Random Processes in Physics Notes

Dominik Szablonski

December 11, 2023

Contents

Chapter 1

Introduction to Randomness and Probability

It is important to distinguish between the different types of processes which we consider to be "random".

"Truly Random Processes"	"Chaotic Processes"	"Human Driven Processes"
Radioactivity Quantum effects	Dice rolling Weather	Road traffic Financial markets

In the course we treat these all as truly random processes.

1.1 Definitions

Probability

Definition. *These are the outcomes we observe.*

Statistics

Definition. *This is experimentally counting outcomes.*

In this definition, probability is the **number of a given outcome taken to an infinite sample size.**

Probability Theory

Definition. *Predicting what may happen.*

This may be defined mathematically by:

$$\text{Probability} = \frac{\text{Outcomes of Interest}}{\text{All possible outcomes}}$$

1.2 Basic Mathematical Notation of Probability

Odds Ratio

If an outcome in a sample happens N times, and **doesn't** happen M times, its odd ratio is written as:

$$N : M$$

Fraction

If we take outcomes like above, we can write probability P as:

$$P = \frac{N}{N+M}$$

Further probabilistic notation is shown below:

Concept	Notation
Outcome A	$P(A)$
Not A	$P(\overline{A})$
A and B	$P(A \cap B)$
A or B	$P(A \cup B)$

1.3 Related Outcomes

Suppose two events, A and B...

Independent

Definition. *The two events are independent if when A happens, it does not effect the probability of B occurring.*

Mutually Exclusive

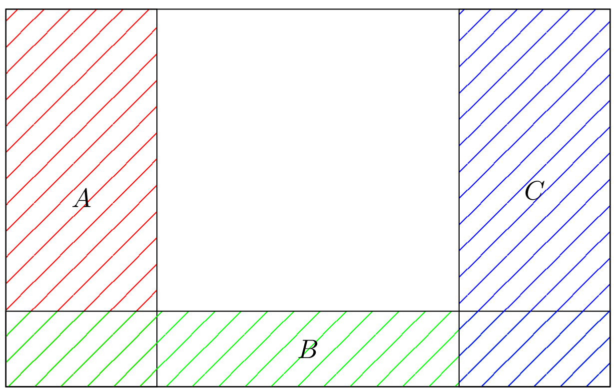
Definition. *The two events are mutually exclusive if when A happens the probability of B occurring is 0, and vice versa.*

1.4 Set Visualisation and Sample Space

Definition. *The range of all possible outcomes.*

In a sample space, the probability P of an outcome is equal to the fraction of the sample space corresponding to that outcome.

To represent a sample space, we may draw sample space diagrams. Usually, these are only able to represent events which are mutually exclusive and/or independent of each other. An example is shown below for events A, B, and C. Events A and B, and B and C are independent of each other, while A and C are mutually exclusive.



1.5 Axioms of Probability

These axioms are assumptions we make on which we build our understanding of probability theory.

1. For all outcomes x...
$$0 \leq P(x) \leq 1$$
2. An outcome either happens or it doesn't happen.
$$P(A) + P(\overline{A}) = 1$$
$$P(A \cup B) = 1$$
3. The probability of a set of mutually exclusive outcomes is teh sum of the probability of each outcome.
$$P(A_1 \cup A_2 \cup A_3 ... A_n) = \sum_{i=1}^{n} P(A_i)$$

Using these axioms we can derive various probabilistic properties. Some of these are shown below.

Generally...

$$P(A) = P(A \cap B) + P(A \cap \overline{B}) \tag{1.1}$$

$$\begin{aligned} P(A \cup B) &= P(A \cap \overline{B}) + P(\overline{A} \cap B) + P(A \cap B) \\ &= P(A) + P(B) - P(A \cap B) \end{aligned} \tag{1.2}$$

For independent outcomes:

$$P(A \cap B) = P(A)P(B) \tag{1.3}$$

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= P(A) + P(B) - P(A)P(B) \\ \text{This resembles a quadratic} &\rightarrow (1 - P(A))(1 - P(B)) \\ &= 1 - P(A) - P(B) + P(A)P(B) \\ \implies P(A \cup B) &= 1 - (1 - P(A))(1 - P(B)) \\ &= 1 - P(\overline{A})P(\overline{B}) \end{aligned} \tag{1.4}$$

As we have derived equation 4, we can generally say for n independent outcomes...

$$P(A_1 \cup A_2 \cup A_3 ... A_n) = 1 - \prod_{i=1}^{n} (1 - P(A_i))$$

1.6 Permutations and Combinations

A Priori

Definition. *relating to or derived by reasoning from self-evident propositions.*

It is derived from logical examination of an event. A lot of the times it involves the counting of all possible outcomes, or making assumptions about events. For example, assuming that a die is symmetrical in its physical properties.

If we unable to assume any information or are not given any information about an event, the best a priori assumption we are able to make is that the probability of an outcome is 50%, either it happens or doesn't happen.

Empirical/Posteriori

Definition. *relating to reasoning or knowledge which proceeds from observations or experiences to the deduction of probable causes.*

1.6.1 Permutations

When calculating the total number of permutations, we are interested in how many ordered subsets of size m there are in a set of size n.

$$\frac{n!}{(n-m)!} = {}^{n}P_m \tag{1.5}$$

We should also know that in the case $m = n$, $P_n = n!$.

