Probability 2 - Notes

Dom Hutchinson

May 6, 2019

Contents

1 General

Definition 1.1 - *σ-Field, $\mathcal{F}$*
A *σ-Field* is a collection of subsets of the *Sample Space* which can be used to establish a formal definition of the probability distribution of the *Sample Space*. $\mathcal{F}$ is a *σ-Field* if

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

ii) $\forall \ \{A_1, A_n\} \subseteq \mathcal{F}, \ \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$; and,

iii) $\forall \ A \in \mathcal{F}, \ A^c \in \mathcal{F}$.

The events, $A_i \in \mathcal{F}$, are said to be *$\mathcal{F}$-Measurable*. $\emptyset \in \mathcal{F}$ is known as the *Impossible Event* & $\Omega \in \mathcal{F}$ is known as the *Certain Event*. For a collection of events, $\mathcal{C}$, $\sigma(\mathcal{C})$ is the smallest *σ-Field* that contains $\mathcal{C}$.

Theorem 1.1 - *Properties of σ-Fields*

i) $\forall$ σ-fields $\mathcal{F}_1, \mathcal{F}_2, \ \mathcal{F}_1 \cap \mathcal{F}_2$ is a σ-field.

ii) 2^{Ω} is a σ-field.

Definition 1.2 - *Probability Measure, $\mathbb{P}$*
A *Probability Measure* is a function $\mathbb{P} : \mathcal{F} \to [0, 1]$ which satisfies

i) $\mathbb{P}(\emptyset) = 0$ & $\mathbb{P}(\Omega) = 1$; and,

ii) If $A_1, \dots, A_n \in \mathcal{F}$ are pair-wise disjoint then $\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i)$. ($\sigma$-Additivity)

Definition 1.3 - *Probability Space*
A *Probability Space* is a triple formed of a *Sample Space*, Ω; a *σ-Field*, $\mathcal{F}$, on Ω; and, a *Probability Measure*, $\mathbb{P}$, on $\mathcal{F}$. Denoted $(\Omega, \mathcal{F}, \mathbb{P})$.

Definition 1.4 - *Random Variable*
A *Random Variable*, X on a *Probability Space* $(\Omega, \mathcal{F}, \mathbb{P})$ is a function $X : \Omega \to \mathbb{R}$.

Definition 1.5 - *Filtration, $\mathcal{F}_t$*
A *Filtration* is a family of *σ-Fields*, $\mathcal{F}_t = \{\mathcal{F}_t : t \geq 0\}$ st $\mathcal{F}_{t_1} \subset \mathcal{F}_{t_2} \ \forall \ 0 \leq t_1 \leq t_2$.

Definition 1.6 - *Stochastic Process*
A *Stochastic Process* is a collection of random variables which represent the state of a system at different times (*e.g.* $X = \{X_t\}_{t \in \Delta}$ with $\Delta \subseteq \mathbb{R}$ where X_t is the state of the system at time t). The *Stochastic Process* has an associated *Filtration*, $\mathcal{F}_t$, st X_t is *$\mathcal{F}_t$-Measureable* (X is *Adapted* to $\mathcal{F}_t$). The *State Space* of a *Stochastic Process* is the set of all possible values at a specific time.
N.B. -*Stochastic Processes* can be in discrete or continuous time.

Definition 1.7 - *Time-Homogenous*
A *Stochastic Process*, X, is *Time-Homogenous* if $\mathbb{P}(X_{n+1} = j | X_n = i)$ depends on i & j, but not on n.

Definition 1.8 - *Sequences*
Let $X = A_1, A_2, \dots$ be a sequence of events. If $A_n \subseteq A_{n+1} \ \forall \ n \in \mathbb{N}$ then X (*i.e.* A_n occurs $\Rightarrow A_{n+1}$ occurs) is an *Increasing Sequence*. If $A_{n+1} \subseteq A_n \ \forall \ n \in \mathbb{N}$ then X is a *Decreasing Sequence*.

