) &= \mathbb{P}(U > v + u | X_{t+u} = i \text{ and further information about the past}) \\
&= \mathbb{P}(U > v)\text{by Markov property}
\end{aligned}$$

Since U has last of memory property & is on continuous time it must have an exponential distribution.
Let $U \sim Exp(\lambda_i)$. We shall relate the values of λ_i & h_{ij} to the generator matrix G.
For $i \neq j$. Since $g_{ij} = p'_{ijj}(0)$ then

$$\begin{aligned}
g_{ij} &= \lim_{\delta\to 0} \tfrac{p_{ij}(delta) - p_{ij}(0)}{\delta} \\
&= \lim_{\delta\to 0} \tfrac{p_{ij}(delta) - 0(0)}{\delta} \\
&= \lim_{\delta\to 0} \tfrac{\mathbb{P}(\text{in time}(0,\delta),\text{leave state } i, \text{jump to } j, \text{ and nothing else}) + o(\delta)}{\delta} \\
&= \lim_{\delta\to 0} \tfrac{\mathbb{P}(U<\delta)h_{ij} + o(\delta)}{\delta} \\
&= \lim_{\delta\to 0} \tfrac{(1-e^{-\lambda_i \delta})h_{ij} + o(\delta)}{\delta} \\
&= \lambda_i h_{ij} \text{ by l'Hôpital's rule}
\end{aligned}$$

For $i=j$

$$\begin{array}{rcl} g_{ii} & = & \lim_{\delta\to0}\frac{p_{ii}(delta)-p_{ii}(0)}{\delta} \\ & = & \lim_{\delta\to0}\frac{e^{-\lambda_i\delta}+o(\delta)-1}{\delta} \end{array}$$

$= -\lambda_i$ by l'Hôpital's rule

Rearranging this we see that

$$\begin{array}{rrcl} & \lambda_i & = & -g_{ii} \\ & h_{ij} & = & \dfrac{g_{ij}}{\lambda_i} \\ & & = & -\dfrac{g_{ij}}{g_{ii}} > 0 \\ \implies & \sum_{j\in S} h_{ij} & = & \sum_{j\neq i} -\frac{g_{ij}}{g_{ii}} + h_{ii} \\ & & = & \sum_{j\neq i} -\frac{g_{ij}}{g_{ii}} \\ & & = & 1 \end{array}$$

Thus the matrix H is a stochastic matrix.
The chain X_t proceeds with a sequence of jumps around the state space.

Definition 4.8 - *Jump Chain*
Let $S_1 < S_2 < \dots$ denote the times when the continuous-time Markov Chain X jumps.
The discrete-time Markov Chain formed by $(X_{S_1}, X_{S_2}, \dots)$ is called the *Jump Chain*.
N.B. This is AKA *Embedded Markov Chain*.

Example 4.6 - *Jump Chain for Linear Births & Deaths*
Define

$$p_{ij}(h) = \begin{cases} o(h) & j < i-1 \ or\ j > i+1 \\ \mu ih + o(h) & j = i-1 \\ 1-(\lambda+\mu)ij + o(h) & j = i \\ \lambda ih + o(h) & j = i+1 \end{cases}$$

and

$$g_{ij} = \begin{cases} 0 & j < i-1 \ or\ j > i+1 \\ \mu i & j = i-1 \\ -(\lambda+\mu)i & j = i \\ \lambda i & j = i+1 \end{cases}$$

Note that $U \sim Exp(-g_{ii}) \sim Exp((\lambda+\mu)i)$.
Hence, the jump chain transition probabilities are

$$h_{ij} = \begin{cases} 0 & j < i-1 \ or\ j > i+1 \\ \frac{\mu}{\lambda+\mu} & j = i-1 \\ \frac{\lambda}{\lambda+\mu} & j = i+1 \\ 0 & j = i \end{cases}$$

4.4 Class Structure, Recurrence & Stationary Distributions

Definition 4.9 - *Equivalent Definitions of Communication*
If $i \neq j$ the following are equivalent

i) $i \to j$;

ii) $i \to j$ in the jump chain;

iii) $\exists i \neq i_1 \neq \dots \neq i_n \neq j$ st $g_{ii_1}g_{i_1i_2}\dot{g}_{i_nj} > 0$;

iv) $p_{ij}(t) > 0 \ \forall \ t$

Definition 4.10 - *Continuous Time Recurrence*
A state i is called *Recurrent* in continuous time if $\mathbb{P}(\{t \geq 0 : X_t = i\} \ is \ unbounded | X_0 = i) = 1$.

Definition 4.11 - *Continuous Time Transience*
A state i is called *Transient* in continuous time if $\mathbb{P}(\{t \geq 0 : X_t = i\} \ is \ unbounded | X_0 = i) = 0$.

Theorem 4.3 -
A state i is recurrent for the continuous time Markov Chain X_t iff it is recurrent for the jump chain Y_n.

Proof 4.3 - *Theorem 4.3*
Suppose $X_0 = Y_0 = i$ and let $_n$ be the n^{th} jump time of X_t.
Suppose i is recurrent for Y_n so that $Y_n = i$ for infinitely many n with probability 1.
Hence $X_{T_n} = Y_n = i$ for infinitely many n, and the set $\{t : X_t = i\}$ is unbounded with probability 1.
So i is recurrent for X_t.

Conversely, suppose that i is transient for Y_n.
So $\mathbb{P}(Y_n = i \ infinitely \ often) = 0$ and with probability 1 there xists a maximal n st $Y_n = i$.
Therefore $X_t \neq i \ \forall \ t > T_{n+1}$.
Thus i is transient for X_t.

Theorem 4.4 - *Stationary Distribution & Generator*
If P_t has generator G then π is stationary iff $\pi G = 0$.

Proof 4.4 - *Theorem 4.4*
We have

$$\begin{array}{rcl} \pi \text{ is stationary} & \Leftrightarrow & \pi P_t = \pi \ forall \ t \geq 0 \\ & \Leftrightarrow & \frac{d}{dt}\pi P_t = \frac{d}{dt}\pi \\ & \Leftrightarrow & \pi P_t' = 0 \\ & \Leftrightarrow & \pi G P_t = 0 \text{ by backwards equations} \end{array}$$

But for $t = 0, \ P_0 = I \implies \pi G P_0 = \pi G = 0$.

Theorem 4.5 -
Let G be the generator of a chain whose jump chain has transition matrix H.
Then π is stationary for G iff $vH = v$ where $([v]_i = \pi_i g_{ii}$.

Proof 4.5 - *Theorem 4.5*

$$\begin{array}{rrcl} & vH & = & v \\ \implies & vH - vI & = & 0 \\ & [v(H-I)]_j & = & \sum_{i \in S} v_i(h_{ij} - \delta ij) \\ & & = & v_j(h_{jj} - \delta_{jj} \sum_{i \neq j} v_i(h_{ij} - 0) \\ & & = & (-\pi_j g_{jj})(-1) + \sum_{i \neq j} -\pi_i g_{ii} \left(\frac{-g_{ij}}{g_{ii}}\right) \\ & & = & \sum_{i \in S} \pi_i g_{ij} \\ & & = & [\pi G]_g \\ \implies & (\pi G) = 0 & \Leftrightarrow & vH = v \\ & & \Leftrightarrow & \pi \text{ stationary} \end{array}$$

Theorem 4.6 -
For an irreducible continuous time Markov chain

$$p_{ij} \xrightarrow{t\to\infty} \pi_j$$

5 Brownian Motion

5.1 Basic Notions

Remark 5.1 - *Brownian Motion Intuition*
Brownian Motion can be though of as the motion of a particle suspended in fluid, moving randomly but continuously about $\mathbb{R}^n$ with $n \in \mathbb{N}$.