1.6.2 Combinations

In this case, we are interested in the number of unique permutations in a subset of size m within a set of size n. In this case, permutations like $\{AB\}$ and $\{BA\}$ would be counted as the same permutation.

$$\frac{n!}{(n-m)!m!} = {}^{n}C_m = \binom{n}{m} \tag{1.6}$$

We are able to use this to generalise the binomial expansion.

$$(x+y)^n = \sum_{m=0}^{n} \binom{n}{m} x^m y^{n-m} \tag{1.7}$$

Proof.

$$\sum_{m=0}^{n} \binom{n}{m} = 2^n$$

.

Using 1.7, we can set $x = y = 1$.

$$\sum_{m=0}^{n} \binom{n}{m} \cdot 1 \cdot 1 = (1+1)^n$$
$$\sum_{m=0}^{n} \binom{n}{m} = 2^n$$

□

Proof.

$$\sum_{m=0}^{n} \binom{n}{m} = n2^{n-1}$$

Using 1.7, we can set $y = 1$.

$$\sum_{m=0}^{n} \binom{n}{m} x^m = (1+x)^n$$

Take derivatives with respect x.

$$\sum_{m=0}^{n} m\binom{n}{m} x^{m-1} = n(1+x)^{n-1}$$

Set $x = 1$

$$\sum_{m=0}^{n} m\binom{n}{m} = n(1+1)^{n-1}$$
$$\sum_{m=0}^{n} m\binom{n}{m} = n2^{n-1}$$

□

Pascal's Triangle

We are able to use pascal's triangle to derive certain properties of combinations.

$$\binom{n}{0} = 1$$
$$\binom{n}{n} = 1$$
$$\binom{n}{1} = \binom{n}{n-1}$$
$$\binom{n}{m} = \binom{n}{n-m}$$

Most events we analyse experimentally are not independent. When events are dependent, we are able to calculate:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{1.8}$$

This describes the probability of the outcome A **given that** B has already occurred.

1.7 Baye's Theorem

This theorum allows us to manipulate 1.8 so that we are able to calculate dependent probabilities more easily.

$$\begin{aligned} P(A|B) = \frac{P(A \cap B)}{P(B)} \qquad P(B|A) &= \frac{P(A \cap B)}{P(A)} \\ P(B)P(A|B) &= P(A)P(B|A) \\ P(A|B) &= \frac{P(A)}{P(B)} \cdot P(B|A) \end{aligned} \tag{1.9}$$

$$\begin{aligned} P(B) &= P(B \cap A) + P(B \cap \overline{A}) \\ &= P(B|A)P(A) + P(B|\overline{A})P(\overline{A}) \\ P(A|B) &= \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\overline{A})P(\overline{A})} \end{aligned} \tag{1.10}$$

Chapter 2

Discrete and Continuous Random Variables

Probability Distribution

Definition. *Encompasses all values of a an outcome of a random variable together with the probability each each possible value of that variable.*

2.1 Discrete Random Variables

This type of distribution applies to countable outcomes.

We can create a discrete random variable, such as X. We can then say X can take a range of values such that

$$R_x = \{x_1, x_2, x_3, ..., x_n\}$$

Further, we can say,

$$X : S \to \mathbb{R}$$

This is shorthand for saying the function of the discrete random variable X is a subset of a sample space S which maps to the set of real numbers.

2.1.1 Discrete Probability Distribution

We can then ascribe probability to our discrete random variable, using the following mapping.

$$\{x_i\} \to \{P_i\}$$
$$P_i = P(X = x_i)$$
$$0 \leq P_i \leq 1 \quad \forall i$$
$$\sum_i P_i = 1$$

Note For any set A which is a subset of which is a subset of R_x, the probability that X is an element of A can be found using the distribution,

$$P(X \in A) = \sum_{i \in A} P_i$$

2.1.2 Expectation Value

For a discrete random variable X with a range $R_x = \{x_1, x_2, ..., x_i\}$ with a probability distribution $P_i = P(X = x_i)$, the mean or expected value is defined as:

$$E(X) = \langle X \rangle = \bar{X} = \sum_i x_i P_i \tag{2.1}$$

2.2 Variance and Standard Deviation

$$\begin{aligned}\sigma &\implies \text{Standard Deviation}\\ \sigma^2 &\implies \text{Variance}\end{aligned}$$

2.2.1 Deriving Variance Equation

$$\begin{aligned}\sigma^2 &= \langle (X - \langle X \rangle) \rangle\\ &= \left\langle X^2 - 2X\langle X \rangle + \langle X \rangle^2 \right\rangle\\ \text{Using linearity of expectation} &\implies \langle X + Y \rangle = \langle X \rangle + \langle Y \rangle\\ \sigma^2 &= \left\langle X^2 \right\rangle - 2\left\langle \langle X \rangle X \right\rangle + \langle X \rangle^2\end{aligned}$$

Note: X is a constant real number.

$$\begin{aligned}&\implies \langle \langle X \rangle X \rangle = \langle X \rangle \langle X \rangle = \langle X \rangle^2\\ &\& \quad \left\langle \langle X \rangle^2 \right\rangle = \langle X \rangle^2\\ \therefore \sigma^2 &= \left\langle X^2 \right\rangle - 2\langle X \rangle^2 + \langle X \rangle^2\\ &= \left\langle X^2 \right\rangle - \langle X \rangle^2 \quad \square\end{aligned}$$

2.3 Geometric Distribution

Definition. *Represents the number of failures before you get a success in a series of trials.*

Assumptions:

1. There are only 2 outcomes
2. Each trial is independent
3. The probability of success is the same for each trial

$$p = \text{Success} \qquad q = 1 - p = \text{Failure}$$

$$P_n = pq^n \implies \text{Probability of } n \text{ failures before a succes.} \tag{2.2}$$

$$\langle n \rangle = \frac{p}{q} \tag{2.3}$$

$$\sigma^2 = \frac{1-p}{p^2} \tag{2.4}$$

2.3.1 Proof of Law of Total Probability for the Geometric Distribution

$$\sum_{n=0}^{\infty} P_n = \sum_{n=0}^{\infty} pq^n = p\sum_{n=0}^{\infty} q^n$$

$$\sum_{r=0}^{\infty} x^r = \frac{1}{1-x}$$

$$\therefore \sum_{n=0}^{\infty} P_n = p\frac{1}{1-q} = \frac{p}{p} = 1 \quad \square$$

2.3.2 Deriving the Expectation Value for the Geometric Distribution

$$
\begin{aligned}
\langle n \rangle = \sum_{n=0}^{\infty} n P_n &= p \sum_{n=0}^{\infty} n q^n \\
&= pq \sum_{n=0}^{\infty} n q^{n-1} \\
nq^{n-1} \to \frac{d}{dq}\left(q^n\right) & \\
&= pq \sum_{n=0}^{\infty} \frac{d}{dq}(q^n) \\
&= pq \frac{d}{dq}\left(\sum_{n=0}^{\infty} q^n\right) \\
&= pq \frac{d}{dq}\left(\frac{1}{1-q}\right) \\
&= \frac{pq}{(1-q)^2} \\
&= \frac{pq}{p^2} \\
&= \frac{q}{p} \quad \square
\end{aligned}
$$

2.3.3 Deriving the Standard Deviation and Variance for the Geometric Distribution

$$
\begin{aligned}
\left\langle n^2 \right\rangle &= \langle n(n-1) \rangle + \langle n \rangle \\
\sigma^2 &= \langle n(n-1) \rangle + \langle n \rangle - \langle n \rangle^2 \\
\langle n(n+1) \rangle &= \sum_{n=1}^{\infty} n(n-1) P_n \\
&= \sum_{n=0}^{\infty} n(n-1) p q^n \\
&= pq^2 \sum_{n=0}^{\infty} n(n-1) q^{n-2} \\
n(n-1)q^{n-2} &= \frac{d^2}{dn^2}(q^n) \\
\langle n(n+1) \rangle &= pq^2 \sum_{n=0}^{\infty} \frac{d^2}{dn^2}(q^n) \\
&= pq^2 \frac{d^2}{dn^2}\left(\frac{1}{1-q}\right) \\
&= pq^2 \frac{2}{1(1-q)^2} \\
\sigma^2 &= \frac{2pq^2}{(1-q)^3} + \frac{q}{p} - \frac{q^2}{p^2} \\
&= \frac{2q^2}{p^2} + \frac{pq}{p^2} - \frac{q^2}{p^2} \\
&= \frac{q(q+p)}{p^2} = \frac{1-p}{p^2} \quad \square.
\end{aligned}
$$

2.4 Continuous Random Processes

Definition. *A continuous random process is a random process with a sample space that can be mapped onto* $\mathbb{R}$.

When describing a continuous random process, we shouldn't ask what $P(X = x)$ is, as that value is infinitesimally small. However, it is sensible to ask questions such as what is $P(X > x)$.
Further, the shape of a continuous random process is defined by the **Probability Density Function (PDF)**. A PDF describes the shape of the distribution, and **does not indicate the probability of an event** Generally,

$$P(x_1 < X < x_1) = \int_{x_1}^{x_2} f(x)dx. \tag{2.5}$$

We also have a cumulative distribution function, such that,

$$C(x) = P(X \leq x). \tag{2.6}$$

A PDF should be normalised such that,

$$\int_{-\infty}^{\infty} f(x)dx = 1. \tag{2.7}$$

The expectation value of a continuous random process is given as,

$$\langle X \rangle = \int_{-\infty}^{\infty} xf(x)dx \tag{2.8}$$

2.5 Uniform Distribution

This distribution is defined by the PDF,

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & x < a, x > b \end{cases} \tag{2.9}$$

2.6 Normal Distribution

Theorem 2.6.1 (Central Limit Theorem). *For any*[1] *independent random variable (eg,* X_i*) with mean* μ_x *and variance* σ_x^2*, if* Y *is the mean of any value drawn for* X_i *i.e.,*

$$Y = \frac{\sum\limits_{i=1}^{n} X_i}{n}$$

then as $n \to \infty$*,* Y *will approach the normal distribution with a PDF given by,*

$$f(y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(\frac{-(y-\mu^y)^2}{2\sigma_y^2}\right) \tag{2.10}$$

with $\mu_y = \mu_x$ *and* $\sigma_y^2 = \dfrac{\sigma_x^2}{N}$

2.7 The Cauchy Distribution

It has a PDF,

$$f(x) = \frac{1}{\pi(1+x)^2} \tag{2.11}$$

It has no mean, no variance, and has large tails which mean that extreme values are likely.

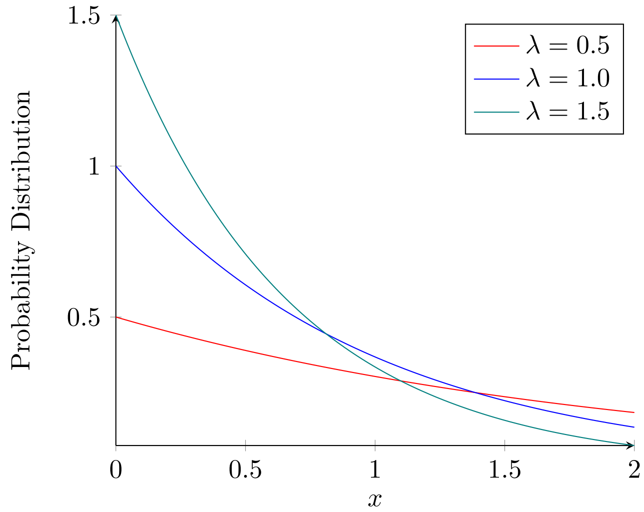
[1] As long as X has a well defined mean and variance.