Theorem 1.2 - *Continuity of Probability*
Let $A_1, A_2, \dots$ be an *Increasing Sequence* of events. Define $A := \bigcup_{n=1}^{\infty} A_n$. Then

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

Let $B_1, B_2, \dots$ be a *Decreasing Sequence* of events. Define $B := \bigcap_{n=1}^{\infty} B_n$. Then

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

2 Markov Chains

Definition 2.1 - *Markov Property*
A *Stochastic Process* $X := \{X_n\}_{n\in\mathbb{N}}$ or $X := \{X_t\}_{t\in\mathbb{R}}$ has the *Markov Property* if values only depend on the value immediately before them (in time) and nothing earlier. Equivalently

$$\begin{array}{rcl} X_{n+1} & = & f(X_n) \\ \mathbb{P}(X_{n+1} = x_{n+1}|\mathcal{F}_n) & = & \mathbb{P}(X_{n+1} = x_{n+1}|X_n = x_n) \\ \mathbb{P}(X_{n+1} = x_{n+1}|X_n = x_n, \dots, X_0 = x_0) & = & \mathbb{P}(X_{n+1} = x_{n+1}|X_n = x_n) \\ \mathbb{P}(X_t = j|\mathcal{F}_s) & = & \mathbb{P}(X_t = j|X_s) \end{array}$$

Definition 2.2 - *Markov Chain*
A *Markov Chain* is a state space process with the *Markov Property.*

Definition 2.3 - *Transition Matrix*
A *Transition Matrix* are a *Stochastic Matrix* which describes the probabiltiy of transitioning between each state. For a *Continuous Time* process we define different *Transition Matrices* for each time t.

Discrete	$[P]_{ij} := p_{ij}$
Continuous	$[P_t]_{ij} := p_{ij}(t) = \mathbb{P}(X_t = j\|X_0 = i)$.

Definition 2.4 - *Stationary Distributions*
Let $X := \{X_n\}_{n\in\mathbb{N}}$ (or continuous) be a *Markov Chain*, on state space S, with transition matrix P . Let π be a <u>horizontal vector</u>. π is a *Statiomnary Distribution* of X if

i) $\pi_j \geq 0\ \forall\ j \in S$;

ii) $\sum_{j\in S} \pi_j = 1$;

iii) $\pi = \pi P_t$ or $\pi = \pi P \Leftrightarrow \pi_j = \sum_{i\in S} \pi_i p_{ij}\ \forall\ j \in S$

An *Irreducible Markov Chain* on S has a *Stationary Distribution* iff all the states in S are positive reccurent. In this case $\pi_i = \frac{1}{m_{ii}}$ (For continuous time this is equivalent to $p_{ij}(t) \xrightarrow{t\to\infty} \pi_j$).
<u>N.B.</u> -π is a stationary distribution for G iff $\boldsymbol{\nu} H = \boldsymbol{\nu}$ where H is transition matrix for jump chain & $\nu_i = -\pi_i g_{ii}$.

2.1 Discrete Time

Definition 2.5 - *Communication*
A state $i \in S$ *Communicates* with state $j \in S$ $(i \to j)$ if $\exists\ n \in \mathbb{N}$ *st* $p_{ij}(n) > 0$. State $i, j \in S$ *Intercommunicate* if $i \to j$ & $j \to i$. If two states *Intercommunicate* and one is *Recurrent*, then both are *Recurrent*. *Intercommunication* is an *Equivalence Relation*. The state space can be partitioned into *Communicating Classes* $\{E_1, \dots, E_n\}$ st $\forall\ i, j \in E_k\ i \leftrightarrow j$.

Definition 2.6 - *Closed & Irreducible*
Let $C \subseteq S$ be a set of states. C is *Closed* if $p_{ij} = \ \ \forall\ i \in C,\ j \notin C$. C is *Irreducible* if

$i \leftrightarrow k \; \forall \; i, j \in C$. A *Closed, Singleton Set* $\{i\}$ is called an *Absorbing State*. If C is a *Communicating Class* and is not clossed then all states in C are *Transient*.

Remark 2.1 - *Partitioning State Space*
The state space can be uniquely partition st $S = T \cup C_1 \cup \cdots \cup C_n$ where T is a set of *Transient* states & each C_i is an *Irreducible-Closed* set of *Recurrent* States.

Definition 2.7 - *Period*
Let $j \in S$ be a state st $p_{jj}(n) > 0, \; n \in \mathbb{N}$. Define $\mathcal{N}_j := \{n \geq 1 : p_{jj}(n) > 0\}$. The *period* of j is defined as $d_j := \gcd(\mathcal{N}_j)$, if $d_j = 1$ then j is said to be *Aperiodic*. For a *Communicating Class* $C \subseteq S, \; d_i = d_j \; \forall \; i, j \in C$.