Definition 5.1 - *Multivariate Normal Distribution*
An $n \times 1$ random vector $\boldsymbol{X} = (X_1, \dots, X_n)^t$ has *Multivariate Normal Distribution* if its joint density is

$$f(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^t\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right)$$

where Σ is a symmetric positive-definite matrix of size $n \times n$ and $\boldsymbol{\mu}$ is a $n \times 1$ vector.
N.B. - This distribution is denoted $N(\boldsymbol{\mu}, \Sigma)$.
N.B. - This is also known as *Multinormal Distribution*.

Remark 5.2 - *Positive-Definite Matrix*
Let A be a symmetric positive-definite matrix of size $n \times n$.
Then A has n eigenvalues which are all positive.
Furthermore, we can define $A = P^{-1}DP$ where D is a diagonal matrix of the eigenvalues & P is a unitary matrix.

Theorem 5.1 - *Properties of Multivariate Normal Distribution*
If $\boldsymbol{X} \sim N(\boldsymbol{\mu}, \Sigma)$ then

i) $\mathbb{E}(\boldsymbol{X}) = \mu$ (*i.e.* $\mathbb{E}(X_i) = \mu_i$;

ii) The i, j^{th} entry of Σ is the $Cov(X_i, X_j)$.

Theorem 5.2 - *Distribution of Linear Transformation of Multinormal Distribution*
Let $m \leq n$.
Define $\boldsymbol{X} \sim N(\boldsymbol{\mu}, \Sigma)$ to be a $n \times 1$ random vector, $\boldsymbol{c} \in \mathbb{R}^m$ & $B \in \mathbb{R}^{m \times n}$ with $Rank(B) = m$. Then

$$\boldsymbol{Y} = \boldsymbol{c} + B\boldsymbol{X} \implies \boldsymbol{Y} \sim N(\boldsymbol{c} + B\boldsymbol{\mu}, B\Sigma B^t)$$

Remark 5.3 - *MultiNormal Distribution of Independent Variables*
Let $\boldsymbol{X} = (X_1, \dots, X_n)^t$ with $X_1, \dots, X_n$ be independent with distributions $X_i \sim N(\mu_i, \sigma_i^2)$.
Then $\boldsymbol{X} \sim N(\boldsymbol{\mu}, Diag_n(\sigma_i^2))$ where $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)^t$ and $Diag_n(\sigma_i^2) = \begin{pmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n^2 \end{pmatrix}$.

Theorem 5.3 - *Central Limit Theorem*
Let $X_1, X_2, \dots$ be independent identically distribution random variables with $\mathbb{E}(X_i) = \mu$ & $Var(X_i) = \sigma^2 \neq 0$.
Define $Z_n = \dfrac{1}{\sigma\sqrt{n}} \sum_{i=1}^{n} (X_i - \mu)$.

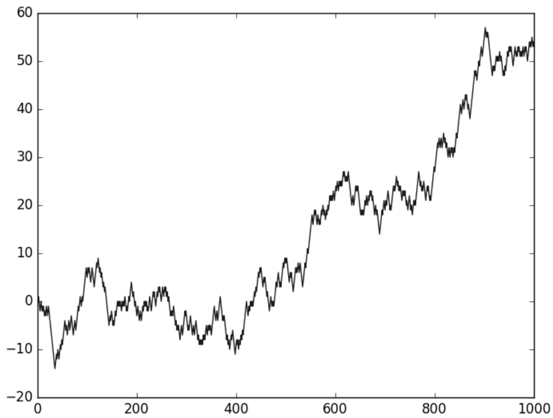
Then as $n \to \infty$ distribution of Z_n converges to $Z_n \sim N(0,1)$.

$$\lim_{n\to\infty} \mathbb{P}(|_n \leq x) = \Phi(x)$$

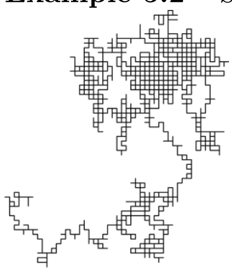
5.2 Definition & Construction of Brownian Motion

Definition 5.2 - *Simple Symmetric Random Walk, SSRW*
Let $Y_1, Y_2, \dots$ be independent random variables taking values of ± 1 with probability of $1/2$.
Then $S_n = \sum_{i=1}^n Y_i$ is known as a *Simple Symmetric Random Walk.*

Example 5.1 - *Simple Symmetric Random Walk in* $\mathbb{R}$



Example 5.2 - *Simple Symmetric Random Walk in* $\mathbb{R}^2$



Definition 5.3 - *Brownian Motion - 1D*
Let $\mathcal{F}_t$ be a filtration.
Brownian Motion is an adapted *stochastic process* $W = \{W_t\}_{t\geq 0}$ where

i) $W_0 = x$ for some $x \in \mathbb{R}$.

ii) W has independent and stationary Normal increments:

  (a) $W_{y+u} - W_t$ is independent of $\mathcal{F}_t \ \forall \ t, u \geq 0$;
  (b) $W_{y+u} - W_t$ & $W_{s+u} - W_s$ has the same distribution $\forall \ s, t, u \geq 0$ with $s \leq s+u \leq t+u$;
  (c) $W_{t+u} - W_t \sim N(0, u)$.

iii) W has continuous paths *i.e.* $t \mapsto W_t(\omega)$ is a continuous function of $t \ \forall \ \omega \in \Omega$.

Definition 5.4 - *Standard Brownian Motion*
$W = \{W_t\}_{t\geq 0}$ is said to be a *Standard Brownian Motion* if $W_0 = 0$.

Proposition 5.1 - *Distribution of Increments of General Brownian Motion*
Let W_t be *Brownian Motion* with $W_t = x$ then

$$(W_t | W_0 = x) \sim N(x, t)$$

Proposition 5.2 - *Distribution of Increments of Standard Brownian Motion*
Let W_t be *Standard Brownian Motion* then

$$W_t = W_t - W_0 \sim N(0,t)$$

Remark 5.4 - *Transition Density*
Brownian Motion has *Transition Density*

$$p(t,x,y) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2t}(x-y)^2}$$

where $\mathbb{P}(W_t \in (y, y+\Delta y)|W_0 = x) = p(t,x,y)\Delta y + o(\Delta y)$.
i.e - The probability that a *Brownian Motion* starting at x ends up in interval $(y, y+\Delta y)$ at time t is $\approx p(t,x,y)\Delta y$.

Proposition 5.3 - *Constructing the Brownian Motion of a Simple Symmetric Random Walk*
Let Y_t be a simple symmetric random walk & $S_n = \sum_{i=0}^n Y_i$, assuming $S_0 = 0$.
Each step Y_i has mean 0 & variance 1 thus by the central limit theorem $\frac{1}{\sqrt{n}} S_n$ converges to $N(0,1)$ distribution.
Compressing both time & space results in a process that converges to *Brownian Motion.*
Define $X^n = \{X_t^n\}_{t\in[0,1]}$ by setting $X_t^n = \frac{1}{\sqrt{n}} S_{nt}\ \forall\ t = \frac{j}{n}$ and use linear interpolation in between.
Then X^n converges to Brownian Motion.

