Mathematics of Uncertainty

Lead instructor: Amrik Sen

Co-instructors: Sanjay Bose
Jitesh Panchal (Module 5.2)

Instructional Manager: Vanita Sharma

Teaching Assistants: Arpit Shukla
Manraj Singh Ghuman
Sandeep Kumar

Tutorials
Worksheet
Mini-projects

Mathematics of Uncertainty

Module 1: Probability Basics and Definitions	(6)
Module 2: Probability Distributions	(6)
Module 3: Discrete Time Markov Chains	(6)
Module 4: Statistical Experiments	(6)
Module 5: Statistics for Complex Problems	(4)

Each module consists of Lectures and a Mini-Project

Study of Probability
in
Ancient India

Movie Recommendation for the Weekend

Watch

"Breaking Vegas: The True Story of the MIT Blackjack Team"
(*Documentary*)

or

"21"
(*Hollywood Movie*)

Both Available on YouTube

Shows what you can do with Mathematics and Probabilities

Birthday Surprise

In a party, make a bet that there will be at least two people in the room with the same birthday!

If the number of people in the party is more than 23, you are likely to win the bet!

Textbook: “*Practical Introduction to Probability and Statistics – A project based conceptual guide to students and practioners*”, by Amrik Sen

Draft Edition, under peer review by Cambridge University Press, will be provided to students on course website

Reference Books:

1. “Weighing the Odds - A course in Probability and Statistics”, by David Williams.
2. “Probability Theory – The Logic of Science”, by E.T. Jaynes.
3. “Introduction to Probability and Statistics for Engineers and Scientists”, by Sheldon M. Ross.
4. “Probability, Random Variables and Stochastic Processes”, Papoulis and Pillai

Assessments:

- Two written exams: 30% Mid Term (proctored) + 20% End Term (proctored) TOTAL Weightage = 50%

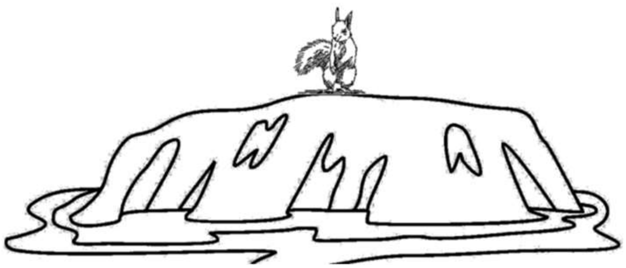
- Five Mini-Projects in the laboratory of 10% each TOTAL Weightage 50%

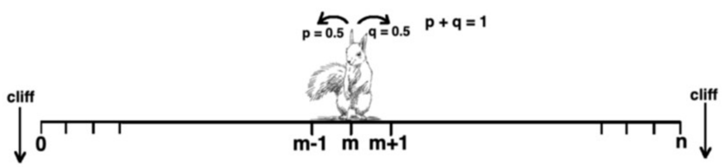
Note that students will have to obtain a "pass" grade separately for the written exam (theory) and the laboratory exams (projects) in order to pass the course.

For details, consult the "Course Brochure" provided in the course website.

The laboratory experiments and projects will use MATLAB (all modules) and also Python (for module 5.2)

Misadventures of Squeaky





Our friend Squeaky is trapped somewhere in the middle of a lonely island hill with sharp cliffs on both sides. Squeaky is excited and jumps around in her merry way. At any given instance, she decides to hop to the left or to the right independent of her past moves. Squeaky is unaware of the impending danger of falling off the cliff.

In this project, we will use calculations based on the principles of conditional probability, the law of total probability, and the law of total expectation to predict her fate. In other words, what are the odds that she will bounce around on the island hill, her left-sided moves balancing out her right-sided moves on an average, and never actually trip and fall off on either side? Or will chance play the devil's role and will she eventually drift off to one side and perish? And if the latter turns out to be true, then what is her life expectancy in terms of the total number of hops starting from her first move? Does a certain initial position on the hill give her the best chance to survive the longest?

In addition to our theoretical calculations, we will also build a computer simulation of her actions to corroborate our result. For convenience, we shall assume that the island is one dimensional, i.e., Squeaky's movements are restricted exclusively to lateral directions(left or right). While we build the computer-simulated solution, we will learn to apply a random number generator using a computer software in order to mimic Squeaky's mental choices to hop either to the left or to the right independent of her past moves.