Theorem 2.1 - *Stationary Distribution & Transition Probabilities*
Let $X := \{X_n\}_{n \in \mathbb{N}}$ be an irreducible aperioic *Markov Chain*, with *Stationary Distribution* π. Then

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

Theorem 2.2 - *Chapman-Kolmogorov Equation*

$$p_{ij}(n) = \sum_{k \in S} p_{ik}(r) p_{kj}(n - r) \; \forall \; i, j \in S, \; n \in \mathbb{N} \; \& \; r \in [0, n]$$

2.2 Continuous Time

Definition 2.8 - *Generator Matrix*
A *Generator Matrix* is an alternative way of displaying continuous-time *Markov Chains*. A *Generator Matrix* G is defined st

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

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

Definition 2.9 - *Jump Chain*
A *Jump Chain* $\{Y_n\}_{n \in \mathbb{N}}$, for a *Continuous-Time Markov Chain* $\{X_t\}_{t \in \mathbb{R}^{\geq 0}}$, is a *Discrete-Time Markov Chain* which describes the different states X_t moves to. $Y_n := X_{S_n}$ where S_n is an arrival time.

Definition 2.10 - *Recurrence & Transience - Continuous Time*
A state $i \in S$ is *Recurrent* if $\mathbb{P}(\{t \geq 0 : X_t = i\}$ is unbounded$|X_0 = i) = 1$. A state is only recurrent in continuous time iff it is recurrent for the *Jump Chain* Y_n. A state $i \in S$ is *Transient* if $\mathbb{P}(\{t \geq 0 : X_t = i\}$ is unbounded$|X_0 = i) = 0$.

Definition 2.11 - *Poisson Process, N_t*
Let N_t be the number of events to have occured by time t. Then $\{N_t\}_{t \in \mathbb{R}^{\geq 0}}$ is a *Poisson Process* with rate λ if

i) $N_0 = 0$;

ii) $N_t \in \mathbb{N} \; \forall \; t \in \mathbb{R}^{\geq 0}$;

iii) $N_{t+s} - N_t$ depends on s only (Stationary Increments);

iv) $N_{t_2} - N_{t_1}, \dots, N_{t_n} - N_{t_{n-1}}$ are all independent. (Independent Increments);

v) $\forall\ t,h>0$ we have

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

For $t\geq 0,\ N_t\sim \text{Po}(\lambda t) \implies \mathbb{E}(N_t)=\text{Var}(N_t)=\lambda t$. The filtration $\mathcal{F}_t$ of a *Poisson process* is generated by the process itself.

Definition 2.12 - *Arrival Times*
For a *Poisson Process* $\{N_t\}_{t\in\mathbb{R}^{\geq 0}}$ we define $S_i:=\inf\{t\geq 0: N_t=i\}$ to be the *Arrival Time* of the occurence of the i^{th} event. Since $N_t\sim\text{Po}(\lambda t)$ then $S_1\sim\text{Exp}(\lambda)$. Since $S_n=\sum_{j=1}^i T_j$ then $S_i\sim\Gamma(i,\lambda)$.

Definition 2.13 - *Inter-Arrival Times*
For a *Poisson Process* $\{N_t\}_{t\in\mathbb{R}^{\geq 0}}$ we define $T_1=S_1,\ T_i=S_i-S_{i-1}\ \forall\ i\in\mathbb{N}^{\geq 2}$ to be the *Inter-Arrival Time* of events. *Inter-Arrival Times* are independent of each other. $T_i\sim\text{Exp}(\lambda)$.

Proposition 2.1 - *Waiting Times*
Let T be a random variable that models how long customers have to wait. We want T to have the following property $\mathbb{P}(T\in(t,t+h]|T>t)=\lambda h+o(h)$ where $\lim_{h\to 0}\frac{o(h)}{h}=0$ & λ is the average rate that customers are servered at. Thus T is modelled by $T\sim\text{Exp}(\lambda)$. If we have multiple *servers* who serve customers at different rates we can model them as indepdent random variables $T_i\sim\text{Exp}(\lambda_i)$ where λ_i is the rate of the i^{th} server.