Example 5.3 - *Expected Position of Brownian Motion*
Suppose a particle is at position 1.7 at time $t = 2$. What is the expected position at time $t = 4$?

$$\begin{array}{rcl}
\mathbb{E}(W_4|W_2 = 1.7) & = & \mathbb{E}(W_4 - W_2 + W_2 - W_0|W_2 = 1.7) \\
& = & \mathbb{E}(W_4 - W_2|W_2 = 1.7) + \mathbb{E}(W_2 - W_0|W_2 = 1.7) \\
& = & \mathbb{E}(W_4 - W_2) + 1.7 \\
& = & 0 + 1.7 = 1.7
\end{array}$$

Example 5.4 - *Probability of Position of Brownian Motion*
Suppose the price of a produce moves according to $X_t = \sigma W_t + \mu t$ with $\sigma^2 = 4$ & $\mu = -5$.
Given that $X_8 = 4$ what is the probability that the price is below 1 at time 9?

$$\begin{array}{rcl}
\mathbb{P}(X_9 < 1|X_8 = 4) & = & \mathbb{P}(X_9 - X_8 < -3|X_8 = 4) \\
& = & \mathbb{P}(X_9 - X_8 < -3 \\
& = & \mathbb{P}(X_1 = X_0 < -3) \\
& = & \mathbb{P}(X_1 < -3) \\
& = & \mathbb{P}(2W_1 - 5 \times 1 < -3) \\
& = & \mathbb{P}(W_1 < 1 \\
& = & \Phi(1) \\
& = & 0.8413
\end{array}$$

5.3 Properties of Brownian Motion

Proposition 5.4 - *Properties of Standard Brownian Motion*
Let W be a *Standard Brownian Motion* then

i) $\forall\ t \geq 0,\ \mathbb{E}(W_t) = 0$ & $Var(W_t) = t$;

ii) $\forall\ 0 \leq s \leq t,\ Cov(W_s, W_t) = s$;

iii) $-W_t$ is a *Standard Brownian Motion*;

iv) For a fixed $s > 0$ the process $X = \{X_t\}_{t\geq 0}$ defined by $X_t = W_{t+s} - W_s$ is also a *Standard Brownian Motion*;

v) For any $\alpha > 0$ the process $Y = \{Y_t\}_{t\geq 0}$ defined by $Y_t = \frac{1}{\sqrt{\alpha}}W_{\alpha t}$ is a *Standard Brownian Motion*.

N.B. - v) is known as the *Scaling Property*.

Proof 5.1 - *Proposition 5.4*

i) Since $W_t \sim N(0,t)$ then $\mathbb{E}(W_t) = 0$ & $Var(W_t) = t$;

ii) Let $0 \leq s \leq t$ then

$$\begin{aligned} Cov(W_s, W_t) &= Cov(W_s, W_t - W_s + W_s) \\ &= Cov(W_s, W_t - W_s) + Cov(W_s, W_s) \\ &= 0 + Var(W_s) \\ &= s \end{aligned}$$

iii) We check $-W_t$ has the properties of *Standard Brownian Motion* as defined in **Definition 5.3** & **Definition 5.4**

(a) $-W_0 = -0 = 0$;

(b) $-W_{t+u} - (-W_t) = -(W_{t+u} - W_t)$ we know $W_{t+u} - W_t \sim N(0,u)$.
So $-(W_{t+u} - W_t) \sim N(0,u)$ by symmetry of the normal distribution & it is independent of $\mathcal{F}_t$.

(c) $-W_t$ is continuous if W_t is continuous.

iv) We check $X_t = W_{t+s} - W_s$ has the properties of *Standard Brownian Motion* as defined in **Definition 5.3** & **Definition 5.4**

(a) $X_0 = W_s - W_s = 0$;

(b) $X_{t+u} - X_t = (W_{t+u+s} - W_s) + (W_{t+s} - W_s) = W_{t+u+s} - W_{t+s} \sim N(0,u)$.
This is independent of $\mathcal{F}_{t+s}$.

(c) X_t is the difference of two continuous functions, so is continuous.

v) We check $Y_t = \frac{1}{\sqrt{\alpha}}W_{\alpha t}$ has the properties of *Standard Brownian Motion* as defined in **Definition 5.3** & **Definition 5.4**

(a) $Y_0 = \frac{1}{\sqrt{\alpha}}W_0 = \frac{1}{\sqrt{\alpha}} \times 0 = 0$;

(b) $Y_{t+u} - Y_t = \frac{1}{\sqrt{\alpha}}(W_{\alpha(t+u)} - W_{\alpha t})$
We know $W_{\alpha t+\alpha u} - W_{\alpha t} \sim N(0, au)$ so $Y_{y+\alpha} - Y_t \sim N(0,u)$.
For $t_1 \leq t_2 \leq t_3 \leq t_4 \implies \alpha t_1 \leq \alpha t_2 \leq \alpha t_3 \leq \alpha t_4$.
$Y_{t_2} - Y_{t_1} = \frac{1}{\sqrt{\alpha}}(W_{\alpha t_2} - W_{\alpha t_1})$
$Y_{t_4} - Y_{t_3} = \frac{1}{\sqrt{\alpha}}(W_{\alpha t_4} - W_{\alpha t_4})$.
These are independent of each other as W_t is a *Brownian Motion* & the time gaps don't overlap.

(c) Y_t is continuous, since W_t is continuous.

Remark 5.5 - *Brownian Motion is not differentiable*

Definition 5.5 - *Gaussian Process*
A *Gaussian Process* is a Continuous-Time Stochastic Process with continuous sample paths & finite dimensional distributions that are multivariable normal.
A *Gaussian Process* is completely determined byits mean function & auto-covariance function.

Theorem 5.4 - *All States are Recurrent in Standard Brownian Motion*
For a *Standard Brownian Motion* W we have

$$\mathbb{P}\left(\sup_{t\geq 0} W_t = \infty, \inf_{t\geq 0} W_t = -\infty\right) = 1$$

N.B. This implies that *Brownian Motion* is recurrent.

5.4 The Reflection Principle & First Passage Time

Theorem 5.5 - *The Reflection Principle*
Let τ_a be the first passage time of a *Standard Brownian Motion*, W_t and define

$$\widetilde{W_t} = \begin{cases} W_T & for\ t \leq \tau \\ a - (W_t - a) = 2a - W_t & for\ t > \tau \end{cases}$$

Then $\widetilde{W_t}$ is also a *Standard Brownian Motion.*

Theorem 5.6 - *Density of First Passage Time*
The density of τ_a is given by

$$\frac{|a|}{\sqrt{2\pi t^3}} e^{-\frac{a^2}{2t}},\ t \geq 0$$

Proof 5.2 - *Theorem 5.6*
Let 0.
First, observe that $\{\tau_a \leq T\} = \{\sup_{0\leq t\leq T} W_t \geq a\}$.
Indeed, if $\tau_a \leq T$, this means the process hit a by time T.
This implies that $\sup_{0\leq t\leq T} W_t$ is at least a by continuity of the process.
On the other hand, if $\sup_{0\leq t\leq T} W_t \geq a$, this imples that the process hit a for the first time by T.
By continuity

$$\begin{array}{rcl}
\mathbb{P}(\tau_a \geq T) & = & \mathbb{P}(\sup_{0\leq t\leq T} W_t \geq a) \\
& = & \mathbb{P}(\sup_{0\leq t\leq T} W_t \geq a, W_t > a) + \mathbb{P}(\sup_{0\leq t\leq T} W_t \geq a, W_t < a) \\
& = & \mathbb{P}(\sup_{0\leq t\leq T} W_t \geq a, W_t > a) + \mathbb{P}(\sup_{0\leq t\leq T} \widetilde{W_t} \geq a, \widetilde{W_t} > a) \\
& = & \mathbb{P}(\sup_{0\leq t\leq T} W_t \geq a, W_t > a) + \mathbb{P}(\sup_{0\leq t\leq T} \widetilde{W_t} \geq a, W_t > a) \text{Since } \widetilde{W_t} \text{ is S. Brownian Motion} \\
& = & 2\mathbb{P}(\sup_{0\leq t\leq T} W_t \geq a, W_t > a) \\
& = & 2\mathbb{P}(W_t > a) \\
{}^{[1]} & = & 2\int_a^\infty \frac{1}{\sqrt{2\pi T}} e^{-\frac{1}{2t}x^2} dx \\
& = & \int_0^T \sqrt{\frac{a^2}{2\pi u^3}} e^{-\frac{a^2}{2u}} du \\
\implies f_{\tau_a}(T) & = & \sqrt{\frac{a^2}{2\pi T^3}} e^{-\frac{a^2}{2t}}
\end{array}$$