Deterministic Outcomes

Permutations	*when the order matters*
Combinations	*when the order does not matter*

Permutations:

* Number of permutations of n different things $n! = n \times (n-1) \times .... \times 2 \times 1$
* Consider a situation where there are a total of n objects of r different types and where the number of objects of type k is n_k with $k = 1, 2, ..., r$

and $n = n_1 + n_2 + ... + n_r$

If we assume that the objects of the same type are indistinguishable from each other then the number of different ways in which the objects can be arranged is

$$\frac{n!}{n_1!n_2!......n_r!}$$

* If we have n different things then the number of permutations we get by taking only r ($r \leq n$) of them at a time are $P_r^n = n(n-1)\ldots..(n-r+1) = \frac{n!}{(n-r)!}$

For example, if we have nine slots and four differently colored balls, then the number of different arrangements possible are $\frac{9!}{5!} = 3024$

Combinations:

Consider when the order of arrangement is not important.

For example when we want to choose *r* items out of *n* in any order, then the number of ways in which this can be done is given by

$$C_r^n = \binom{n}{r} = \frac{n!}{r!(n-r)!}$$

For example, if we have nine slots and four identically colored balls, then the number of different arrangements possible are $\binom{9}{4}$=126

Useful Result: $C_r^n(r!) = P_r^n$

Permutations and Combinations, as discussed in the previous slides, belong to a class of experiments that have deterministic outcomes, i.e. there are a finite and fixed of ways of arranging or combining items.

There may be experiments where the outcomes are not certain. For example, if we consider the experiment of putting five differently colored balls in five slots, we may ask what is the probability that the first slot is filled with a ball of a specific color.

This is where the study of probability and statistics comes in.

Probability:	Giving a measure of likelihood that an event will occur
Statistics:	Dealing with the collection (sampling), organization, analysis and interpretation of data to make inferences and forecasts. (It relies on the principles of probability.)

Some Definitions

Probability: The Probability of an Event is the measure of the likelihood that the event will occur, e.g. *the probability that it will rain today is 0.75*

Statistics: This is the branch of mathematics that deals with the collection (sampling), organization, analysis and interpretation of data including making inferences and forecasts.

Probability deals with predicting the likelihood of future events, while statistics involves the analysis of the frequency of past events

Probability Space (Definition):

A probability space is a triple $(\Omega, \mathcal{F}, P)$ where Ω is a sample space, $\mathcal{F}$ is a $\sigma -$ algebra (sigma algebra) of events and *P* is a probability measure on $\mathcal{F}$

- The sample space Ω is the set of all possible outcomes of a probabilistic experiment
- The $\sigma -$ algebra $\mathcal{F}$ is the collection of all subsets of Ω to which we are able/willing to assign probabilities; these subsets are called events
- The probability measure *P* is a function that associates a probability to each of the events belonging to the $\sigma -$ algebra $\mathcal{F}$

Example: Probabilistic experiment to get a ball from an urn containing two balls, one red (*R*) and one blue (*B*)

$\Longrightarrow$ Sample Space $\Omega = \{R, B\}$ and a possible $\sigma -$ algebra $\mathcal{F}$ of events is $\mathcal{F} = \{\emptyset, \Omega, \{R\}, \{B\}\}$ where

$\emptyset$ is the empty set (nothing happens)	$P(F) = 0$ if $F = \emptyset$
Ω**:** either *R* or *B* is extracted	= 0.5 if *F*={*R*}
{*R*}: Red ball is extracted	= 0.5 if *F*={*B*}
{*B*}: Blue ball is extracted	= 1 if *F*= Ω

Axioms of Probability

Axioms are regarded as *a priori propositions* whose veracity is accepted universally without requiring validation by demonstration, i.e. they are accepted without proof.

