The overview of random walk and its application

Xiaofeng Xu*

Department of Mathematics, University of California, Davis

*Corresponding author: xfxxu@ucdavis.edu

Keywords: Random Walk, Stochastic Processes, Markov chain, Self-Avoiding Walk, Spacey Random Walks, Bernoulli Excursion, Gaussian Random Walks.

Abstract: Stochastic processes are regarded as a dynamic component of probability theory. By understanding the theory of stochastic processes, people can investigate random phenomena under a given time basis. Random phenomenon concept and investigation is vital in the development of mathematical models. Generally, stochastic processes represent a group of random variables that have been indexed against another variable or set of factors. Poisson phenomena, for instance, the radioactive decay, Markov processes, and time series, with the indexing parameter relating to time, are some of the most fundamental forms of stochastic processes. The type of changes in the variables concerning time is the focus of this indexing, which might be continuous or discrete. The random walk phenomena are an example of a Markov chain process I will be researching in this paper. Although the terminology is most commonly used to refer to a subcategory of Markov chains, several time-dependent phenomena are referred to as random walks, with a prefix emphasizing their unique qualities. Random walks may be used to represent a variety of phenomena in humans and animals, such as diffusion, interactions, and attitudes, as well as to extract information about key entities or dense clusters of entities in a network. For decades, random walks have been investigated on regular lattices and networks with various architectures. This research paper shifts the focus to multiple types of random walks, processes, and their applications in statistics and real life.

1. Introduction

In research from Masuda, the term was introduced by Karl Pearson during the period of analysis of the "Gambler's ruin "problem by several scholars such as Bernoulli, Pascall, and others [1]. Scientist Albert Einstein discovered the stochastic processes in the form of Brownian motion phenomena of collision of particles in continuous time due to a collision with particles. Probability theory has been a great contributor to the theoretical developments on stochastic processes such as random walks, and other contributors include computer science, operational research, and statistical physics. If T is a set known as the index set (which may be conceived of as time), then a stochastic process is a set or collection of random variables $X(t),\ t \in T. X(t)$ is a stochastic model with a discrete variable if T is a denumerable endless sequence. $X(t)$ is a stochastic model with a continuous variable if T is a discrete or continuous interval. T is the period included in the definition above, and $X(t)$ is the measurement at period t [2].

Random Walks are a typical example of a Markov chain. However, not all random walks can be referred to depict the Markov chain. The term Markov chain refers to a statistical model that changes based on statistical criteria from one state to another. A Markov chain is distinguished because the present state exclusively determines the probability of shifting to any given state and the amount of time passed. The state vector, or collection of all potential states, may include anything: characters, numerals, weather patterns, soccer scores, or stock market performance. Random walks models are represented in numerous ways and use various approaches depending on the field the model applies to. This research paper investigates self-avoiding walk phenomena, the random lattice walk, Bernoulli execution, Gaussian random walks, and spacey random walks. Each of the above random approaches has its applications in real life and the field of study.

Published by CSP © 2021 the Authors

The benefits accrued to random walks and their applications are numerous. RW models have also been used in a variety of fields, including animal movement patterns and foraging, neurological firing dynamics, and decision-making in the brain, polymer chains, development of the field of study using RWs on problem space, ranking systems, phylogenetic analysis, regression problems and extraction of features from high-dimensional data, such as dispersion maps, stock market descriptions, and even sports metrics. In the subsequent sections, I will review how the concept is applied in the above-mentioned applications.

1.1 Stochastic model

A stochastic process is a collection of random variables indexed by a mathematical set. A random walk is a stochastic process that is commonly characterized as the sum of independent and identically distributed random variables or random vectors in Euclidean space, which implies that it is discrete time. Each probability and random process has its own element in the set. The index set indexes the random variables. The index set was originally a subset of the real line, interpreting time. The stationary distribution of a non-Markovian stochastic process is provided by the tensor eigenvector's spacey random walk. Random walk with vertex reinforcement is a vertex-reinforced random walk with continuous dynamics.

