Study of Random Walks

Arunabh Ghosh, *IIT Bombay,*

***Abstract*—A random walk is a stochastic or random process, that describes a path that consists of a succession of random steps on some mathematical space. The objective of this project is to study the properties of these random walks and use it to model real life processes and compare it to practical simulations.**

***Keywords*—Markov Chain, Stochastic**

I. Random walk in 1D space

A. *Variation of distribution with number of steps*

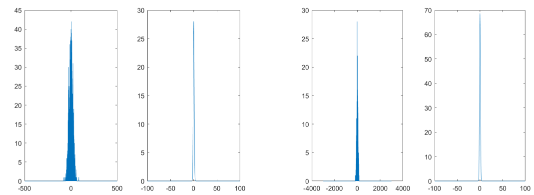
The simple random walk process is a minor modification of the Bernoulli trials process. Nonetheless, the process has a number of very interesting properties, and so deserves a section of its own. On simulating, we observe as number of steps increases, the position distribution of the particle tends to a normal distribution as expected by **Central Limit Theorem**.

```
p = 0.5; % The probability of +1
number_of_steps = 100000:100:100000;
[1, r] = size(number_of_steps);
number_of_simulations = 10000;
steps = [-1,+1];
size_of_steps = 2;
c = 1;
figure

for k=1:r
    num_line = -number_of_steps(k):number_of_steps(k);
    disp(number_of_steps(k));
    frequency = zeros(1, number_of_steps(k)*2 + 1);

    for i=1:number_of_simulations
        pos = 0;
        for j = 1:number_of_steps(k)
            index = randperm(2, 1);
            pos = pos + steps(index);
        end
        frequency(pos+ number_of_steps + 1) = frequency(pos+ number_of_steps + 1) + 1;
    end

    subplot(1, 2, 1)
    plot(-20:20, frequency(number_of_steps(k) - 20 + 1:number_of_steps(k) + 20 + 1));

    subplot(1, 2, 2)
    fplot(@(x) sqrt(number_of_steps(k)/2*pi)*exp(-x^2/2), [-100, 100]);
end
```

B. *At t=∞*

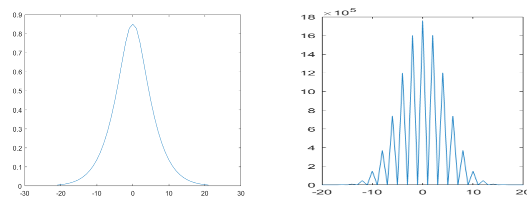
The position of distribution of the particle is calculated by the following equation -

$$P(x=a) = p^a[\sum_{i=0}^{\infty} \binom{a+2i}{i} (p(1-p))^i] \quad (1)$$



(a) 500 steps (b) 3000 steps

Fig. 1: Comparison with the normal distribution



(a) Mathematical Analysis (b) Practical Analysis

Fig. 2: Position distribution at t=∞

II. Random walk in 2D space

Let the particle now be in a 2D plane, with equal probability of going in any direction. We will first calculate mathematically the position distribution of the particle and a simulation would be run to validate the results.

A. *Mathematical Analysis*

Let's analyze the x coordinate of the particle. In one step, the particle moves by 1 unit (magnitude) moves in any direction (uniform distribution). Let **p** denote the position of the particle. In one step, the update step is:

$$\mathbf{p} = \mathbf{p} + [cos(\theta)\hat{x} + sin(\theta)\hat{y}] \quad (2)$$

where θ is the angle that the direction makes with the x-axis. Therefore the x coordinate after **N** steps would be:

$$X = \sum_{i=1}^{N} cos(\theta_i) \quad (3)$$

$cos(\theta)$ is a function of random variable θ which is uniformly distributed from 0 to 2π. Using the standard probabilistic methods, the distribution of Y = $cos(\theta)$ is:

$$f_Y(y) = \frac{1}{\pi\sqrt{1-y^2}} \quad (4)$$

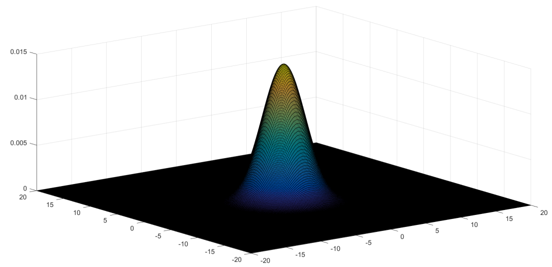


Fig. 3: Mathematical Analysis

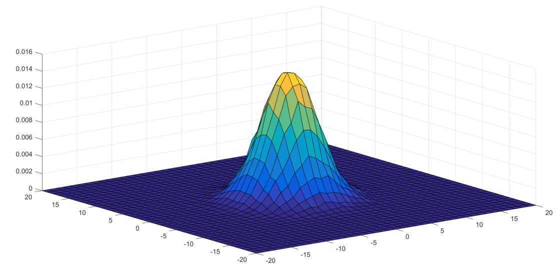


Fig. 4: Practical Analysis

The distribution of the X coordinate is nothing but the distribution of the sum of independent and identically distributed random variables. Therefore the distribution of X is given by the **recursive convolution** of distribution functions.