2.8 Exponential Distribution

Definition. *A continuous distribution which defines the time needed for an event to occur.*

Its PDF takes the form

$$f(x) = \lambda e^{-\lambda x}. \tag{2.12}$$



2.8.1 Derivation of PDF

Proof. Consider a time interval, Δt,

$$\begin{aligned} P(\text{decay}) &\propto \Delta t \\ &= \lambda \Delta t \\ \lambda &= \frac{1}{\tau} \\ \tau \rightarrow &\text{ Characteristic Decay Time} \\ P(\text{no decay}) &= 1 - \lambda \Delta t \end{aligned}$$

Now consider the probability that there is no decay **after** a time $n\Delta t$,

$$\begin{aligned} P(\text{no decay after } n\Delta t) &= (1 - \lambda \Delta t)^n \\ \implies P(T > t), t &= n\Delta t \end{aligned}$$

Where T is an arbitrary time at which point we think decay will occur, and t is the point in time immediately after which the event has a chance to occur. This resembles the survival probability.

$$
\begin{aligned}
P(T > t) &= \lim_{n\to\infty}(1-\lambda\Delta t)^n = \lim_{n\to\infty}\left(1-\frac{\lambda t}{n}\right)^n \\
P(T > t) &= y \\
\ln y = n\ln\left(1-\frac{\lambda t}{n}\right) &= \frac{\left(1-\frac{\lambda t}{n}\right)}{\frac{1}{n}} \\
\lim_{n\to\infty}\ln y = \lim_{n\to\infty}\frac{\left(1-\frac{\lambda t}{n}\right)}{\frac{1}{n}} &\to \text{ Takes the form of } \frac{0}{0} \\
\therefore \text{ use L'Hopital's Rule,} \\
\lim_{n\to\infty}\ln y = \lim_{n\to\infty}\frac{\frac{\lambda t}{n^2}}{-\frac{1}{n^2}\left(1-\frac{\lambda t}{n}\right)} &= -\lambda t \\
\therefore \lim_{n\to\infty}\ln y = -\lambda t &\iff \lim_{n\to\infty} y = e^{-\lambda t} \\
\therefore P(T > t) &= e^{-\lambda t} = 1 - P(T \leq t)
\end{aligned}
$$

$P(T \leq t)$ takes the form of the cumulative function;

$$
\begin{aligned}
P(T > t) &= 1 - c_T(t) \\
f_T(t) &= \frac{d}{dt}c_T(t) \\
c_T(t) &= 1 - P(T > t) = 1 - e^{-\lambda t} \\
\therefore f_T(t) &= \lambda e^{-\lambda t}
\end{aligned}
$$

□

2.8.2 Properties of the Exponential Distribution

Law of Total Probability

Proof.

$$
\begin{aligned}
\int_o^{infty} f(t)dt &= \lambda\int_0^\infty e^{-\lambda t}dt \\
I(t) = \int_0^\infty e^{-\lambda t}dt &= \left[-\frac{1}{\lambda}e^{-\lambda t}\right]_0^\infty \\
&= \frac{1}{\lambda} \\
\int_o^\infty f(t)dt &= \lambda\cdot\frac{1}{\lambda} = 1
\end{aligned}
$$

□

Expectation Value

Proof.

$$
\begin{aligned}
\langle t\rangle = \int_o^\infty tf(t)dt &= \lambda\int_0^\infty te^{-\lambda t}dt \\
I(\lambda) &= \int_0^\infty e^{-\lambda t}dt = \frac{1}{\lambda} \\
\frac{d}{d\lambda}I(\lambda) = -\frac{1}{\lambda^2} &= \frac{d}{d\lambda}\int_0^\infty e^{-\lambda t}dt \\
\text{Using Leibniz rule} \rightarrow &= \int_0^\infty \frac{\partial}{\partial\lambda}e^{-\lambda t}dt \\
&= -\int_0^\infty te^{-\lambda t}dt \\
\therefore \langle t\rangle &= -\lambda\frac{d}{d\lambda}(I(\lambda)) \\
&= -\lambda\cdot\frac{1}{\lambda^2} \\
&= \frac{1}{\lambda} = \tau
\end{aligned}
$$

□

Variance and Standard Deviation

Proof.

$$
\begin{aligned}
\sigma^2 &= \left\langle t^2\right\rangle - \langle t\rangle^2 \\
\left\langle t^2\right\rangle &= \int_0^\infty t^2 f(t)dt \\
&= \lambda\int_0^\infty t^2e^{-\lambda t}dt \\
\frac{d^2}{d\lambda^2}I(\lambda) = \frac{2}{\lambda^3} &= \frac{d^2}{d\lambda^2}e^{-\lambda t}dt \\
&= \int_0^\infty \frac{\partial^2}{\partial\lambda^2}e^{-\lambda t}dt \\
&= \int_0^\infty t^2e^{-\lambda t}dt \\
\therefore \left\langle t^2\right\rangle &= \lambda\cdot\frac{d^2}{d\lambda^2}I(\lambda) \\
&-\lambda\cdot\frac{2}{\lambda^3} = 2\tau^2 \\
\sigma^2 &= 2\tau^2 - \tau^2 = \tau^2 \\
\sigma &= \tau
\end{aligned}
$$

□

2.8.3 Memorylessness

Past events do not effect future behaviour.