Axioms are useful as they allow deduction of realizable experiences

1. $P(E) \geq 0$ for all $E \in \mathcal{F}$ (non-negativity of probability)

2. $P(\Omega) = 1$ (unitarity)

3. $P(\bigcup_{i=1}^{\infty} E_i)=\sum_{i=1}^{\infty} P(E_i)$ for a countable sequence of disjoint events E_1, E_2 ......
(σ-additivity)

Supplementary Properties of the probability measure *P* that are helpful while performing calculations

1. For $E_1, E_2 \in \mathcal{F}$, we have $P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$
   This can be generalized to *n* events E_1, E_2 ....E_n by induction. Principle of Inclusion-Exclusion

2. If E_1, E_2 are independent events (i.e. $E_1 \perp E_2$), then $P(E_1 \cap E_2) = P(E_1)P(E_2)$

3. If A^c is the event complementary to the event *A*, then $P(A^c) = 1 - P(A)$

4. The probability of the *impossible event* is zero, i.e., $P\{\emptyset\} = 0$

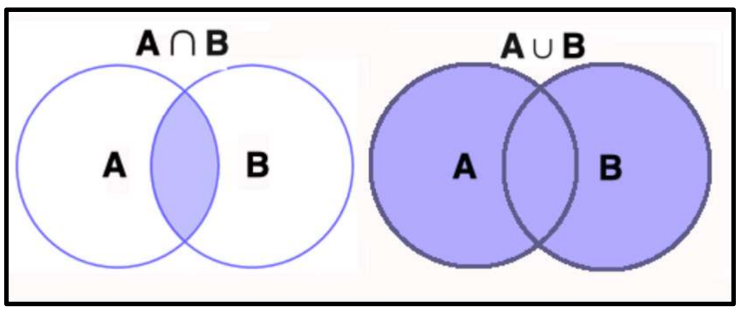
5. It is important to distinguish between *mutually exclusive* (disjoint) events and *independent* events

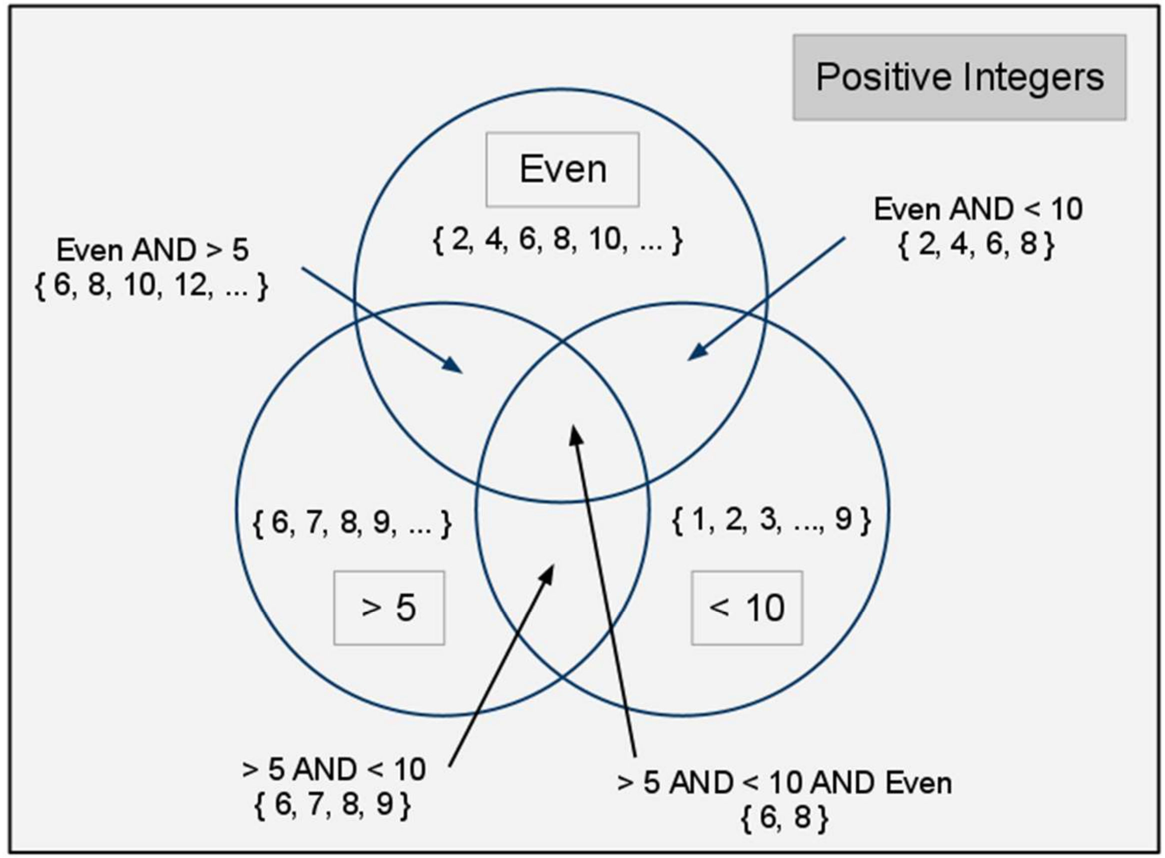
Mutually Exclusive	if $E_1 \cap E_2 = \{\emptyset\} \Rightarrow P(E_1 \cap E_2) = 0$
Independent Events	$P(E_1 \cap E_2) = P(E_1)P(E_2)$ and $P(E_1 \mid E_2) = P(E_1)$