Random Process the result, i.e., the observed value at each time, is a random variable. According to Karl Pearson Each random variable in the same mathematical space, known as the state space. As an example, this state-space is an integer or a real line. The increment of a stochastic process is the difference between two index values, which are often regarded as two points in time. Due to inherent unpredictability, a stochastic process might have multiple outcomes, and one of those possibilities is known as a sample function or realization. It is said to be discrete time if its index set comprises a limited or countable number of items, such a finite set of numbers, integers, or natural numbers. Time is continuous if the index specified is a genuine line interval. There are two kinds of stochastic processes: discrete-time and continuous-time. Because the index set is uncountable in continuous-time stochastic processes, discrete-time stochastic processes are regarded simpler to understand. The stochastic process is also known as a random sequence if the index set is made up of integers. A discrete or integer-valued stochastic process has a state space made up of integers or natural numbers. A real-valued stochastic process or a process with continuous state space has a state space that is the real line. The stochastic process is called a vector process if the state space is dimensional.

2. Methods OF RANDOM WALKS

2.1 Simple Random Walk

Simple random walk is the most basic model for random walks on numeral lattice Z. A random walk is a collection of independent and identically distributed random variables S_r, with $S_0 = 0$, defined as

$$S_r = Y_r \quad k = 1\, Y_k \tag{1}$$

where Y_k are independent and identically distributed random variables (i.i.d.). If $Y_k = \pm 1$ then the random walk is simple, with

$$P(Y_k = 1) = p \text{ and } P(X_k = -1) = 1 - p = q \tag{2}$$

If each of the nearby nodes has the same probability, the object is categorized as symmetric. Simple random walk on Z: As the name implies, this is the most basic of all random walks. Here, X_1 has values in the range +1, -1, and the walk S_n, which began at 0, is therefore restricted to the subset of all numerals Z. X_1 frequently takes both values with equal probability,

i.e. $$P\,(X_1 = 1) = P(X_1 = -1) \tag{3}$$

The walk then alternates between jumping left and right. The "simple symmetric random walk" is a better name for this instance, although the term "symmetric" is generally often deleted. The walk is called biased in the other instances, for as when $P(X_1 = 1)$ with p = $\frac{1}{2}$. When p $> \frac{1}{2}$, the skew will be to the right, and when $p < \frac{1}{2}$, the prejudice will be to the left. Consider a person or a particle walking along an axis, each discrete time step moving one unit to the right (with probability) or one unit to the left (with probability). The walker may decide whether to go right or left at each step by tossing a coin with the possibility of landing on heads.

2.2 Lattice Random Walk

The lattice random walk model is one that walks randomly on a regular grid, in which at each step the location jumps to another site according to some probability distribution. In a simple random walk, this position can only jump to adjacent positions of the grid, forming a grid path. In a simple symmetric random walk on a locally finite lattice, the probability of jumping to the position of each of its immediate neighbors is the same. The best studied example is a random walk on a D-dimensional integer lattice (sometimes called a hypercube lattice). If the state space is limited to a finite dimension, the random walk model is called a simple boundary symmetric random walk, and the transition probability depends on the position of the state, because in the edge and corner states, the motion is finite. A lattice is a structure of points arranged in a certain order. When modeling random walks on lattices, the dimensions differ on each type of random lattice walk. Lattice random walks range from one-dimension, two-dimension, and three-dimensional lattice walks. The three approaches seek to generate a formula to calculate the dispersion and the average first passage instances of an element in a random walk from the initial point to an arbitrary lattice point on a time-bound lattice with given boundary conditions. Usually, the amount of time is equal in proportion to lattice points quantity. When n is big, the number of different locations visited after M steps on an M-dimensional lattice:

$$(with\, m \geq 3)\;\; a_1 n + a_2\, n^{1/2} \; + \; a_3 + \; a_4 n^{-1/2} + \cdots \tag{4}$$

When $m \; = \; 3$, the constants a_1 and a_2 have been provided for walks on a simple cubic lattice, and a_1 and a_2 have been provided for simple and face-centered cubic lattices, respectively.