Proof. Let's consider,

$$\begin{aligned} P(T > a + b | T > a) &= \frac{P(T > a + b \cap T > a)}{P(T > a)} \\ &= \frac{P(T > a + b)}{P(T > a)} \\ &= \frac{e^{-\lambda(a+b)}}{e^{-\lambda a}} \\ &= \frac{e^{-\lambda a} e^{-\lambda b}}{e^{-\lambda a}} \\ &= e^{-\lambda b} = P(T > b) \end{aligned}$$

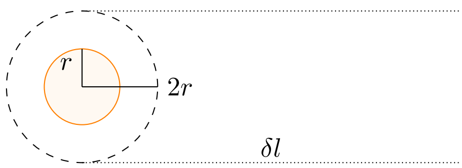
□

2.9 Collisions in a Gas

We need to try and find an expression that a gas molecule will collide after travelling a distance l. Consider the probability that the particle will travel a distance $l + \delta l$.

$$P(l + \delta l) = P(l) \cdot P(\delta l).$$

Let's say the gas occupies a volume V. As the particle travels, it creates a volume ΔV where collisions may occur.



The dotted region is where a collision could occur. We can then define,

$$\Delta V = \pi(2r)\delta l = 4\pi r^2 \delta l.$$

We can then define the probability that a particle is not present in the colliding volume,

$$P = \frac{V - \Delta V}{V} = 1 - \frac{\Delta V}{V},$$

for N particles,

$$P_N = \left(1 - \frac{\Delta V}{V}\right) \simeq 1 - N\frac{\Delta V}{V} \text{ by the binomial expansion.} \tag{2.13}$$

We can then define the probability of a particle experiencing no collision after travelling a distance l,

$$P(l + \delta l) = P(l) \cdot \left\{1 - \frac{N\Delta V}{V}\right\}.$$

To get further insight, we may look at the Taylor expansion of $P(l + \delta l)$.

$$\begin{aligned} P(l + \delta l) &= P(l) + \delta l \frac{dP(l)}{dl} + \cdots \\ &= P(l) \cdot \left\{1 - \frac{N\Delta V}{V}\right\} \\ \therefore \delta l \frac{dP(l)}{dl} &= -\frac{N\delta V}{V} P(l) \\ &= -\frac{4\pi r^2 \delta l N}{V} P(l) \\ \frac{dP(l)}{dl} &= -\frac{4\pi r^2 N}{V} P(l). \end{aligned}$$

We can then define the mean-free path of the particle, λ by,

$$\frac{1}{\lambda} = \frac{4\pi r^2 N}{V},$$

so,

$$\begin{aligned}
\frac{dP(l)}{dl} &= -\frac{P(l)}{\lambda} \\
\int \frac{1}{P(l)} dP(l) &= \int -\frac{1}{\lambda} dl \\
\ln P(l) &= -\frac{l}{\lambda} + c \\
P(l) &= e^{-\frac{l}{P(l)}} \\
&\rightarrow \text{ Constant disappears because } P(0) = 1.
\end{aligned}$$

We can differentiate the above expression to derive the PDF.

$$\frac{1}{\lambda} e^{-\frac{l}{\lambda}}. \tag{2.14}$$

2.10 Non-Constant Hazard Rate

This is the case for when the decay "constant" varies. We can think of this as the probability of an event occurring in the time interval $(t, \Delta t)$,

$$f(t + \Delta t) = P(t) \cdot \alpha(t)\Delta t,$$

where $\alpha(t)\Delta t$ is the probability that the event occurs after Δt. We further know that $f(t) = -\frac{dP}{dt}$, so

$$\begin{aligned}
\frac{dP}{dt} &= -f(t) = -P(t)\alpha(t) \\
\frac{1}{P}\frac{dP}{dt} &= -\alpha(t) \\
\int_{P(0)}^{P(t)} \frac{dP'}{P'} &= -\int_0^t \alpha(t')dt' \\
P(0) &= 1 \\
\implies P(t) &= \exp\left(-\int_0^t \alpha(t')dt'\right)
\end{aligned}$$

Chapter 3

The Poisson Distribution

It applies to independent events where the mean rate (hazard rate) of occurrence is known. It can help predict the number of events that occur in a certain time/space.

Table 3.1: Exponential vs Poisson Distribution

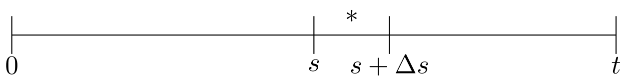
Exponential	Poisson
Probability of an event taking a time t to occur	Probability of getting k events in time t

3.1 Derivation

We will calculate a probability $P_k(t)$. We will also assume a constant hazard rate α. Our methodology will be to find $P_0(t) \to P_1(t) \to P_k(t)$. We already know the probability of no events occurring,

$$P_0(t) = e^{-\alpha t} \implies \text{Survival Probability.}$$

Now, consider a time window between 0 and t. In this time window, there occurs an event between time s and Δs.