[1] consider the following substitution

$$\begin{aligned}
\text{Set} \quad u &= \frac{a^2T}{x^2} \\
\implies du &= -\frac{2a^2T}{x^3}dx \\
\& \quad x^6 &= \frac{a^6T^3}{u^3}
\end{aligned}$$

Theorem 5.7 - *For any fixed $a \neq 0$ we have $\mathbb{E}(\tau_a) = \infty$*

Proof 5.3 - *Theorem 5.7*
We only prove this for $a > 0$.
We compute $\mathbb{E}(\tau_a)$ using the standard trick of integrating the tail

$$\mathbb{E}(\tau_a) = \int_0^\infty \mathbb{P}(\tau_a > t)dt = \int_0^\infty \left(1 - \frac{2}{\sqrt{2\pi t}}\int_a^\infty ext(-x^2/2t)dx\right)dt$$

Using the formula for $\mathbb{P}(\tau_a > t)$ we have found in the previous truth.
But

$$\begin{aligned}
1 - \frac{2}{\sqrt{2\pi t}}\int_a^\infty ext(-x^2/2t)dx &= \frac{2}{\sqrt{2\pi t}}\int_0^a exp(-x^2/2t)dx \\
\text{Let } y = x/\sqrt{t} \implies &= \frac{2}{\sqrt{2\pi}}\int_0^{a/\sqrt{t}} exp(-y^2/2)dy
\end{aligned}$$

Plugging into the formula for $\mathbb{E}(\tau_a)$ we obtain

$$\begin{aligned}
\mathbb{E}(\tau_a) &= \frac{2}{\sqrt{2\pi}}\int_0^\infty \int_0^{a/\sqrt{t}} e^{-y^2/2}dy\ dt \\
&= \frac{2}{\sqrt{2\pi}}\int_0^\infty \int_0^{a^2/y^2} e^{-y^2/2}dy\ dy \\
&= \frac{2a^2}{\sqrt{2\pi}}\int_0^\infty \frac{1}{y^2}e^{-y^2/2}dy \\
&\geq \frac{2a^2}{\sqrt{2\pi}}\int_0^1 \frac{1}{y^2}e^{-y^2/2}dy \\
&\geq c\frac{2a^2}{\sqrt{2\pi}}\int_0^1 \frac{1}{y^2}dy \\
&= \infty
\end{aligned}$$

5.5 Martingales

Definition 5.6 - *Discrete-Time Martingale*
A stochastic process $Y = \{Y_n\}_{n\in\mathbb{N}}$ is called a *Discrete-Time Martingale* with respect to a filtration $\mathcal{F}_n$ if $\forall\ n \in \mathbb{N}$

i) $\mathbb{E}(|Y_n|) < \infty$;

ii) $\mathbb{E}(Y_{n+1}|\mathcal{F}_n) = Y_n$.

Definition 5.7 - *Continuous-Time Martingale*
A stochastic process $Y = \{Y_t\}_{t\geq 0}$ is called a *Continuous-Time Martingale* with respect to a filtration $\mathcal{F}_n$ if $\forall\ t \geq s \geq 0$

i) $\mathbb{E}(|Y_t|) < \infty$;

ii) $\mathbb{E}(Y_t|\mathcal{F}_s) = Y_s$.

Definition 5.8 - *Supermartingale*
$Y = \{Y_n\}_{n\in\mathbb{N}}$ is called a *Supermartingale* with respect to a filtration $\mathcal{F}_n$ if $\forall\ n$

i) $\mathbb{E}(|Y_n|) < \infty$;

ii) $\mathbb{E}(Y_{n+1}|\mathcal{F}_n) \leq Y_n$.

$Y = \{Y_t\}_{t\geq 0}$ is called a *Supermartingale* with respect to a filtration $\mathcal{F}_n$ if $\forall\ n$

i) $\mathbb{E}(|Y_n|) < \infty$;

ii) $\mathbb{E}(Y_{n+1}|\mathcal{F}_n) \geq Y_n$.

Remark 5.6 - *Common Filtrations*
Often $\mathcal{F}_n$ or $\mathcal{F}_t$ are taken to be the filtration generated by the process itself, $\{Y_n : n \in \mathbb{N}\}$ or $\{Y_s : 0 \leq s \leq t\}$.

Remark 5.7 - *Iterated Expectation of Martingale*
By the *Law of Iterated Expectation*

$$\mathbb{E}(Y_{n+1}) = \mathbb{E}(\mathbb{E}(Y_{n+1}|\mathcal{F}_n)) = \mathbb{E}(Y_n)\ \forall\ n \in \mathbb{N}$$

N.B. For *supermartingales* replace $=$ with $\leq$ & $\geq$ respectively.

Example 5.5 - *Simple Symmetric Random Walk is a Martingale*
Let $X_1, X_2, \dots$ be IID random variables with $\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = \frac{1}{2}$.
Fixing k let $Y_0 = k$ and for $n \geq 1$ define $Y_n = k + X_1 + \dots + X_n$.
Then $\{Y_n\}_{n\in\mathbb{N}}$ is a simple symmetric random walk which starts at k.
Let $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$.

$$\begin{array}{rcl}
\mathbb{E}(|Y_n|) & = & \mathbb{E}(|k + X_1 + \dots + X_n|) \\
& \leq & \mathbb{E}(|k|) + \mathbb{E}(|X_1|) + \dots + \mathbb{E}(|X_n|) \\
& = & |k| + n\mathbb{E}(X_1) \\
& = & |k| + n \\
& < & \infty\ \forall\ n \in \mathbb{N} \\
\mathbb{E}(Y_{n+1}|\mathcal{F}_n) & = & \mathbb{E}(Y_n + Y_{n+1}|\mathcal{F}_n) \\
& = & \mathbb{E}(Y_n|\mathcal{F}_n) + \mathbb{E}(X_{n+1}|\mathcal{F}_n) \\
& = & Y_n + \mathbb{E}(X_{n+1}) \\
& = & Y_n
\end{array}$$

Thus Y_n is a martingale.

Definition 5.9 - *Stopping Time of Filtration*
Let X be a stochastic process with the associated filtration $\mathcal{F}_n$ (or $\mathcal{F}_t$).
Then T is said to be a *Stopping Time of* $\mathcal{F}_n$ (or $\mathcal{F}_t$) if for every k (or s) the event $\{T \leq k\}$ (or $\{T \leq s\}$) if $\mathcal{F}_k$-measurable (or $\mathcal{F}_s$-measurable).
N.B. Stopped martingales are still martingales.

Theorem 5.8 - *Stopped Discrete-Time Martingale*
Let $Y = \{T_n\}_{n\in\mathbb{N}}$ be a *super-martingale* with respect to $\mathcal{F}_n$ and let T be a stopped time.
Then $Z = \{Z_n\}_{n\in\mathbb{N}}$ is defined as $Z_n = Y_{T\wedge n} = \begin{cases} Y_n & n \leq t \\ Y_T & n > t \end{cases}$.