Proposition 2.2 - *First Expected*
Suppose we have n servers whose serving times $T_i\sim(\lambda_i)$. Then

$$\begin{array}{rcl}\min(T_1,\dots,T_n) & \sim & \text{Exp}\left(\sum_{i=1}^n\lambda_i\right)\\ \mathbb{P}\big(T_i=\min(T_1,\dots,T_n)\big) & = & \lambda_i\frac{1}{\sum_{i=1}^n\lambda_i}\end{array}$$

Definition 2.14 - *Birth & Death Processes*
Birth Death Processes are continuous time processes where we consider a population of indepdent individuals where population members either: give birth (increase population by 1); or, die (decrease population by 1). We define

i) N_t to be the population size at time t (with $N_0=1$);

ii) $p_n(t):=\mathbb{P}(N_t=n)$;

iii) $S_i:=\inf\{t\geq 0: N_t=i\}$; and,

iv) $T_i:=S_1-S_{i-1}$.

Definition 2.15 - *Generalised Birth & Death Process*
A *Continuous Time Stochastic Process*, $\{N_t\}_{t\in\mathbb{R}^{\geq 0}}$, is a *Generalised Birth & Death Process* if

i) $N_t\in\mathbb{N}\ \forall\ t\in\mathbb{R}^{\geq 0}$; and,

ii) For $\lambda_n,\mu_n\geq 0\ \forall\ n$ we have

$$\begin{array}{rcl}\mathbb{P}(N_{t+h}-N_t=\ \ 1|N_t=n) & = & \lambda_n h+o(h)\\ \mathbb{P}(N_{t+h}-N_t=-1|N_t=n) & = & \mu_n h+o(h)\\ \mathbb{P}(N_{t+h}-N_t=\ \ 0|N_t=n) & = & 1-(\lambda_n+\mu_n)h+o(h)\end{array}$$

iii)
$$\begin{array}{rcl} p'_n(t) & = & \lambda_{n-1}p_{n-1}(t) - (\lambda_n + \mu_n)p_n(t) + \mu_{n+1}p_{n+1}(t) \ n \geq 1 \\ p'_0(t) & = & -\lambda_0 p_0(t) + \mu_1 p_1(t) \end{array}$$

Proposition 2.3 - *Pure Immigration*
For a *Birth & Death Process* with a $\lambda_n = \lambda,\ \mu_n = 0\ \forall\ n \in \mathbb{N}$, for $\lambda \in [0,1]$, we have that $\{N_t\}_{t\in\mathbb{R}^s}$ is a *Poisson Process.*

Proposition 2.4 - *Linear Birth Process (Yule Process)*
The *Yule Process* is the *Linear Birth Process.* This is a *Birth & Death Process* where $\lambda_n = n\lambda,\ \mu_n = 0\ \forall\ n \in \mathbb{N}$ for $\lambda, \mu \in [0,1]$. Then

i) $N_t \sim \text{Geo}(e^{-\lambda t})$;

ii) $p_n(t) = e^{-\lambda t}(1 - e^{-\lambda t})^{n-1}$;

iii) $\mathbb{E}(S_i) \simeq \frac{1}{\lambda}\ln n$;

iv) $T_i \sim \text{Exp}(i\lambda)$;

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

Proposition 2.5 - *Linear Birth & Death Process*
The *Linear Birth & Death Process* is a *Birth & Death Process* where $\lambda_n = n\lambda,\ \mu_n = n\mu\ \forall\ n \in \mathbb{N}$ for $\lambda, \mu \in [0,1]$. Then

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

Proposition 2.6 - *Holding Times*
Consider a *Continous-Time Markov Chain* on state space S. Suppose the chain enters state $i \in S$ at time t, define $U := \inf\{s \geq - : X_{t+s} \neq i\} \sim \text{Exp}(\lambda_i)$ to be *Holding Time* of the chain in state i. We define H to be the matrix st $h_{ij} = \mathbb{P}(X_{t+U} = j | X_t = i)$ (probability of transitioning to state i from state j). Then $h_{ij} = -\frac{g_{ij}}{g_{ii}} = \frac{g_{ij}}{\lambda_i}$.

Definition 2.16 - *Kolmogorov Equations*
Forwards Equations $P'_t = P_t G$ since $P_0 = I \implies P_t = e^{tG}$
Backwards Equations $P'_t = GP_t$.
Further, we derive that

$$\frac{d^k}{dt^k}P_t\bigg|_{t=0} = G^k \text{ for } k \geq 0 \implies g_{ij} = p'_{ij}(0)$$

3 Random Walks

Definition 3.1 - *Random Walk*
A *Random Walk* is a discrete-time stochastic process where the value of the process increases or decreases by 1, only (it always changes value). *Random Walk*s have the *Markov Property.*

*Random Walk*s can be defined to include *Absorbing Barriers* which are values that if the process ever achieves, the process stops.