Mutually Exclusive events cannot happen concurrently.
Independent Events may happen concurrently but the outcome of one does not affect the outcome of the other

Venn Diagram:

A diagram representing mathematical or logical sets pictorially as circles or closed curves within an enclosing rectangle (the universal set). Common elements of the sets being represented by intersections of the circles.





Example: Rolling two dice concurrently

The experiment comprises throwing two independent dice. The sample set is the Cartesian product comprising the following ordered pairs –

Ω = {1, 2, 3, 4, 5, 6} x {1, 2, 3, 4, 5, 6}
= {1, 1}, {1, 2}, {1, 3}, ..., {1, 6}, {2, 1}, {2, 2}, ..., {2, 6}, {6, 1}, {6, 2}, ..., {6, 5}, {6, 6}

We may be interested in knowing the odds that the sum of the outcomes from each die is greater than or equal to 10.

In this case, the event set is $\mathfrak{E}$ = { (4, 6), (5, 5), (5, 6), (6, 4), (6, 5), (6,6) }

with Probability $\frac{|\mathfrak{E}|}{|\Omega|} = \frac{6}{36}$

Random Variable

Ω Sample space of all possible outcomes
$\mathfrak{F}$ Event space from Ω
P Probability measure

Consider a probability space $(\Omega, \mathfrak{F}, P)$. A random variable is a measurable function, $X: \Omega \rightarrow \mathbb{R}$, that maps each outcome to a real number, i.e.

$$\{\omega \in \Omega; X(\omega) \leq x\} \in \mathfrak{F}, \qquad x \in \mathbb{R}$$

A random variable may be *discrete* or *continuous* depending on whether it takes on discrete values or continuous values

Example:

(a)For a throw of the dice, we assign to the six outcomes f_i i =1,…,6 the numbers $\boldsymbol{X}(f_i) = 10i$ to get the random variables

$$X(f_1) = 10 \ldots \ldots X(f_6) = 60$$

(b) In the same experiment, we assign the number 1 to every even outcome and 0 to every odd outcome. This gives the random variables

$$X(f_1) = X(f_3) = X(f_5) = 0 \qquad X(f_2) = X(f_4) = X(f_6) = 1$$

Example: A Coin Tossing Experiment

(a) Consider a simple experiment of tossing a fair coin $\Omega_1 = \{H, T\}$

We associate the random variables $X(H) = 1, X(T) = 0$ with each outcome.

(b) If we toss the coin twice, the outcomes in terms of events are

$$\Omega_2 = \{HH, HT, TH, TT\}$$

We can associate a random variable with each outcome by counting the number of HEADs in each case, i.e.

$$X(HH) = 2, X(HT) = 1, X(TH) = 1, X(TT) = 0$$

Indicator Random Variable

In some calculations, it is convenient to define an **Indicator Random Variable** as a random variable which takes binary values 1/0 for TRUE/FALSE.

$$\mathbb{I}_{A(\omega)} \equiv \mathbb{I}_{(\omega \in A)} = \begin{cases} 1; \ \omega \in A \\ 0; \ \omega \notin A \end{cases}$$

Example: Consider a fair coin being tossed twice, where the outcome of interest is one where there is at least one HEAD. In this case, we can define $A = \{HH, HT, TH\}$ and use the Indicator Random Variable to represent when at least one H appears.