2.3 One Dimensional Lattice Random Walk

When evaluating random walks on a line, i.e., within a one-dimensional space, the possible random walk models are continuous-time and discrete-time models. Discrete models are characterized by a change in state variables only at a measurable number of points in a given period, while in continuous models. In contrast, in continuous models, the change in state variables is constant rather abrupt. A series of random variables $X_n\colon (\Omega, \mathcal{F})(\mathbb{R}, \mathcal{B})$ indexed by $n \,\epsilon\, \mathbb{Z}_+$, where $\mathcal{B}$ is the Borel-field, constitutes a discrete-time real-valued random walk. Such sequences are written as $(X_n,\; n \geq 0)$, with the assumption that the duration reference n is an integer.

2.4 Two Dimensional Lattice Random Walk

Random walk models in two dimensions and three-dimension space follow identical laws, with the exception that in two dimensions and three dimensions, the object is free to wander over a two-dimensional or three-dimensional space, respectively. However, the total distance traveled from the beginning position is about $\sqrt{N}$, whereby N represents the total of steps taken—considering a summation of N random-orientated double vectors on a two-dimensional plane. Let each vector's phase be randomized using phasor notation. Assume you're going in either direction with N unit steps. A discrete-time authentic random walk is defined as a sequence of random variables $X_n\colon (\Omega, \mathcal{F})(\mathbb{R}, \mathcal{B})$ indexed by$n \,\epsilon\, \mathbb{Z}_+$, where B is the Borel-field, which gives the location z in the complex plane after N steps. Such sequences are expressed as $(X_n,\;\; n \geq 0)$, with the assumption that the duration reference n is an integer. A discrete-time real-valued random walk is composed of a series of random variables $X_n\colon (\Omega, \mathcal{F})(\mathbb{R}, \mathcal{B})$ indexed by $n \,\epsilon\, \mathbb{Z}_+$, where B is the Borel-field. With the presumption that the period reference n is an integer, such sequences are expressed as $(X_n,\;\; n \geq 0)$.

$$Z = \sum_{j=1}^{N} e^{i\theta_j} \tag{5}$$

Studies show that when the number of iterations tends to infinity, a random walk on a two-dimensional lattice has a unity chance of reaching any point [3].

2.5 Gaussian Random Walk

These are random walks with a normal probability distribution. It is asymmetric probability distribution around the mean, indicating that data close to the average occur more often than data distant from the mean. The normal distribution will show as a bell curve on a graph. The Central Limit Theorem motivates the Gaussian model. According to this theory, means derived from distinct, equal variance random variables have nearly normal distributions, irrespective of the kind of distribution used to sample the data as long as it has a limited variance. The components of the classical Gaussian distribution are conventionally denoted by the letter Z, implying that Z equals $N\epsilon[0,1]$. Mode, mean and median terms are essential in Gaussian distribution. The mean results from averaging all the figures, the midpoint of the distribution is the median while mode represents the value most frequent during observation [4]. In a gaussian walk model, the figures of median, mean, and mode are almost similar. The diagram below shows a gaussian or normal distribution indicating value percentages within a standard deviation from the mean.

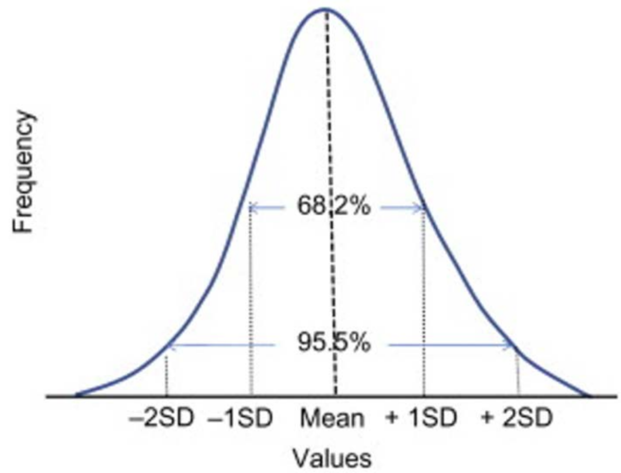

Figure 1. Gaussian Distribution showing percentage values within a standard deviation from mean

Generally, the Gaussian random walk is the sum of a series of independent identically distributed fuzzy numbers, X_i from the reverse cumulative standard deviation having an average equal to zero and from the initial reverse cumulative standard deviation.