Definition 3.2 - *Probability of Visiting*
Consider a time-homogenous stochastic process X on state space S. Let $i, j \in S$ be a state. We define

$$f_{ij} := \mathbb{P}(\text{We will visit } j | X_0 = i) \quad \& \quad f_{ij}(n) := \mathbb{P}(X_n \text{ is first visit to } j | X_0 = i)$$

This can be equated with *n-Step Transition Probabilities* as

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

If we are guaranteed to return to a state $s \in S$ (*i.e.* $f_{ss} = 1$) then s is *Recurrent*, else (*i.e.* if $f_{ss} < 1$) s is *Transient*. A state j is *Recurrent* iff $P_{ij}(1) = \infty \ forall\ i$. If a state j is *Transient* then $p_{ij}(n) \xrightarrow{n \to \infty} 0$.

Proposition 3.1 - *Generating Functions*
For *n-Step Transition Probabilities*, $p_{ij}(n)$, & *First Passage Probabilities*, $f_{ij}(n)$, we define the following generating functions

$$\begin{array}{ll} & P_{ij}(s) = \sum_{i=1}^{\infty} p_{ij}(n) s^n \qquad\qquad F_{ij}(s) = \sum_{n=1}^{\infty} f_{ij}(n) s^n \\ \implies & P_{ij}(s) = \mathbb{1}_{i=j} + F_{ij}(s) P_{jj}(s) \end{array}$$

These can be used to relate $f_{ij}(n)$ with $p_{ij}(n)$.

Proposition 3.2 - *Mean Time of First Passage*
Consider a time-homogenous stochastic process X on state space S. We can calculate the *Mean Time of First Passage* from state $i \in S$ to state $j \in S$ as

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

<u>N.B.</u> -m_{ii} is called the *Mean Time of Return.*

Proposition 3.3 - *Gambler's Ruin*
The *Gambler's Ruin* is a common scenario used to demonstrate a *Random Walk.* A gambler starts with £k & stops gambling if they reach £0 or £N (*i.e.* There are absorbing barries at 0 & N). Each time £1 is gambled & either it is lost (with probability 1) or £1 is won (with the bet returned), with probability $q := 1 - p$.

$$p_k := \mathbb{P}(\text{Gambler ever ruined} | X_0 = k) \implies p_k = p_{k+1} p + p_{k-1} q = \frac{\left(\frac{q}{p}\right)^k - \left(\frac{q}{p}\right)^N}{1 - \left(\frac{q}{p}\right)^N}$$

Theorem 3.1 - *Probability of Ruin against Upper Barrier*
Let p_k be the probability of ruin for a gambler starting at £k & no upper barrier. Let $p_k^{(N)}$ be the probability of ruin for a gambler with an upper barrier at N. Then

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

Definition 3.3 - *Stopping Time*
Let $X = \{X_n\}_{n \in \mathbb{N}}$ be a *Stochastic Process.* Non-negative integer-valued random variable T is a

Stopping Time Process for X if $\forall\ n \in \mathbb{N}$ the event $\{T \leq n\}$ id determined by $X_0, \dots, X_n$ only. Often a *Stopping Time* is an event where a first passage occurs.

Theorem 3.2 - *Wald's Lemma*
Let $Z_1, Z_2, \dots$ be a sequence of iid rv with $\mathbb{E}(|Z_n|) < \infty$ & $X_n := \sum_{i=1}^n Z_i$. If T is a *Stopping Time Process* if $X = \{X_n\}_{n\in\mathbb{N}}$ with $\mathbb{E}(T) < \infty$. Then

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

4 Brownian Motion / Weiner Process

Definition 4.1 - *Simple Symmetric Random Walk*
A *Simple Symmtric Random Walk* is a process $S_1, S_2, \dots$ where $S_i = \sum_{j=1}^i Y_j$ with $Y_1, Y_2, \dots$ being independent rvs which take values $\{-1, 1\}$ with probability $\frac{1}{2}$.
Definition 4.2 - *Brownian Motion*
Brownian Motion is a *continuous-time stochastic process* which models random motion in continuous space $\mathbb{R}^n, n \in \mathbb{N}$.
Let $\mathcal{F}_t$ be a filtration. An *Adpated Stochastic Process* $W := \{W_t\}_{t\geq -}$ is a *Brownian Motion* if

i) $W_0 = x,\ x \in \mathbb{R}$;

ii) $W_{t+u} - W_t \sim \mathcal{N}(0, u)$ is independent of $\mathcal{F}_t\ \forall\ t, u \geq 0$ ($\implies W_t \sim \mathcal{N}(0, t)$); and,

iii) $t \mapsto W_t(\omega)$ is a continuous funciton of $t\ \forall\ \omega \in \Omega$.