Proof 5.4 - *Theorem 5.8*
Notice that $Y_n = Y_0 + \sum_{i=1}^n (Y_i - Y_{i-1})$ and $Z_n = Y_0 + \sum_{i=1}^n \mathbb{1}_{\{i \leq T\}}(Y_i - Y_{i-1})$. Then

$$\begin{array}{rcl}\mathbb{E}(Z_{n+1}|\mathcal{F}_n) - Z_n & = & \mathbb{E}(Z_{n+1} - Z_n|\mathcal{F}_n) \\ & = & \mathbb{E}(\mathbb{1}_{\{n+1 \leq T\}}(Y_{n+1} - Y_n)|\mathcal{F}_n) \\ & = & \mathbb{1}_{\{n<T\}}\mathbb{E}(Y_{n+1} - Y_n|\mathcal{F}_n)\end{array}$$

Thus Z_n is a supermartingale.

Theorem 5.9 - *Optional Stopping Theorem - Discrete Time*
Let $Y = \{Y_n)_{n\in\mathbb{N}}$ be a discrete-time martingale with respect to $\mathcal{F}_n$ and let T be a stopping time of $\mathcal{F}_n$.
If any of the following holds then $\mathbb{E}(Y_T) = \mathbb{E}(Y_0)$.

i) T is (almost surely) bounded - $\exists K \in \mathbb{R}^+ \ st\ \mathbb{P}(T < K) = 1$.

ii) T is (almost surely) finite and $\exists\ K > 0$ st $|Y_{T\wedge n}| < K\ \forall\ n \geq 0$.

iii) $\mathbb{E}(T) < \infty\ \&\ \exists\ K \in \mathbb{R}^+\ st\ |Y_n - Y_{n-1}| \leq K\ \forall\ n < T$.

iv) T is (almost surely) finite and $\mathbb{E}(|Y_T|) < \infty\ \&\ \mathbb{E}(Y_n\mathbb{1}_{\{T>n\}} \to 0$ as $n \to \infty$.

Theorem 5.10 - *Martingale Convergence Theorem - Discrete Time*
Let $Y = \{Y_n\}_{n\in\mathbb{N}}$ be a supermartingale wrt $\mathcal{F}_n$.
Suppose $\exists\ A > 0$ st $\mathbb{E}(|Y_n|) \leq A\ \forall\ n \in \mathbb{N}$.
Then $\exists$ a random variable Y_∞ st

$$\mathbb{P}\left(\lim_{n\in\infty} Y_n = Y_\infty\right) = 1$$

i.e For almost every ω, $lim_{n\to\infty}Y_n(\omega) = Y_\infty(\omega)$ or $\lim_{n\to\infty} Y + n = Y_\infty$.

Theorem 5.11 - *Existsnce of Z_∞*
If Z_n is a non-negative supermartingale wrt $\mathcal{F}_n$ then $\exists$ a random variable Z_∞ st $Z_n \to Z_\infty$.

Proof 5.5 - *Theorem 5.11*

$$\begin{array}{rcl}\mathbb{E}(|Z_n|) & = & \mathbb{E}(Z_n) \text{ as } Z_n \geq 0 \\ & = & \mathbb{E}(\mathbb{E}(Z_n|\mathcal{F}_{n-1})) \\ & \leq & \mathbb{E}(Z_{n-1}) \text{ since } Z_n\ mathrmis\ a\ supermartingale \\ & \leq & \mathbb{E}(Z_0) \\ & < & \infty \text{ as } \mathbb{E}(|Z_nZ) < \infty\ \forall\ n \text{ by definition of martingale}\end{array}$$

It follows from *Martingale Convergence Theorem* that $\exists\ Z_0$ st $Z_n \to Z_\infty$.

Remark 5.8 -
Standard Brownian Motion does not satisfy the assumption in the *Martingale Convergence Theorem*.
It may, or may not, satisfy the assumption in the *Optional Stopping Theorem*.

5.6 Application of Gambler's Ruin

Proposition 5.5 -
Let $X_1, X_2, \ldots$ be iid random variables st $\mathbb{P}(X_i = 1) = p\ \&\ \mathbb{P}(X_i = -1) = 1 - p =: q$.
Let $S_0 = k$ where $1 \leq k \leq N - 1$ and for $n \geq 1$ set $S_n = k + X_1 + \cdots + X_n$.
Thus $\{S_n\}_{n\in\mathbb{N}}$ is an unrestricted random walk starting at k.

Let $T = \min\{n : S_n = 0 \text{ or } S_n = N\}$ with T taken to be ∞ if $1 \leq S_n \leq N-1 \ \forall \ n \in \mathbb{N}$.
Let $Y_n = S_{T \wedge n}$.
Then $\{Y_n\}_{n \in \mathbb{N}}$ is a random walk with absorbing barriers at 0 and N.
In particular before stopping ($i.e$ $1 \leq i \leq N-1$)

$$\mathbb{P}(Y_{n+1} = i-1 | Y_n = i) = q \ \mathbb{P}(Y_{n+1} = i+1 | Y_n = i) = p$$

Also

$$\mathbb{P}(Y_{n+1} = 0 | Y_n = 0) = 1 \ \mathbb{P}(Y_{n+1} = N | Y_n = N) = 1$$

Let $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$.

We check that Y_n is a martingale wrt $\mathcal{F}_n$ if $p = q = \frac{1}{2}$.
Since $p = q = \frac{1}{2}$

$$\begin{array}{rrcl} \implies & \mathbb{E}(Y_{n+1} | Y_n = i) & = & i \ \forall \ i \in [1, N) \\ , & \mathbb{E}(Y_{n+1} | Y_n = 0) & = & 0 \\ \& & \mathbb{E}(Y_{n+1} | Y_n = N) & = & 1 \end{array}$$

Hence $\mathbb{E}(Y_{n+1} | \mathcal{F}_n) = \mathbb{E}(Y_{n+1} | Y_n) \ \forall \ n$ by *Markov Property* & $\mathbb{E}(|Y_n|) \leq N < \infty$.
So Y_n is a martingale.

Theorem 5.12 - *Properties of Gambler's Ruin Martingale*
Suppose $p = \frac{1}{2} = q$. then

i) $\mathbb{P}(T < \infty) = 1$

ii) $\mathbb{P}(Y_T = N) = \frac{Y_0}{N} = \frac{k}{N}$; and,

iii) $\mathbb{E}(T) = Y_0(N - Y_0) = k(N-k)$.