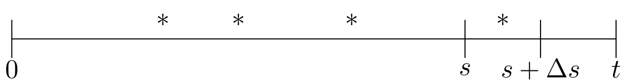
We can consider the probability of this event occurring in the range between s and $s + \Delta s$, keeping in mind that $\Delta s << 1$.

$$\begin{aligned}
\Delta P_1 &= P_0(s) P_1(\Delta s) P_0(t-s) \\
&= e^{-\alpha s} \alpha \Delta s e^{-\alpha(t-s)} \\
&= e^{-\alpha t} \alpha \Delta s.
\end{aligned}$$

We now want to find the probability across the entire window up to t,

$$\begin{aligned}
P_1 &= \sum_0^t \Delta P_1 \\
&= \sum e^{-\alpha t} \alpha \Delta \delta s \text{ let } \Delta S \to 0 \\
&= \int_0^t e^{-\alpha t} \alpha ds \\
P_1(t) &= \alpha t e^{-\alpha t}.
\end{aligned}$$

We can now consider an outcome where there are k events in the time window between 0 and t, where the last event occurs in the range s and $s + \Delta s$.



We can then define the probability of $k+1$ events occurring as,

$$P_{k+1}(t) = \int_0^t P_k(s)P_0(t-s)\alpha ds.$$

Iterating this for different values of k, we eventually come across a pattern which generalises to,

$$P_k(t) = \frac{(\alpha t)^k}{k!}e^{-\alpha t}. \tag{3.1}$$

This is often written with $\lambda = \alpha t$,

$$P_k(t) = \frac{\lambda^k}{k!}e^{-\lambda}.$$

3.2 Properties of the Poisson Distribution

- Increasing αt results in...
  - An increased mean
  - Increase in width
  - Change in the shape of the distribution
- It is a discrete distribution.

3.2.1 Law of Total Probability

$$\begin{aligned}
\sum_{k=0}^{\infty} P_K &= \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}e^{-\lambda} \\
&= e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} \\
&\implies \text{Series expansion of } e^{\lambda} \\
&\implies e^x = \sum_{n=0}^{\infty}\frac{x^n}{n!} \\
\sum_{k=0}^{\infty} P_K &= e^{-\lambda}\cdot e^{\lambda} \\
&= 1 \quad \square
\end{aligned}$$

3.2.2 Expectation Value

$$\begin{aligned}
\langle k\rangle = \sum_{k=0}^{\infty} kP_k &= \sum_{k=0}^{\infty} k\frac{\lambda^k}{k!}e^{-\lambda} \\
&= e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{(k-1)!} \implies \text{Not valid} \\
&= \lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!} + 0 \\
&= \lambda e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^n}{n!} \quad n = k-1 \\
&= \lambda e^{-\lambda}e^{-\lambda} \\
&= \lambda = \alpha t
\end{aligned}$$

3.2.3 Standard Deviation and Variance

$$\begin{aligned}
\sigma^2 &= \left\langle k^2 \right\rangle - \langle k \rangle^2 \\
\left\langle k^2 \right\rangle &= \langle k(k-1) \rangle + \langle k \rangle \\
\left\langle k^2 \right\rangle &= e^{-\lambda} \sum_{k=0}^{\infty} k(k-1) \frac{\lambda^k}{k!} \\
&= e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda}{(k-2)!} \\
&= \lambda^2 e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} \text{ Let } n = k-2 \\
&= \lambda^2 e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \\
&= \lambda^2 e^{-\lambda} e^{\lambda} \\
&\implies \left\langle k^2 \right\rangle = \lambda^2 + \lambda \\
&\implies \sigma^2 = \lambda^2 + \lambda - \lambda^2 \\
\sigma^2 &= \lambda
\end{aligned}$$

3.3 Sums of Poisson Variables

Lets suppose 2 events, which occur at a rate μ and λ. For the first event, m outcomes occur, for the second it's n. What is the probability of the total of the events occurring, P_{m+n}? We know,

$$\begin{aligned}
P_n &= \frac{\mu^n}{n!} e^{-\mu} \\
P_m &= \frac{\lambda^m}{m!} e^{-\lambda}
\end{aligned}$$

Let's consider the sum of the events, $s = m + n$. We can get a combination of events occurring,

$$m = s, n = 0; m = s-1, n = 1; \cdots ; m = 0, n = s$$

We can generalise this as $m = s - k, n = k$ where $k = 0, 1, 2, \cdots , s$. We can then consider the following probability,

$$\begin{aligned}
P([m = s-k] \cap [n = l]) &= P([m = s-k])P([n = l]) \\
&= \left(\frac{\lambda^{s-k}}{(s-k)!} e^{-\lambda} \right) \left(\frac{\mu^k}{k!} e^{-\mu} \right) \\
&= \frac{\lambda^{s-k}}{(s-k)!} \frac{\mu^k}{k!} e^{-(\lambda+\mu)} \\
&\implies P_s = \sum_{k=0}^{s} \frac{\lambda^{s-k}}{(s-k)!} \frac{\mu^k}{k!} e^{-(\lambda+\mu)}
\end{aligned}$$

Recalling the binomial expansion,

$$(a+b)^s = \sum_k \frac{s!}{(s-k)!k!} a^{s-k} b^k$$

$$\begin{aligned}
P_s &= \frac{e^{-(\lambda+\mu)}}{s!} \sum_{k=0}^{s} \frac{s!}{(s-k)!k!} \lambda^{s-k} \mu^k \\
&= \frac{e^{-(\lambda+\mu)}}{s!} (\lambda+\mu)^s \\
&= \frac{(\lambda+\mu)^s}{s!} e^{-(\lambda+\mu)}
\end{aligned}$$

We can notice, then, that the sum of events is simply the regular Poisson distribution with the mean being the sum of the mean of the two other events, $\langle s \rangle = \lambda + \mu$.