$$Z = \sum_{i=0}^{n} X_i \tag{6}$$

However, the distribution for the sum of two equal variances independent random variables,

$$Z = X + Y, is\ obtained\ by\ (\mu_X + \mu_Y, \sigma^2{}_X + \sigma^2{}_Y) \tag{7}$$

$$\mu_X = \mu_Y = 0 \text{ and } \sigma^2{}_X = \sigma^2{}_Y = \sigma^2 \tag{8}$$

In this case yield $\mathcal{N}\ (0, 2\sigma^2)$ (9)

For a Gaussian random walk, this is just the standard deviation of the translation distance distribution after n steps. Therefore, if μ is equal to zero, and since the root mean square (RMS) translation is a standard deviation, the root mean square after n steps will fall to $\pm\sigma$. Similarly, the translation distance after n steps has a 50% chance of falling between $\pm 0.6745\ \sigma$.

3. Applications AND DISCUSSION

3.1 Self-Avoiding Random Walk

A self-avoiding walk is a path that never overlaps itself as it moves from one point to another. Because such routes are typically thought to occur on lattices, steps are limited to a certain number of orientations and lengths. Self-avoiding walk portrays critical traits in statistical mechanics and provides a model of linear polymers in the polymer science field. Critical exponents that have a descriptive and inferential statistical link with critical exponents in ferromagnetism models such as the Ising and $||4$ spin models are used to represent the critical behavior of the self-avoiding walk. The crucial exponents are widely established in physics, yet they nonetheless pose significant mathematical challenges [5]. Weakly self-avoiding walks and strictly self-avoiding walks are the two main self-avoiding walks. Suppose a self-avoiding walk that begins at the origin makes the first step in a positive horizontal plane, is not limited to nonnegative nodes only, but must undertake an up movement prior to taking the first downwards step. The size of such paths $y = 1,2, \dots$ paths are $1,2,5,13,36$ ....

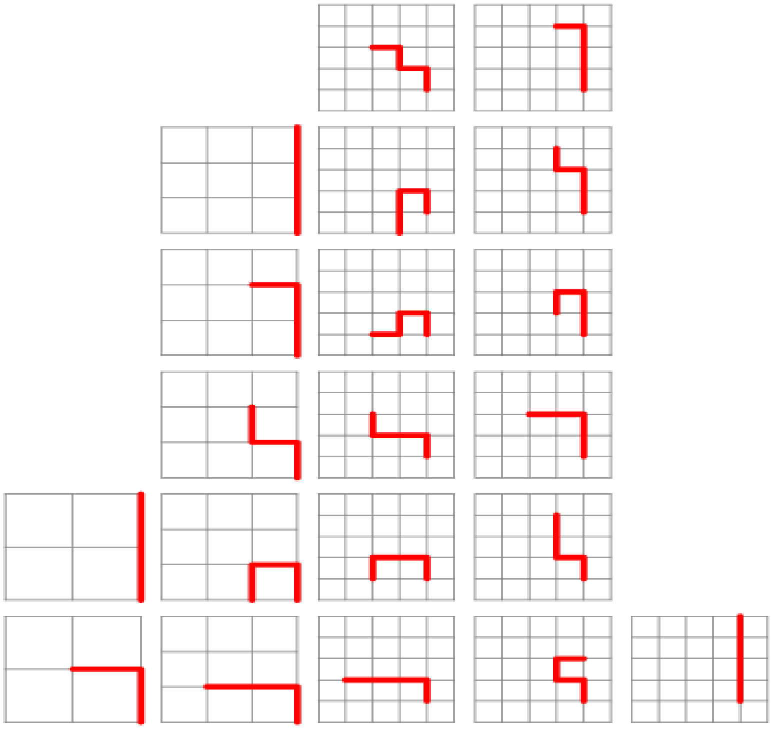

Figure 2. Example of a random walk on a one-dimensional lattice.