Then, the Indicator Random Variable $\mathbb{I} = \begin{cases} 1 \ \ if \ HH, HT, TH \\ \quad 0 \ \ if \ \ TT \end{cases}$

We will use the Indicator Variable in a example that we will subsequently consider

Examples of Continuous Random Variables

1. Inter-arrival time of buses at a bus station
2. Received signal strength (signal power) from a wireless station
3. The Air Pressure in a car tyre
4. Temperature, Pressure, Humidity etc.

Conditional Probability

Consider two random events **A** and **B**.

Conditional Event: We define **A|B** as the conditional event that "A occurs **given** that event B has occurred"

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

$A \cap B$	Events *A* and *B* both occur
Notation: $A \cap B$ also written as *AB* (*A* **AND** *B*)	

Independent Events: Events A and B are independent ($\mathbf{A} \perp \mathbf{B}$) if the occurrence of event B does not affect the occurrence of event A, or vice versa. In that case,

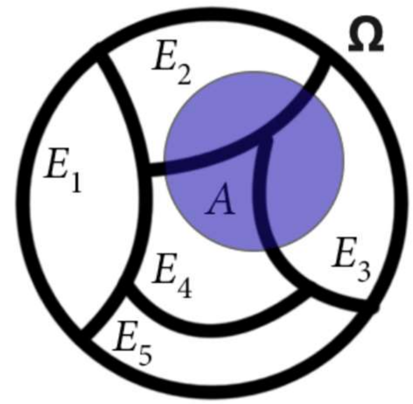
$P(A|B) = P(A)$, $P(B|A) = P(B)$, $P(A \cap B) = P(AB) = P(A)P(B)$

Law of Total Probability

The sample space Ω may be partitioned into *k* disjoint sets (i.e. events) where *i*=1, 2, …., *k.*

The probability of a certain event $A \subset \Omega$ can then be computed by the weighted sum of the conditional probabilities $P(A|E_i)$ where the weights are given by the probability of the partitioning events $P(E_i)$.

This is the Law of Total Probability



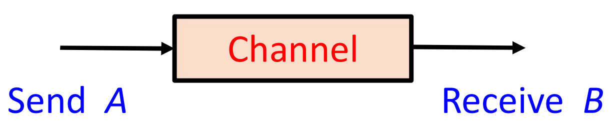
$$P(A) = \sum_{i=1}^{k} P(A \mid E_i) P(E_i)$$

Bayes' Theorem: $$P(A\,|\,B)=\frac{P(B\,|\,A)P(A)}{P(B)}$$

Baye's Theorem allow calculation of the *posterior probability* $P(A|B)$ by using the conditional probability $P(B|A)$ and the probabilities of the individual events $P(A)$ and $P(B)$ – these are known as the *prior probabilities*.

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

Notationally,
$$A\bigcap B=B\bigcap A=AB$$



Channel Characterization
Study what the channel does by sending different signals (A) and observing what are received (B), i.e., get $P(B|A)$.

Receiver Operation
When actually using the channel, observe what is received (B) and infer from that what is the most likely thing (A) that was sent. This is essentially an application of Bayes' Rule by also knowing $P(A)$ and $P(B)$

Example: Bayes' Theorem and the Law of Total Probability

A factory unit uses three automatic bolt threading machines (rollers), each accounting for 20%, 30%, and 50% of the factory output of ready-to-use bolts for the aerospace industry. The precision rating (number of non-defective parts produced per one hundred) of each of the rollers is 95%, 97%, and 99% respectively.

If a part is picked up at random from the production line and found to be defective, what is the probability that it was produced by the second machine?

Solution:

A_i : Event that a randomly picked bolt is made by the i^{th} machine, i=1, 2, 3

B : Event that a randomly chosen part is defective

From the given data, we get that –

$$P(A_1) = 0.2, P(A_2) = 0.3, P(A_3) = 0.5$$

and $$P(B|A_1) = 0.05, P(B|A_2) = 0.03, P(B|A_3) = 0.01$$

We need to find $P(A_2|B)$ i.e., given that we have a defective part (Event B), the probability that it came from machine A_2

Using Bayes' Theorem –

$$P(A_2|B) = \frac{P(B|A_2)P(A_2)}{P(B)}$$

$$= \frac{(0.03)(0.3)}{\sum_{i=1}^{3} P(B|A_i)P(A_i)}$$

$$= \frac{0.009}{(0.05)(0.2)+(0.03)(0.3)+(0.01)(0.5)}$$

= 0.375

Continuing with the previous example ....