Proof 5.6 - *Theorem 5.12*

i) We know that $Y_n = S_{T \wedge n}$ is a Martingale & $Y_n \in [0, N] \ \forall \ n \in \mathbb{N}$.
Hence $\mathbb{E}(|Y_n|) < N+1$ so by *Martingale Converge Theory* $\exists Y_\infty$ st $Y_n \to Y_\infty$ almost surely.
if $T(\omega) = \infty$ then $Y_n(\omega) \in [1, N-1] \ \forall \ n \in \mathbb{N}$ and $|Y_{n+1}(\omega) - Y_n(\omega)| = 1$.
This violates the *Cauchy Criterion* for convergence, so the process Y_n does not converge.
So $\mathbb{P}(T = \infty) \leq \mathbb{P}(Y_n \text{ does not converge}) = 0$ by the *Martingale Converge Theorem*.
$\implies \mathbb{P}(T < \infty) = 1$.

ii) We know that $T < \infty$ as $|Y_n < N+1 \ \forall \ n > 0$.
Hence assumption *ii)* of the *Optional Stopping Theorem* holds ($expect(Y_t) = \mathbb{E}(Y_0) = k$.
$\implies 0.\mathbb{P}(Y_t = 0) + N.\mathbb{P}(Y_t = N) = k$
$\implies \mathbb{P}(Y_T = N) = \frac{k}{N}$.

iii) Define $Z_n := S_n^2 - n$. We have already proved this to be a Martingale.
We have

$$\begin{array}{rcl} \mathbb{E}(T) & = & \sum_{k=1}^{n} \mathbb{P}(T > k) \\ & = & \sum_{r=0}^{\infty} (\mathbb{P}(T > rN) + \mathbb{P}(T > rN12) + \dots + \mathbb{P}(T > rN + n)) \\ & \leq & \sum_{r=0}^{\infty} n.\mathbb{P}(T > rN) \end{array}$$

$$\begin{array}{rrcl}
\text{For } r=1 & \mathbb{P}(T>N) & = & \mathbb{P}(S_n\in[1,N-1],n\leq N)\\
& & = & 1-\mathbb{P}(T\leq N)\\
& & \leq & 1-\mathbb{P}(A)\text{ where } A=\{-1,+1,\dots,-1\}\subset\{T\leq N\}\\
& & = & 1-\frac{1}{2^N}\\
\text{For } r=2 & \mathbb{P}(T>2N) & = & \mathbb{P}(S_n\in[1,n-1],n\leq 2N)\\
& & = & \mathbb{P}\big(S_n\in[1,N-1],n\in[1,N]\big)\\
& & \times & \mathbb{P}\big(S_n\in[1,N-1],n\in[N+1,2N]\big|S_n\in[1,N-1]\big)\\
& & \leq & \mathbb{P}(T>N)\mathbb{P}(T'>N)\\
& & = & \big(1-\frac{1}{2^N}\big)^2\\
\text{So} & \mathbb{P}(T>rN) & \leq & \big(1-\frac{1}{2^N}\big)^r\ \forall\ r\in[1,N]
\end{array}$$

Since $\big|1-\frac{1}{2^N}\big|<1 \implies \mathbb{E}(T)<\infty$.
Since T is bounded, by the first assumption, the *Optional Stopping Theorem* holds. Thus

$$\begin{array}{rrcl}
& \mathbb{E}(Z_T) & = & \mathbb{E}(Z_0)=S_0^2=k^2\\
\implies & k^2 & = & \mathbb{E}(S_T^2-T)\\
& & = & \mathbb{E}(S_T^2)-\mathbb{E}(T)\\
& & = & 0^2.\mathbb{P}(S_T=0)+N^2.\mathbb{P}(S_T=N)-\mathbb{E}(T)\\
& & = & N^2.\frac{k}{N}-\mathbb{E}(T)\\
\implies & \mathbb{E}(T) & = & k(N-k)\\
\& & |Z_n-Z_{n-1}| & = & |S_n^2-S_{n-1}^2+1|\\
& & \leq & |S_n^2|+|S_{n-1}^2|+1\\
& & \leq & 2N^2+1\ \forall\ n
\end{array}$$

Theorem 5.13 - *Special Martingale*
Let $Y=\{Y_n\}_{n\in\mathbb{N}}$ defined by $Y_n=k+X_0+\dots+X_n$, be the absorbed random walk on $[0,N]$ with $p\neq q$, $\mathcal{F}_n=\sigma(X_1,\dots,X_n)$.
Let $V_n=(q/p)^{Y_n}$, then $\{V_n\}_{n\in\mathbb{N}}$ is a *Martingale* wrt $\mathcal{F}_n$.

Proof 5.7 - *Theorem 5.13*
Since $Y_n\in[0,N]$ then

$$\mathbb{E}(|V_n|)=\mathbb{E}(V_n)=\mathbb{E}\left(\left(\frac{q}{p}\right)^{Y_n}\right)\leq\max\left\{\left(\frac{q}{p}\right)^0,\left(\frac{q}{p}\right)^N\right\}<\infty$$

$$\begin{array}{rcl}
\mathbb{E}(V_{n+1}|\mathcal{F}_n) & = & \mathbb{E}\left(\left(\frac{q}{p}\right)^{Y_{n+1}}|\mathcal{F}_n\right)\\
& = & \mathbb{E}\left(\left(\frac{q}{p}\right)^{Y_n+1}|Y_n\right)\text{ by Markov Property}\\
\mathbb{E}(V_{n+1}|Y_n=0) & = & \left(\frac{q}{p}\right)^0=\left(\frac{q}{p}\right)^{Y_1}=V_n\\
\mathbb{E}(V_{n+1}|Y_n=N) & = & \left(\frac{q}{p}\right)^N=\left(\frac{q}{p}\right)^{Y_n}=V_n\\
\text{For } i\in[1,N-1] & & \\
\mathbb{E}\left(\left(\frac{q}{p}\right)^{Y_{n+1}}|Y_n=i\right) & = & \left(\frac{q}{p}\right)^{i=1}p+\left(\frac{q}{p}\right)^{i-1}q\\
& = & \left(\frac{q}{p}\right)^i[q+p]\\
& = & \left(\frac{q}{p}\right)^i\times 1\\
& = & V_n
\end{array}$$

V_n is a martingale.

Theorem 5.14 - *Properties from Theorem 5.13*
Let Y be as in previous lemma. Let T be the first time Y hits 0 or N. Then

i) $\mathbb{P}(T < \infty) = 1$; and,

ii) $\mathbb{P}(Y_T = N) = \dfrac{1-(q/p)^k}{1-(q/p)^N}$

Proof 5.8 - *Theorem 5.14*

i) Let $V_n = V_n = \left(\frac{q}{p}\right)^{Y_n}$.
Then V_n is a martingale & $V_n \leq \max\left\{\left(\frac{q}{p}\right)^0, \left(\frac{q}{p}\right)^N\right\}$.
By the *Martingale Convergence Theorem* we know V_n converges, almost surely, to some V_∞.
Now, $\{T = \infty\} = \left\{V_n \subset \left\{\left(\frac{q}{p}\right)^1, \dots, \left(\frac{q}{p}\right)^{N-1}\right\}\ \forall\ n > 0\right\}$.
So $\exists\ \delta > 0$ st $|V_{n+1} - V_n| > \delta\ \forall\ n$ which implies that V_n does not converge.
So $P(T = \infty) = \mathbb{P}\left(V_n \subset \left\{\left(\frac{q}{p}\right)^1, \dots, \left(\frac{q}{p}\right)^{N-1}\right\}\ \forall\ n > 0\right) \leq \mathbb{P}(V_n \text{ does not converge}) = 0$
by *Martingale Convergence Theorem.*

ii) Since $T < \infty$ almost surely, by 1), and $\forall\ n \leq \max\left\{\left(\frac{q}{p}\right)^0, \left(\frac{q}{p}\right)^N\right\}$ then condition *iii)* of *Optional Stopping Distance* holds.

$$\begin{array}{rrcl}
\implies & \mathbb{E}(V_t) & = & \mathbb{E}(V_0) = \left(\frac{q}{p}\right)^k \\
\implies & \left(\frac{q}{p}\right)^0 \mathbb{P}(Y_t = 0) + \left(\frac{q}{p}\right)^N \mathbb{P}(Y_T = N) & = & \left(\frac{q}{p}\right)^k \\
\implies & 1 \times (1 - \mathbb{P}(Y_T = N)) + \left(\frac{q}{p}\right)^N \mathbb{P}(Y_T = N) & = & \\
\implies & \mathbb{P}(Y_T = n) & = & \dfrac{1 - \left(\frac{q}{p}\right)^k}{1 - \left(\frac{q}{p}\right)^N}
\end{array}$$

5.7 Brownian Motion as a Martingale

Theorem 5.15 - *Standard Brownian Motion as a Martingale*
Let $\{W_t\}_{t\geq 0}$ is *Standard Brownian Motion* with filtration $\mathcal{F}_t$ generated by the process itself, then

1) W_t is a martingale.