We can further extend this to multiple means,

$$\langle s \rangle = \sum_i \langle k_i \rangle$$

3.3.1 Central Limit Theorem

A Poisson distribution with a large λ can be considered as the sum of many small λ. From this, we can assume that for large λ, the Poisson distribution approaches the normal distribution.

3.3.2 Non-Constant Hazard Rate

A Poisson distribution will have a mean rate,

$$\begin{aligned} \lambda &= \sum \alpha(t)\Delta t \\ \lambda &= \int_0^t \alpha(t)dt \end{aligned}$$

3.4 Gaussian Limit of the Poisson Distribution

$$P_n = \frac{\lambda^n}{n!}e^{-\lambda}.$$

We cannot easily find the limit of $n!$, so we will use the Stirling approximation.

First, consider $\ln n!$...

$$\begin{aligned} \ln(n!) &= \ln(1) + \ln(2) + \ln(3) + \cdots + \ln(n) \\ &= \sum_{k=1}^{n} \ln(n) \simeq \int_1^n ln(x)dx \\ \implies &\text{ Using the product rule...} \\ \ln(n!) &= n(\ln(n) - 1) + 1. \end{aligned}$$

We need to add a further term to get a more accurate approximation,

$$ln(n!) \simeq n \ln n - n + \ln \sqrt{2\pi n}. \tag{3.2}$$

Now, let,

$$\begin{aligned} f(n) &= \ln(P_n) \\ &= \ln\left(\frac{\lambda^n}{n!}e^{-\lambda}\right) = n\ln\lambda - n\ln n + n - \ln\sqrt{2\pi n} - \lambda \\ &= n\ln\lambda - n\ln n + n - \ln\sqrt{2\pi n} - \lambda \\ f'(n) &= \ln\lambda - \ln n - \frac{1}{2n} \end{aligned}$$

For large λ, $f(n)$ reaches a maximum, therefore $f'(n) = 0$. Now, using the Taylor expansion,

$$
\begin{aligned}
f(n) &\simeq f(\lambda) + (n-\lambda)f'(\lambda) + \frac{(n-\lambda)^2}{2} f''(\lambda) \\
f'(\lambda) &= 0 \\
f''(n) &= -\frac{1}{n} + \frac{1}{2n^2} \\
\text{For large } \lambda,\ f''(\lambda) &= -\frac{1}{\lambda} \\
f(\lambda) &= -\ln\sqrt{2\pi\lambda} \\
f(n) &\simeq -\ln\sqrt{2\pi\lambda} - \frac{(n-\lambda)^2}{2\lambda} = \ln P_n \\
P_n &= \exp\left(-\ln 2\pi\lambda - \frac{(n-\lambda)^2}{2\lambda}\right) \\
&= \frac{1}{\sqrt{2\pi\lambda}} \exp\left(-\frac{(n-\lambda)^2}{2\lambda}\right) \quad \square
\end{aligned}
$$

Chapter 4

Binomial Distribution

The binomial distribution is used to find k successes in n trials. We make the assumptions,

1. 2 outcomes
2. Trials are independent
3. Probability of success is the same
4. There are a fixed number of trials

The probability for an experiment of n trails is

$$P_k = \binom{n}{k} p^k q^{n-k} \tag{4.1}$$

for k successes.

4.1 Law of Total Probability

$$\begin{aligned}
\sum_{k=0}^{n} P_k &= \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} \\
&= (p+q)^n \\
(p+q) &= 1 \\
\therefore \sum_{k=0}^{n} P_k &= 1 \quad \square
\end{aligned} \tag{4.2}$$

4.2 Expectation Value

$$\begin{aligned}
\langle k \rangle &= \sum_{k=1}^{n} k P_k = \sum_{k=0}^{n} k \frac{n!}{(n-k)!k!} p^k q^{n-k} \\
\frac{k}{k!} &= \frac{k}{k(k-1)!} = \frac{1}{(k-1)!} \\
\langle k \rangle &= \sum_{k=1}^{n} \frac{n!}{(k-1)!(n-k)!} p^k q^{n-k}
\end{aligned}$$

now $y = k-1, m = n-1 \implies k = y+1, n = m+1$, so,

$$\begin{aligned}
\langle k \rangle &= \sum_{y=0}^{m} \frac{(m+1)!}{y!(m-y)!} p^{y+1} (1-p)^{(m-y)} \\
&= (m+1)p \sum_{y=0}^{m} \frac{m!}{y!(m-y)!} p^y (1-p)^{m-y} \\
&= np \cdot (p+1-p)^m \\
\langle k \rangle &= np
\end{aligned}$$

4.3 Variance

$$
\begin{aligned}
\sigma^2 &= \left\langle k^2 \right\rangle - \langle k \rangle^2 \\
\left\langle k^2 \right\rangle &= \langle k(k-1) \rangle + \langle k \rangle \\
\langle k(k-1) \rangle &= \sum_{k=0}^{n} k(k-1) P_k \\
&= \sum_{k=2}^{n} k(k-1) P_k \\
&= \sum_{k=2}^{n} \frac{n!}{(k-3)!(n-k)!} P^k)(1-p)^{n-k}
\end{aligned}
$$