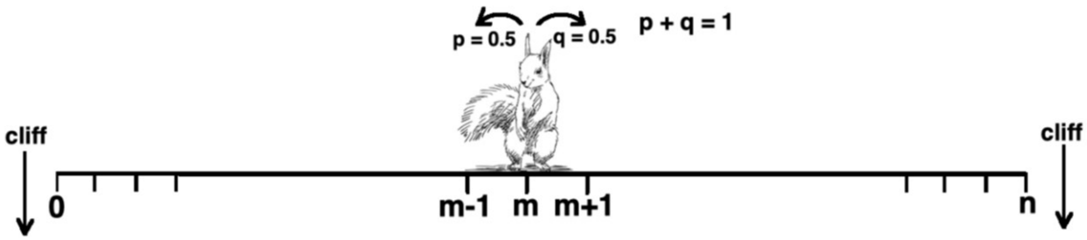
On observing a defective bolt, a naïve conclusion based only on prior belief may be that there is a 30% chance that this part came from Machine 2 because that machine makes 30% of the parts.

However, application of Bayes' Rule gives a much more accurate posterior probability estimate (37.5%) that the defective part came from Machine 2.

Bayesian inference, therefore, enables a much better predictive knowledge of a phenomenon by synthesizing information and data from actual observation and experience.

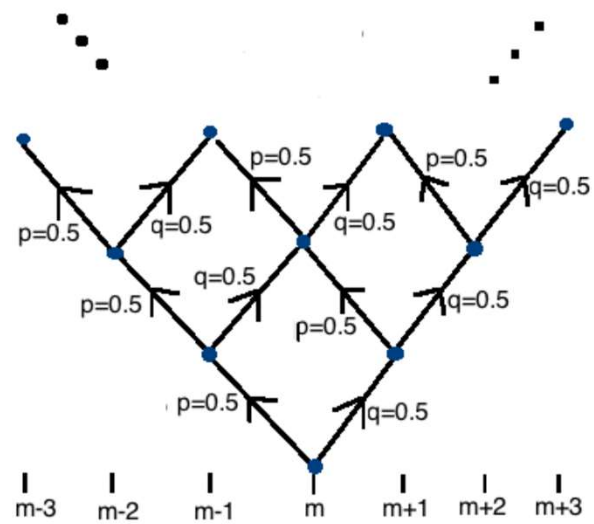
It would be an interesting experience to look back on this calculation and argue logically on why this happened.

Misadventures of Squeaky ………. *continued………..*



1. Squeaky has no "bias", so jumps to the left or to the right are equally likely.
2. However, each jump is of only one step.

Spanning Tree showing possible decision paths starting from location *m*



Squeaky's Demise

Once Squeaky falls off the cliff, he stays there for ever!!

m=0 *falls in the left pit*

or *m*=*n* *falls in the right pit*

Misadventures of Squeaky ………. *continued*………..

W : Event that Squeaky falls into the left pit

$P_m = P_m$(left pit)= P(Squeaky falls into the left pit starting at $X_0=m$) = $P(W|X_0=m)$

$P_0=1$, $P_n=0$ obviously

E : Hop to the Left, E^C : Hop to the Right

Using the law of total probability and conditioning on whether the jump is to the left or right, we get –

$$
\begin{aligned}
P_m &= P(W,E|X_0=m) + P(W,E^C|X_0=m) \\
&= P(W|E, X_0=m)P(E|X_0=m) + P(W|E^C, X_0=m)\,P(E^C|X_0=m) \\
&= P(W|X_1=m\text{-}1)(0.5) + P(W|X_1=m+1)(0.5) \\
&= 0.5P_{m-1} + 0.5P_{m+1}
\end{aligned}
$$

Note that –

E, $X_0=m \rightarrow X_1=m-1$

and E^C, $X_0=m \rightarrow X_1=m+1$

Rewriting $P_m=0.5P_{m-1} + 0.5P_{m+1}$, we get $P_{m+1} - P_m = P_m - P_{m-1}$ for $m=1,\ldots, n-1$ with $P_0=1$, $P_n=0$

Solving Directly! …… works for this problem but may become difficult for a general problem