2) $W_t^2 - t$ is a martingale.

3) $X_t = at + \sigma W_t$ is a martingale iff $a = 0$.

4) $Y_t = e^{at+\sigma W_t}$ is a martingale iff $a = -\frac{\sigma^2}{2}$.

Proof 5.9 - *Theorem 5.15*

3) Let $X_t = at + \sigma W_t$.

$$\begin{aligned}
\mathbb{E}(|X_t|) &= \mathbb{E}(|at+\sigma W_t|)\\
&\leq \mathbb{E}(|at|)+\mathbb{E}(|\sigma W_t|)\\
&= |at|+\sigma\mathbb{E}(|W_t|)\\
&\leq t|a|+\sigma\mathbb{E}(W_t^2)^{1/2}\\
&= t|a|+\sigma t\\
&< \infty\ \forall\ t>0\\
\text{Let } 0\leq s\leq t &\\
\mathbb{E}(X_t|\mathcal{F}_s) &= \mathbb{E}(at+\sigma W_t|\mathcal{F}_s)\\
&= at+\sigma\mathbb{E}(W_t-W_s+W_s|\mathcal{F}_s)\\
&= at+\sigma\mathbb{E}(W_t-W_s)+\sigma W_s\\
&= at+\sigma\times 0+\sigma W_s\\
&= at+\sigma W_s
\end{aligned}$$

We want $\mathbb{E}(X_t|\mathcal{F}_s)=X_s$.Set

$$\begin{aligned}
as+\sigma W_s &= at+\sigma W_s\\
\implies a(s-t) &= 0\\
\implies a &= 0
\end{aligned}$$

4) Let $Y_t = e^{at+\sigma W_t}$.

$$\begin{aligned}
\mathbb{E}(|Y_t|)=\mathbb{E}(e^{at+\sigma W_t}) &\\
&= e^{at}\mathbb{E}(e^{\sigma W_t})\\
&= e^{at}e^{0\times\sigma+\frac{1}{2}t^2\sigma^2}\text{ by moment}-\text{generating funciton}\\
&= e^{at+\frac{1}{2}\sigma^2t^2}\\
&< \infty\ \forall\ t>0\\
\text{Let } 0\leq s\leq t &\\
\mathbb{E}(Y_t|\mathcal{F}_s) &= \mathbb{E}(e^{at+\sigma W_t}|\mathcal{F}_s)\\
&= e^{at}\mathbb{E}(e^{\sigma W_t}|\mathcal{F}_s)\\
&= e^{at}\mathbb{E}(e^{\sigma(W_t-W_s+W_s)}|\mathcal{F}_s)\\
&= e^{at+\sigma W_s}\mathbb{E}(e^{\sigma(W_t-W_s)})\\
&= e^{at+\sigma W_s}e^{0+\frac{1}{2}\sigma^2(t-s)}\\
&= e^{at+\sigma W_s+\frac{1}{2}\sigma^2(t-s)}
\end{aligned}$$

Setting $\mathbb{E}(Y_t|\mathcal{F}_s)$ yields that $a=-\frac{\sigma^2}{2}$.

Proposition 5.6 - *Stopping Time of Brownian Motion*
Let $a,b>0$ and $\tau=\min\{t\geq 0: W_t\in\{a,-b\}\}$. Then

$$\mathbb{P}(W_\tau=a)=\frac{b}{a+b}\text{ and }\mathbb{E}(\tau)=ab$$

Proof 5.10 - *Proposition 5.6*
We know that $\mathbb{E}(\tau)<\infty$, almost surely.
If we define $Z:=\max\{a,b\}$ then $|W_{t\wedge\tau}|\leq Z$.
Then the *Optional Stopping Theorem* holds (*i.e.* $\mathbb{E}(W_\tau)=\mathbb{E}(W_0)=0$.

$$\begin{aligned}
\implies a\mathbb{P}(W_\tau=a)+(-b)\mathbb{P}(W_\tau=-b) &= 0\\
\implies a\mathbb{P}(W_\tau=a)+(-b)(1-\mathbb{P}(W_\tau=a)) &= 0\\
\implies \mathbb{P}(W_\tau=a) &= \frac{b}{a+b}
\end{aligned}$$

Proposition 5.7 - *Which Absorbing Barrier is Hit*
Let $X_t = \mu t + \sigma W_t$ with $\mu < 0$ & $M = \max\{X_t : t \geq 0\}$.
For $a, b > 0$
$$\mathbb{P}(\tau_a > \tau_{-b}) = \frac{1 - e^{-\alpha b}}{e^{\alpha a} - e^{-\alpha b}} \text{ with } \alpha = -\frac{2\mu}{\sigma^2}$$
Thus
$$\mathbb{P}(M \geq a) = e^{-\alpha a}$$

Proof 5.11 - *Proposition 5.7*
First we prove that $e^{\alpha X_t}$ is a *Martingale*.
$$\begin{array}{rcl}
\mathbb{E}(|e^{alphaX_t}|) & = & \mathbb{E}(e^{\alpha X_t}) \\
& = & \mathbb{E}(e^{\alpha\mu t + \sigma\alpha W_t}) \\
& = & e^{\alpha\mu t}\mathbb{E}(e^{\sigma\alpha W_t}) \\
& = & e^{\alpha\mu t}e^{0+\frac{1}{2}\sigma^2\alpha^2 t} \\
& < & \infty\ \forall\ t > 0 \\
\text{Let } 0 \leq s < t & & \\
\mathbb{E}(e^{\alpha X_t}|\mathcal{F}_s) & = & \mathbb{E}(e^{\alpha\mu t + \alpha\sigma W_t}|\mathcal{F}_s) \\
& = & e^{\alpha\mu t}\mathbb{E})e^{\alpha\sigma(W_t - W_s + W_s)}|\mathcal{F}_s) \\
& = & e^{\alpha\mu t + \alpha\sigma W_s}\mathbb{E}(e^{\alpha\sigma(W_t - W_s)}) \\
& = & e^{\alpha\mu t + \alpha\sigma W_s}e^{\frac{1}{2}\alpha^2\sigma^2(t-s)} \\
\text{Setting } \alpha = -\frac{2\mu}{\sigma^2} & & \\
\mathbb{E}(e^{\alpha W_t}|\mathcal{F}_s) & = & e^{-2\frac{\mu^2 t}{\sigma^2} - 2\frac{\mu W_s}{\sigma} + \left(4\frac{\mu^2}{\sigma^2}\right)\left(\frac{t-s}{2}\right)} \\
& = & e^{-2\frac{\mu}{\sigma^2}(\mu s + \sigma W_s)} \\
& = & e^{\alpha X_s}
\end{array}$$

0 Reference

0.1 Notation

Notation 0.1 - *Collection of Events*
A collection of events are denoted by $\mathcal{C}$.

Notation 0.2 - *Infimum*
Let S be a subset of an ordered set T.
The *Infimum* of S is the greatest element in T that is less of equal to all elements of S.
This is denoted by
$$\inf(A)$$

Notation 0.3 - *Minimum*

$$x \wedge y := \min(x, y)$$

Notation 0.4 - *Poisson Process*
A *Poisson Process* is denoted by $\{N_t\}_{t\geq 0}$.
Individual events in this sequence are denoted by N_i.