$$f_X(x) = f_Y * (f_Y * ...(f_Y))(x) \tag{5}$$

We can assume the same distribution for the Y-coordinate by symmetry. The 2D-plot is shown in Figure 3.

```
% All directions are equally possible.
number_of_steps = 20;
number_of_simulations = 100000;
frequency = zeros(2*number_of_steps + 1);

% The directions the particle can go in.
direction_angles = 0:359;

% Mathematical Proof
syms y
f = 1/pi*sqrt(1 - y^2); % The distribution function for X = con(theta)

x = -0.9:0.1:0.9;
f_values = double(subs(f, x)); % The value of the function at various points.

convolved_function = f_values;
% Perform convolution n times.
for i=1:number_of_steps-1
    convolved_function = conv(f_values, convolved_function);
end

[m, n] = size(convolved_function);

prob_surf = convolved_function.' * convolved_function;

rhs = (n-1)/2;
scaling_factor = rhs/number_of_steps;
nl = -rhs:rhs;
nl = nl/9;
x = nl;
y = nl;
z = x*9 + rhs + 1;
surf(x, y, prob_surf(floor(x*9 + rhs + 1), floor(y*9 + rhs + 1))/(125*10^24));
```

B. Practical Simulation

On simulating the random walk of a particle (20 steps) and running it over 100000 times, this is the distribution we get, which is pretty close to the one obtained by a mathematical analysis.

```
% All directions are equally possible.
number_of_steps = 20;
number_of_simulations = 100000;
frequency = zeros(2*number_of_steps + 1);

% The directions the particle can go in.
direction_angles = 0:359;

for i=1:number_of_simulations
    initial_pos = [0, 0];
    for j=1:number_of_steps
        angle = randperm(360, 1);
        initial_pos(1) = initial_pos(1) + cos(angle);
        initial_pos(2) = initial_pos(2) + sin(angle);
    end
    pos_x = floor(initial_pos(1));
    pos_y = floor(initial_pos(2));
    relative_x = pos_x + number_of_steps + 1;
    relative_y = pos_y + number_of_steps + 1;
    frequency(relative_x, relative_y) = frequency(relative_x, relative_y) + 1;
end

x = -number_of_steps:number_of_steps;
y = -number_of_steps:number_of_steps;

figure
surf(x, y, frequency(x + number_of_steps + 1, y + number_of_steps + 1)/number_of_simulations);
```

III. Lord Rayleighs result in D dimensions

In probability theory and statistics, the Rayleigh distribution is a continuous probability distribution for positive-valued random variables. The probability distribution is given by -

$$f(x;\sigma) = \frac{x}{\sigma^2} e^{-x^2/\sigma^2} \tag{6}$$

In this section how the theory of random walkers can be used to model a In this section, we will used the theory of random walks to derive a generalization of Lord Rayleighs result. Beyond 100 steps, Lord Rayleighs result is rather accurate in describing the distribution. It is impressive how the complicated collection of random walkers tends toward a simple, smooth distribution, at least in the central region.

Consider a random walker, who initially starts at the origin in dimensions. At each step, the walker moves by an amount Δ_N , chosen from a probability distribution *p(**r**)*. We have assumed each step is independently and identically distributed and each direction is isotropic so *p(**r**)* is a function of radial distance only. After N steps, let $P_N(\mathbf{R})$ be the probability distribution function. Thus the following probability distribution function can be given:

$$P_{N+1}(\mathbf{R}) = \int p(\boldsymbol{r}) P_N(\mathbf{R} \text{ - } \mathbf{r}) d^d\mathbf{r} \tag{7}$$

As N $\rightarrow \infty$, $P_N(\mathbf{R})$ varies on scales much larger than R, therefore we Taylor expand inside the integral to obtain:

$$P_{N+1}(\mathbf{R}) = P_N(\mathbf{R}) + \frac{\mathbf{r}.\mathbf{r}}{2d}\nabla^2 P_N(\mathbf{R}) + ... \tag{8}$$

If we assume steps are taken at intervals of Δt, and defining time as $t = N\Delta t$:

$$\frac{P_{N+1}(\mathbf{R}) - P_N(\mathbf{R})}{\Delta t} = \frac{\langle r^2 \rangle}{2d\Delta t}\nabla^2 P_N + ... \tag{9}$$

As N $\rightarrow \infty$, the limiting distribution satisfies:

$$\frac{\partial \rho}{\partial t} = D\nabla^2 \rho \tag{10}$$

Solving this differential equation in the Fourier domain and applying the inverse Fourier transform we get:

$$\rho(\mathbf{R}, t) = \frac{e^(R^2/4DT)}{4\pi D t^(d/2)} \tag{11}$$

In the discrete case, as N $\rightarrow \infty$,

$$P_N(\mathbf{R}) \approx \frac{e^{-dR^2/2\langle r^2 \rangle N}}{(2\pi \langle r^2 \rangle N/d)^{d/2}} \tag{12}$$

When the dimension d = 2, and the $\langle r^2 \rangle = a^2$, the distribution becomes:

$$P_N(\mathbf{R}) \approx \frac{e^{-R^2/a^2 N}}{\pi a^2 N} \tag{13}$$

which agrees with with Lord Rayleighs solution. Thus we have shown that Lord Rayleigh's distribution can be accurately modelled using a large collection of random variables. It is impressive how the complicated collection of random walkers tends toward a simple, smooth distribution, at least in the central region. This can be used to model the diffusion of gases where the expected distance travelled scales according to the square root of the number of steps.

REFERENCES

[1] https://www.stat.washington.edu/ñehemyl/files/UW_MATH-STAT395_functions-random-variables.pdf

[2] https://www.dartmouth.edu/ chance/teaching_aids/books _articles/probability_book/Chapter7.pdf

[3] https://en.wikipedia.org/wiki/Random_walk

[4] http://nebula.physics.uakron.edu/dept/faculty/jutta/modeling/diff_eqn.pdf