Writing these equations explicitly from m=1, 2, ….., n-1 we get –

$$\left.\begin{array}{l} P_2 - P_1 = P_1 - P_0 = P_1 - 1 \\ P_3 - P_2 = P_2 - P_1 = P_1 - 1 \\ \ldots\ldots \\ P_n - P_{n-1} = P_1 - 1 \end{array}\right]$$

Summing the LHS and RHS of these $(n-1)$ equations, we get

$$0 - P_1 = (n-1)(P_1 - 1) \quad \text{because } P_0 = 1,\ P_n = 0$$

$$\Rightarrow \quad P_1 = 1 - \frac{1}{n},\ P_2 = 1 - \frac{2}{n}, \ldots\ldots, P_m = 1 - \frac{m}{n}, \ldots\ldots, P_{n-1} = 1 - \frac{n-1}{n}$$

Note that if P_m is the probability of falling in the Left Pit, starting from position m, then the probability of falling in the Right Pit starting from that position, is 1-P_m.

We can also get the same result by repeating the above derivation for the event that Squeaky falls into the Right Pit (Try it!)

This is because Squeaky is destined to eventually fall either in the left pit or in the right pit starting from any position m.

Interesting Question – Can Squeaky nevertheless "live to infinity" like Ashwathama?
What do you think?

Solving as a **Homogenous Linear Recurrence** $f(m) = a_1 f(m-1) + a_2 f(m-2) + ... + a_d f(m-d)$

where $a_1, a_2 \; .... a_d, d$ are constants

Substituting the guess, $f(m) = x^m$, gives and dividing this by x^{m-d} gives the

$$x^m = a_1 x^{m-1} + a_2 x^{m-2} + \cdots + a_d x^{m-d}$$

Characteristic Equation for this

$$x^d = a_1 x^{d-1} + a_2 x^{d-2} + \cdots + a_{d-1} x + a_d$$

The solutions to the homogenous linear recurrence are defined by the roots of the characteristic equation. Neglecting boundary conditions, we have the following –

- If r is a non-repeated root of the characteristic equation, then r^m is a solution to the recurrence
- If r is a repeated root with multiplicity k then $r^m, mr^m, m^2 r^m, \ldots\ldots, m^{k-1} r^m$ are all solutions to the recurrence
- Every linear combination of these solutions is also a solution

To use this approach, we write $P_m = 0.5P_{m-1} + 0.5P_{m+1}$ as $P_{m+2} - 2P_{m+1} + P_m = 0$

The Characteristic Equation for this (on trying a solution of the form x^m) is $x^2 - 2x + 1 = 0$

This has a double root at $x = 1$ $\Rightarrow$ the homogenous solution has the form

$$P_m = a(1^m) + bm(1^m) = a + bm$$

The Boundary Conditions for this are $P_0 = 1$ and $P_n = 0$ gives $a = 1, b = -\frac{1}{n}$

Therefore, P(Squeaky falling into the left pit starting from m) = $1 - \frac{m}{n}$ $\quad m = 0,1, \dots, n-1, n$
and P(Squeaky falling into the right pit starting from m) = $\frac{m}{n}$

Note that, as you would logically expect, Squeaky has a higher probability of falling in the left pit if she starts from a position closer to the left pit (and vice versa)

What happens if $p \neq q$ (but, of course, $p + q = 1$) ?

In that case, we get –

$$P_m = \frac{\left(\frac{1-p}{p}\right)^n - \left(\frac{1-p}{p}\right)^m}{\left(\frac{1-p}{p}\right)^n - 1} \qquad m = 0, 1, \ldots\ldots, n$$

So, poor Squeaky is still sure to die

Try obtaining this, both by the Direct Calculation Method
and by Solving the Homogenous Recurrence

We now know that Squeaky will eventually
Stop Squeaking
but for how long will he still be with us?

To calculate Squeaky's Lifetime, we need to learn few other things first –

1. How to calculate expectations of random variables?

2. How to solve a Non-homogenous Recurrence equation?

Weekend Diversion!

Remember Signor Fibonacci (1170-1240)! He learnt the secrets of ZERO from the Ancient Indians (actually via the Arab mathematicians) and taught the West the joys of having nothing!!