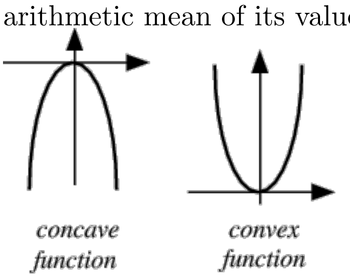
Notation 0.5 - *Power Set*
The power set of set S is denoted by 2^S *or* $\{0,1\}^S$.

Notation 0.6 - *Sample Space*
The *Sample Space* of a variable is denoted by Ω.

Notation 0.7 - *Transition*
A transition between state x & y is generally denoted by p_{xy}.

0.2 Definitions

Definition 0.1 - *Convex & Concave Functions*
A *Convex Function* is a continuous function whose value at the midpoint of every interval in its domain does not exceed the arithmetic mean of its values at the ends of the interval.
A *Concave Function* whose value at the midpoint of every interval in its domain exceeds teh arithmetic mean of its values at the ends of the interval.



Definition 0.2 - *Co-Variance*
Co-Variance is a measure of the joint variability of two random variables.
A greater magnitude of *Co-Variance* corresponds to the two variables having similar behaviour.
A positive *Co-Variance* means that as one variable increases when the other tends to.
A negative *Co-Variance* means that as one variable decreases when the other tends to increase.

$$Cov(X, Y) = \mathbb{E}\left((X - \mathbb{E}(X)(Y - \mathbb{E}(Y)\right)$$

Definition 0.3 - *Index Set*
An *Index Set* is a set whose members are used to index members of another set.

Definition 0.4 - *Indicator Function*
The *Indicator Function* of an event returns 1 or 0 to denote whether a given event event occurs

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

Definition 0.5 - *Moment Generating Function*
For random variable X with probability mass function $p_X(k)$ has a *Moment Generating Function*

$$m_X(\theta) = \mathbb{E}(e^{\theta X}) = \sum_k p_x(k)e^{\theta k}$$

Definition 0.6 - *Probability Generating Function*
For random variable X with probability mass function $p_X(k)$ has a *Probability Generating Function*

$$P_X(s) = \mathbb{E}(s^X) = \sum_K p_X(k)s^k$$

Definition 0.7 - *Right-Continuous Function*
A *Right-Continuous Function* is one in which no jump occurs when the limit point is approached from the right hand size.

Definition 0.8 - *Unitary Matrix*
P is a *Unitary Matrix* if $P^{-1} = P^*$ where P^* is the hermitian matrix of P.

0.3 Theorems

Theorem 0.1 - *Cauchy Criterion for Convergence*
A sequence $\{a_n\}$ converges iff

$$\forall\, \varepsilon > 0\ \exists N \in \mathbb{N}\ st\ \forall\ m, n \in \mathbb{N} \text{ with } m, n > N\ |a_m - a_n| < \varepsilon$$

Theorem 0.2 - *Cauchy-Schwarz Inequality*
Let X & Y be random variables with finite variance, then

$$\mathbb{E}(|XY|) \leq \sqrt{\mathbb{E}(X^2)\mathbb{E}(Y^2)}$$

Theorem 0.3 - *Conditional Probability*

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \bigcap B)}{\mathbb{P}(B)}$$

Theorem 0.4 - *Covariance Identities*
The following are identities concerning the *Covariance*

$$\begin{array}{rcl} Cov(X+Y) & = & \mathbb{E}(XY) - \mathbb{E}(X) + \mathbb{E}(Y) \\ Cov(X,a) & = & 0 \\ Cov(X,aY) & = & aCov(X,Y) \\ Cov(X,Y+Z) & = & Cov(X,Y) + Cov(X,Z) \end{array}$$

Theorem 0.5 - *Expectation of Expectation of Conditional*

$$\mathbb{E}(\mathbb{E}(Y|X)) = \mathbb{E}(Y)$$

Theorem 0.6 - *Expected Value of Indicator Function*

$$\mathbb{E}(1_A) = \mathbb{P}(A)$$

Theorem 0.7 - *Jensen's Inequality*
If g is a convex function, then

$$\mathbb{E}(g(X)) \geq g(\mathbb{E}(X))$$

If g is a concave function, then

$$\mathbb{E}(g(X)) \leq g(\mathbb{E}(X))$$

Theorem 0.8 - *L'Hôpitals Rule*
If $\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{0}{0}or= \frac{\infty}{\infty}$ then

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}$$

Theorem 0.9 - *Markov's Inequality*
Let X be a non-negative random variable, then $\forall c > 0$

$$\mathbb{P}(X > c) \leq \frac{\mathbb{E}(X)}{c}$$

Theorem 0.10 - *Probability of Event as Integral*

$$\mathbb{P}(x \in A) = \int_y \mathbb{P}(x \in A | Y = y) F_Y(y) dy$$

Theorem 0.11 - *Stirling's Formula*
For $n \in \mathbb{N}$

$$n! \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \text{ as } n \to \infty$$

Theorem 0.12 - *Sum of Exponentials*
Let $X_1, \dots, N_n \sim Exp(\lambda)$. Then

$$X_1 + \cdots + X_n \sim \Gamma(b, \lambda)$$

0.4 Probability Distributions

Definition 0.9 - *Binomial Distribution*
Let X be a discrete random variable modelled by a *Binomial Distribution* with n events and rate of success p.

$$\begin{array}{rcl} p_X(k) & = & \binom{n}{k} p^k (1-p)^{n-k} \\ \mathbb{E}(X)np = & \& & Var(X) = np(1-p) \end{array}$$

Definition 0.10 - *Gamma Distribution*
Let T be a continuous randmo variable modelled by a *Gamma Distribution* with shape parameter α & scale parameter λ. Then

$$\begin{array}{rcl} f_T(x) & = & \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} \quad \text{for } x > 0 \\ \mathbb{E}(T) = \frac{\alpha}{\lambda} & \& & Var(T) = \frac{\alpha}{\lambda^2} \end{array}$$

N.B. $\alpha, \lambda > 0$.

Definition 0.11 - *Exponential Distribution*
Let T be a continuous random variable modelled by a *Exponential Distribution* with parameter λ. Then

$$\begin{array}{rcll} f_T(t) & = & \lambda e^{-\lambda t} & \text{for } t > 0 \\ F_T(t) & = & 1 - e^{-\lambda t} & \text{for } t > 0 \\ \mathbb{E}(X) = \frac{1}{\lambda} & \& & Var(X) = \frac{1}{\lambda^2} & \end{array}$$

N.B. Exponential Distribution is used to model the wait time between decays of a radioactive source.

Definition 0.12 - *Normal Distribution*
Let X be a continuous random variable modelled by a *Normal Distribution* with mean μ & variance σ^2.
Then

$$\begin{array}{rcl} f_X(x) & = & \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \\ F_X(x) & = & \frac{1}{\sqrt{2\pi\sigma^2}} \int\limits_{-\infty}^{x} e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy \\ M_X(\theta) & = & e^{\mu\theta + \sigma^2\theta^2(1/2)} \\ \mathbb{E}(X) = \mu & \& & Var(X) = \sigma^2 \end{array}$$

Definition 0.13 - *Poisson Distribution*
Let X be a discrete random variable modelled by a *Poisson Distribution* with parameter λ. Then

$$\begin{array}{rcll} p_X(k) & = & \frac{e^{-\lambda}\lambda^k}{k!} & \text{For } k \in \mathbb{N}_0 \\ \mathbb{E}(X) = \lambda & \& & Var(X) = \lambda & \end{array}$$

N.B. Poisson Distribution is used to model the number of radioactive decays in a time period.