Now, $y = k-2, m = n-2$, so,

$$
\begin{aligned}
\langle k(k-1) \rangle &= n(n-1)p^2 \sum_{y=0}^{m} \frac{m!}{y!(m-y)!} p^y (1-p)^{n+1} \\
&= n(n-1)p^2 \cdot (p+1-p)^m \\
&= n(n-1)p^2 \\
\sigma^2 &= \langle k(k-1) \rangle + \langle k \rangle - \langle k \rangle^2 \\
&= n(n-1)p^2 + np - n^2p^2 \\
&= np(1-p)
\end{aligned}
$$

$$
\sigma^2 = npq \tag{4.3}
$$

4.4 Limit of Binomial

This will consider the binomial distribution for the limit as $N \to \infty$

$$\begin{aligned}
P_k &= \binom{N}{k} p^k (1-p)^{N-k} \\
\langle k \rangle &= Np \\
p &= \frac{\langle k \rangle}{N} \\
P_k &= \frac{N!}{k!(N-k)!} \left(\frac{\langle k \rangle}{N} \right)^k \left(1 - \frac{\langle k \rangle}{N} \right)^{N-k} \\
&= \frac{N!}{k!(N-k)!} \left(\frac{\langle k \rangle /N}{1 - \langle k \rangle /N} \right)^k \left(1 - \frac{\langle k \rangle}{N} \right)^N \\
\frac{N!}{(N-k)!} &= N(N-1)(N-2)\cdots(N-k+1) \\
&= N^k \left(1 - \frac{1}{N} \right) \left(1 - \frac{2}{N} \right) \cdots \left(1 - \frac{(k-1)!}{N} \right) \\
\lim_{N\to\infty} P_k &= N^k \\
\left(1 - \frac{\langle k \rangle}{N} \right) &= \sum_{l=0}^{n} \frac{n!}{l!(N-l)!} \left(1 - \frac{1 - \langle k \rangle}{N} \right)_l \\
&= \sum_{l=0}^{n} \frac{n^l}{l!} \left(- \frac{\langle k \rangle}{n} \right)^l \\
&= \sum_{l=0}^{n} \frac{(- \langle k \rangle)^l}{l!} = e^{-\langle k \rangle} \\
P_k &= \frac{N^k}{k!} \left(\frac{\langle k \rangle}{N} \right)^k \cdot e^{-\langle k \rangle} \\
&= \frac{\langle k \rangle^k}{k!} e^{-\langle k \rangle}.
\end{aligned}$$

From the above derivation, we can see that for the limit $N \to \infty$ and $\langle k \rangle = Np$ is fixed, the binomial distribution approaches the Poisson distribution.

Chapter 5

Diffusion and Walks

5.1 Speed of Diffusion

We can consider the speed of a molecule using the equation,

$$\begin{aligned}\langle KE \rangle &= \frac{3}{2} k_b T \\ \frac{1}{2} m \left\langle v^2 \right\rangle &= \frac{3}{2} k_b T\end{aligned}$$

We can then define the root mean square speed,

$$v_{\mathrm{RMS}} = \sqrt{\langle v^2 \rangle} = \sqrt{\frac{3k_B T}{m}}. \tag{5.1}$$

However, this value is too large. We need a better model.

5.2 Random Walks

We can consider the molecule bouncing off air molecules as it travels. We can split this up into motions with a vector $\mathbf{r}_i$ of a fixed length $|r_i| = \Delta l$.

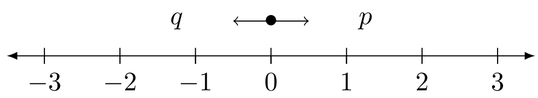
$$\begin{aligned}\mathbf{R} &= \sum_i \mathbf{r}_i \\ \langle \mathbf{R} \rangle &= \sum_i \langle \mathbf{r}_i \rangle\end{aligned}$$

The direction of $\mathbf{r}_i$ is completely random, therefore $\langle \mathbf{r}_i \rangle = 0$, hence, $\langle \mathbf{R} \rangle = 0$. We can attempt to improve our model by using R_{RMS}.

$$\begin{aligned}R_{\mathrm{RMS}} &= \sqrt{\langle \mathbf{R}^2 \rangle} = \left\langle \left(\sum_i \mathbf{r}_i \right)^2 \right\rangle \\ \sum_i \mathbf{r}_i^2 &= \left(\sum_i \mathbf{r}_i^2 \right) + 2 \left(\sum_i \mathbf{r}_i \cdot \mathbf{r}_{i+1} \right) \\ \left\langle \mathbf{r}_i^2 \right\rangle &= |\mathbf{r}_i|^2 = \Delta l^2 \\ \langle \mathbf{r}_i \cdot \mathbf{r}_j \rangle &= 0 \text{ Given that } i \neq j \\ &\Longleftarrow \text{ Directions are equally likely and independent} \\ \left\langle \mathbf{R}^2 \right\rangle = \sum_i^N \left\langle \mathbf{r}_i^2 \right\rangle &= \sum^N \Delta l^2 \\ &= N \Delta l^2 \\ R_{\mathrm{RMS}} &= \sqrt{N} \Delta l\end{aligned}$$

5.3 Random Walker in 1 Dimension

A random walk of this kind can be modelled by the binomial distribution.



We often want to know what the probability of being at position x is. We can model this by considering the probability of moving k steps right, and $N-k$ steps left.

$$P_k = \binom{N}{k} p^k q^{N-k}.$$

We can then say that the amount of steps it takes to get to x is,

$$x = k - (N-k). \tag{5.2}$$

5.3.1 Probability of Returning to the Beginning

Given that $p = q$, we can conclude that the walker will always return to the beginning for large N. This is because the probability of returning to the beginning after exploring N regions is,

$$P(\text{return}) = 1 - \left(\frac{1}{2}\right)^N. \tag{5.3}$$