N.B. -W is a *Standard Brownian Motion* if $x = 0$.

Proposition 4.1 - *Constructing Brownian Motion from a Simple Symmetric Random Walk*
By defining $X_t^n = \frac{1}{\sqrt{n}} S_{nt}$ for $t = \frac{j}{n} \in [0, 1]$ (and linerly interpreting other values) we can show that as $n \to \infty$ X_t^n satisfies the definition for *Brownian Motion.*
Theorem 4.1 - *Transition Density of Brownian Motion*

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

Proposition 4.2 - *Properties of a Standard Brownian Motion*
Let W be a *Standard Brownian Motion*, then

i) $\forall\ W_t \sim \mathcal{N}(0, t) \implies \mathbb{E}(W_t) = 0$ & $\text{Var}(W_t) = t$.

ii) $\forall\ 0 \leq s \leq t$ $\text{Cov}(W_S, W_t) = 0$.

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

iv) $-W_t$ is a *Standard Brownian Motion.*

v) $X = \{X_t\}_{t\geq 0}$ where $X_t := W_{t+s} - W_s$ is a *Standard Brownian Motion* for fixed $s > 0$.

vi) $X = \{X_t\}_{t\geq 0}$ where $X_t := \frac{1}{\sqrt{\alpha}} W_{\alpha t}$ is a *Standard Brownian Motion* for fixed $s > 0/$

Definition 4.3 - *The Reflection Principle*
Let W_t be a *Standard Brownian Motion.* Let $a \in \mathbb{R}^{\geq 0}$ and $\tau_a := \inf\{t > 0 : W_t = a\}$. Define $\widetilde{W_t} = \begin{cases} W_t & \text{for } t \leq \tau_a \\ a - (W_t - a) & \text{for } t > \tau_a \end{cases}$. Then $\widetilde{W_t}$ is a *Standard Brownian Motion.*

Proposition 4.3 - *Properties of First Passage Time*
Let W_t be a *Standard Brownian Motion.* Let $a \in \mathbb{R}^{\geq 0}$ and $\tau_a := \inf\{t > 0 : W_t = a\}$. Then

i) $\mathbb{P}(\tau_a \leq T) = 2\mathbb{P}(W_t > a)$.

ii) $f_{\tau_a}(t) = \frac{|a|}{\sqrt{2\pi t^3}} e^{-\frac{a^2}{2t}}$ for $t > 0$.

iii) $\mathbb{E}(\tau_a) = \infty$ for fixed $a \neq 0$.

Definition 4.4 - *Gaussian Process*
A *Gaussian Process* is a *continuous-time stochastic process* with continuous sample paths (Set of possible values for a stochastic process). and finite dimenisonal distributions that are multivariate normal. A *Gaussian Process* is completely determined by its mean function $\mu_t = \mathbb{E}(X_t)$ & auto-covariance function $\mathbb{E}((X_s - \mu_s)(X_t - \mu_t))$.

5 Martingales

Definition 5.1 - *Discrete-Time Martingales*
A *Discrete-Time Martingale*, wrt a filtration $\mathcal{F}_n$ (often generated by the process itself), is a *Stochastic Process* $Y := \{Y_n\}_{n\in\mathbb{N}}$ where $\forall\ n \in \mathbb{N}$

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

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

Definition 5.2 - *Continuous-Time Martingales*
A *Continuous-Time Martingale*, wrt a filtration $\mathcal{F}_t$ (often generated by the process itself), is a *Stochastic Process* $Y := \{Y_n\}_{n\in\mathbb{N}}$ where $\forall\ 0 \leq s \leq t$

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

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

Definition 5.3 - *Supermartingale & Submartingale*
A *Supermartingale* is a process where $\mathbb{E}(Y_t|\mathcal{F}_s) \leq Y_s$.
A *Submartingale* is a process where $\mathbb{E}(Y_t|\mathcal{F}_s) \geq Y_s$.