He is also famous for giving the world the Fibonacci Sequence

$$f(n) = f(n-1) + f(n-2) \ldots \text{for } n \geq 2 \text{ and } f(0) = 1, f(1) = 1$$

It took SIX centuries for mathematicians to figure out a solution to this of the form

$$f(n) = ab^n + cd^n$$

> You can of course do this in a few hours on a weekend afternoon (less, if you know how to Google) but once you see the solution, you can appreciate why it took so much time for the world to get there!

The Fibonacci Sequence $f(n) = f(n-1) + f(n-2)$ ….for $n \geq 2$ and $f(0) = 1, f(1) = 1$

has the solution -

$$f(n) = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^{n+1}$$

Expected Values of Random Variable *X*

(First Statistical Moment) μ_X or μ Mean

The expected value of a random variable may be computed as –

$\mu_X = E(X) = \sum_{x \in range(X)} xP(X = x)$ *may be simply written as* μ

(Second Statistical Moment) σ_X^2 or σ^2 Variance Var(*X*)

The variance of a random variable may be computed as –

$$\sigma_X^2 = E\left((X-\mu)^2\right) = \sum_{x \in range(X)} (x-\mu)^2 P(X = x)$$

We also have the useful relationship $Var(X) = E(X^2) - \mu_X^2$

Some other useful results

$E(cX) = cE(X)$ where *c* is a constant

$E(X + c) = E(X) + c$ where *c* is a constant

$E(X + Y) = E(X) + E(Y)$

$Var(cX) = c^2 Var(X)$ where *c* is a constant

$Var(X + c) = Var(X)$ where *c* is a constant

$Var(aX \pm bY) = a^2 Var(X) + b^2 Var(Y) \pm 2abCov(X, Y)$ where *a, b* are constants

and $Cov(X, Y) = E\big((X - \mu_X)(Y - \mu_Y)\big)$ *covariance of X and Y*

Example: *Expectation of an **Indicator** random variable*

As defined earlier, an **Indicator Random Variable** as a random variable which takes binary values 1/0 for TRUE/FALSE.

$$\mathbb{I}_{A(\omega)} \equiv \mathbb{I}_{(\omega \in A)} = \begin{cases} 1;\ \omega \in A \\ 0;\ \omega \notin A \end{cases}$$

Therefore, it follows that –

$$E(\mathbb{I}_A) = 1 \times P(A) + 0 \times P(A^C) = P(A)$$

Example: *Expected number of new recruits per n hiring interviews*

Problem Statement: Let us consider that a hiring manager has the responsibility of conducting interviews of n candidates for the post of a service executive over a certain period of time.

The candidates appear for the interviews in a random fashion, i.e. from the perspective of the hiring manager, prior to the interview, there is an equal probability among candidates to be the most suitable candidate.

The hires are made on a rolling basis in the sense that whenever he encounters a better candidate than the existing one, he hires that person and keeps him in the job until a better candidate is found.

How many hires are made in this process?

Can we give an estimate of the cost associated with this *firing-recruiting* process?

Example: *Expected number of new recruits per n hiring interviews …… continued …*

Define an indicator variable as $\mathbb{I}_i = \begin{cases} 1 & \text{when the } i^{th} \text{ candidate is hired} \\ 0 & \text{when the } i^{th} \text{candidate is not hired} \end{cases}$

The i^{th} candidate is hired if he/she is better than the preceding $(i - 1)$ candidates. The probability of this is $p_i = \frac{1}{i}$ since each of these i candidates have an equal chance to be hired.

The total number of hires is then $X = \sum_{i=1}^{n} \mathbb{I}_i$

with $$E(X) = E(\sum_{i=1}^{n} \mathbb{I}_i) = \sum_{i=1}^{n} E(\mathbb{I}_i) = \sum_{i=1}^{n} p_i = \sum_{i=1}^{n} \frac{1}{i} = \log(n) + \mathcal{O}(1)$$

This means that for every n interviews conducted by the hiring manager, approximately $\log(n)$ of them get hired on an average. The cost of the recruitment process is $\mathcal{O}(c_H \log(n))$ where c_H is the hiring cost for every new hire.

$$\gamma = \lim_{n\to\infty} \left(-\log n + \sum_{i=1}^{n} \frac{1}{i} \right) = 0.5772$$ Euler's Constant or Euler-Mascheroni Constant