Proposition 5.1 - *Properties of Martingales*
Let $Y := \{Y_n\}_{n\in\mathbb{N}}$ be a *Discrete-Time Martingale*. Note the following hold, with appropriate alterations, for *Continuous-Time Martingales*, submartingales & supermartingales.

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

Remark 5.1 - *Stopped Martingales are still Martingales*

Theorem 5.1 - *Stopped Supermartingales*
Let $Y := \{Y_n\}_{n\in\mathbb{N}}$ be a *Supermartingale* wrt $\mathcal{F}_n$. Let T be a stoppping time of Y. Then $Z := \{Z_n\}_{n\in\mathbb{N}}$ with $Z_n := Y_{T\wedge n}$ is a *Supermartingale*.
N.B. -A similar results holds for *Continuous-Time Supermartingales*
Theorem 5.2 - *Optional Stopping Theorem - Discrete Time*
Let $Y := \{Y_n\}_{n\in\mathbb{N}}$ be a *Discrete-Time Martingale* wrt $\mathcal{F}_n$, let T be a stopping time of $\mathcal{F}_n$.
If any of the following conditions holds then $\mathbb{E}(Y_T) = \mathbb{E}(Y_0)$

i) T is bounded (*i.e.* $\exists\ K \in \mathbb{R}^{\geq 0}$ *st prob*$(T < K) = 1$).

ii) T is finite and $\exists\ K > 0$ *st* $|Y_{T\wedge n}| < K\ \forall\ n \in \mathbb{N}$.

iii) $\mathbb{E}(T) < \infty$ and $\exists\ K > 0$ *st*$|Y_n - Y_{n-1}| \leq K\ \forall\ n < T$.

iv) T is finite <u>and</u> $\mathbb{E}(|Y_T|) < \infty$ <u>and</u> $\mathbb{E}(Y_n \mathbb{1}_{T>n}) \xrightarrow{n\to\infty} 0$.

Theorem 5.3 - *Martingale Convergence Theorem - Discrete Time*
Let $Y := \{Y_n\}_{n\in\mathbb{N}}$ be a *Supermartingale* wrt $\mathcal{F}_n$. Suppose there exists $A > 0$ st $\mathbb{E}(|Y_n|) \leq A\ \forall\ n \in \mathbb{N}$. Then $\exists$ *rv* Y_∞ st $\mathbb{P}\left(\lim_{n\to\infty} Y_n = Y_\infty\right) = 1$.

Proposition 5.2 - *Gambler's Ruin - Application*
Let $X_1, X_2, \dots$ be iid rv st $\mathbb{P}(X_i = 1) = p$ & $\mathbb{P}(X_i = -1) = q := 1 - p$. Define $S_0 = k$ & $S_n = S_{n-1} + X_n$, then $\{S_n\}_{n\in\mathbb{N}}$ is an unrestricted random walk. Let $T := \min\{n : S_n = 0 \text{ or } S_n = N\}$, then $\{Y_n\}_{n\in\mathbb{N}}$ is a random walk with absorbing barries at 0 & N.

$p = q$ $Y := \{Y_n\}_{n\in\mathbb{N}}$ is a martingale wrt $\mathcal{F}_n := \sigma(X_1, \dots, X_n)$.
The following properties hold

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

ii) $\mathbb{P}(Y_T = N) = \frac{Y_0}{N}$.

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

$p \neq q$ Define $B_n = \left(\frac{q}{p}\right)^{Y_n}$. Then $\{V_n\}_{n\in\mathbb{N}}$ is a martingale wrt $\mathcal{F}_n := \sigma(X_1, \dots, X_n)$.
The following properties hold

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

ii) $\mathbb{P}(Y_T = N) = \frac{1-\left(\frac{q}{p}\right)^k}{1-\left(\frac{q}{p}\right)^N}$.

Proposition 5.3 - *Brownian Motion - Application*
Let $W := \{W_t\}_{t\geq 0}$ be a *Standard Brownian Motion* wrt $\sigma_t := \sigma(W_1, \dots, W_t)$. Then

i) W_t is a martingale.

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

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

iv) $X_t := e^{at+\sigma W_t}$ is a martingale iff $a = -\frac{1}{2}\sigma^2$.

Let $a, b > 0$ and define $\tau := \min\{t \geq 0 : W_t \in \{a, -b\}\}$ then $\mathbb{P}(W_\tau = a) = \frac{b}{a+b}$ and $\mathbb{E}(\tau) = ab$.
Let $X_t := \mu t + \sigma W_t$ with $\mu < 0$ & define $M := \max_{t\geq 0} X_t$. For $a, b > 0$ then

$$\mathbb{P}(\tau_a < \tau_{-b}) = \frac{1 - e^{-\alpha b}}{e^{\alpha a} - e^{-\alpha b}} \text{ with } \alpha := -\frac{2\mu}{\sigma^2}. \text{ Thus } \mathbb{P}(M \geq a) = e^{-\alpha a}$$

0 Reference

0.1 Definitions

Definition 0.1 - *Equivalence Relation*
A process is an *Equivalence Relation* if it is: Reflexive ($a = a$); symmetric ($a = b \implies b = a$); and, Transitive ($a = b,\ b = c \implies a = c$).

Definition 0.2 - *Lack of Memory Property*
A random variable X is said to have the *Lack of Memory Property* if

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

Definition 0.3 - *Positive Definite Matrix*
A *Positive Definite Matrix* is a matrix with strictly positive eigenvalues. A positive-definite matrix A can be defined as $A = P^{-1}DP$ where D is a diagonal matrix & P is a *unitary matrix.*

Definition 0.4 - *Stochastic Matrix*
A matrix M is a *Stochastic Matrix* if $[M]_{ij} \geq 0 \forall\ i, j \in [1, n]$ & $\sum_{j=1}^{n} M_{ij} = 1\ \forall\ i \in [1, n]$.

0.2 Notation

Notation 0.1 - *Transition Probabilities*
For a *Time-Homogenous Stochastic Process* we define *Transition Probabilities* as

$$\begin{aligned} p_{ij} &:= \mathbb{P}(X_1 = j | X_0 = i) = \mathbb{P}(X_{n+1} = j | X_n = i) \\ p_{ij}(n) &:= \mathbb{P}(X_n = j | X_0 = i) \end{aligned}$$

Notation 0.2 - *Stopped*
Let $Y := \{Y_n\}_{n\in\mathbb{N}}$ be a stochastic process. Let T be a stopping time of Y. Then

$$Y_{T\wedge n} := \begin{cases} Y_n & \text{for } n \leq t \\ Y_T & \text{for } n > t \end{cases}$$

0.3 Theorems

Theorem 0.1 - *Cauchy-Schwarz Inequaility*
Let X &Y be random variables with finite variance. Then

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

Theorem 0.2 - *Central Limit Theorem*
Let $X_1, X_2, \dots$ be iid rb with $\mu = \mathbb{E}(X)$ & $\sigma^2 = Var(X) \neq 0$. Then

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

Theorem 0.3 - *Equating Exponential & Binomial Distribution*
For an event that occurs at time $T \sim \text{Exp}(\lambda)$ then $\mathbb{P}(T < t) = 1 - e^{-\lambda t} \simeq \lambda t$ (for small λt). For n of these events (with the same distribution) the number of events that occur before time t is modelled by $N \sim \text{Bi}(n, \lambda t)$.

Theorem 0.4 - *Jensen's Inequailities*
Let g be a *Convex Function.* Then

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

Let g be a *Concave Function*. Then

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

Theorem 0.5 - *Markov's Inequaility*
Let X be a non-negative random variable. Then

$$\mathbb{P}(X > c) \leq \frac{\mathbb{E}(X)}{c} \ \forall\ c > 0$$

Theorem 0.6 - *One-Step Conditioning Argument*
Let A be an event that is depependent on the events X & Y. Then

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

0.4 Probability Distributions

Definition 0.5 - *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 0.6 - *Gamma Distribution*
Let T be a continuous randmo variable modelled by a *Gamma Distribution* with shape parameter α & scale parameter λ. Then

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

<u>N.B.</u> -$\alpha, \lambda > 0$.

Definition 0.7 - *Exponential Distribution*
Let T be a continuous random variable modelled by a *Exponential Distribution* with parameter λ. Then

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

<u>N.B.</u> -Exponential Distribution is used to model the wait time between decays of a radioactive source.

Definition 0.8 - *Normal Distribution*
Let X be a continuous random variable modelled by a *Normal Distribution* with mean μ & variance σ^2.
Then

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

Definition 0.9 - *Poisson Distribution*
Let X be a discrete random variable modelled by a *Poisson Distribution* with parameter λ. Then

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

<u>N.B.</u> -Poisson Distribution is used to model the number of radioactive decays in a time period.