3.2 The Spacey Random Walk

When dealing with high order data such as population dispersion, spacey random walk is the go-to stochastic process model. The tensor eigenvector determines the equilibrium distribution of a non-Markovian probabilistic process called a spacey random walk. The operation is a directed graph random walk with discrete behaviors connected to a continuous stochastic process [6]. The convergence features of these dynamics are investigated, and numerical approaches for determining the stationary distribution are discussed. Coppersmith, Diaconis, and Pemantle presented vertex-reinforced random walks, and the spacey stochastic process is a subset of a larger group of vertex-reinforced probability models studied by Bena m. The relationship between a discrete random walk and an active learning system is a key finding from Bena m's research. In essence, the limiting probability of time spent at each node in the network is related to the dynamical system's long-term behavior. This analysis can be much improved using the vertex-reinforced random walk. For example, proving that the appropriate stochastic model for the method always conforms to a stable equilibrium in the exceptional case of a discrete binary system. There provide adequate requirements for a popular approach for incorporating dynamical systems to converge to a stable distribution when the process includes many states.

3.3 Bernoulli Excursion and its Various Applications

Consider a collection of $2n$-element sequences, each with n elements equal to $+1, n$ elements equal to 1, and the total of the first I members larger than or nil for all $i = 1,\dots,2n$. The n-th Catalan number determines the number of such sequences.

$$C_n = \frac{1}{n+1}\binom{2n}{n} \tag{10}$$

$C_0 = 1, C_1 = 1, C_2 = 2, C_3 = 5, C_4 = 14, C_5 = 42$ are the first few Catalan numerals.

From the C_n sequences, a sequence is picked at random, assuming that all potential sequences are equally likely. Set η_i^+ denote the sum of the first I elements in this selected sequence by I ($i = 1,\dots,2n$). After that, for $i = 1,\dots,2n$, $\eta_{2n}^+ = \eta_0^+ = 0$ and $\eta_i^+ > 0$. The sequence $(\eta_0^+,\dots,\eta_{2n}^+)$ denotes a random walk, also known as a Bernoulli excursion. On the x-axis, one may envision a particle taking a random wander. It begins at $x = 0$ and progresses in $2n$ steps. Depending on if the i^{th} character in the random sequence is +1 or 1, the particle travels an equal distance to the right or to the left at the i^{th} step. For $i = 1,\dots,2n$, the particle's location at the conclusion of the i^{th} step is $x=\eta_i^+$ [7].

3.4 Applications of Random Walks

Bernoulli's excursion is applicable in probability theory. Several functions demand the determining of several working of the Bernoulli's excursion. A practical example of Bernoulli execution example is the single-server queues and random trees. $M|M|1$ The dispersion of the maximum queue size throughout a busy time necessitates the identification of the random variable max $(0,\dots,2n)$ In random trees, with $N + 1$ unlabeled vertices, there are Cn complete binary rooted plane trees. Pick a tree at random, assuming that each of the Cn potential trees is equally likely. The random tree's height has the same distribution as max $(0,\dots,2n)$ in the Bernoulli excursion. Self-avoiding random walks are critical in segmenting color images using the noise reduction algorithm. The basis of the algorithm is on a virtual particle performing a self-avoiding walk [7]. Both the segmentation effect achieved by the method and the capacity to get rid of Gaussian noise is what make the method suitable for the task. self-avoiding walks create networks that help in the study of various properties such as low-energy phonon spectral distribution in linear polymers and conductivity of electricity. Spacey random walks have numerous applications for describing dynamics of various processes, i.e., in the genetic analysis of populations, transportation field, poly urn stochastic processes, and clustering and ranking activities. Colony genetics is the study of the mechanisms of gene dispersion in evolutionary psychology.

4. Conclusion

In summary, this paper concludes random walks to be a Markov process in stochastic modeling. Random walk models have several traits that are used in quantifying their aspects such as recurrence, dispersal distributions and rates of encounter. Random walks models on the lattice are grouped based on the degree of randomness applied to each model I.e., the number of dimensions on which the random element is likely to move. Random models are effective in predictive analysis, simulations, mathematical modeling and statistics as well. However random walk models such as the Gaussian model are based on the probability distribution of data in the random model. Spacey model of random walk is grouped on the basis of developing models when handling higher data populations. In all probability theory models the definition of a random walk differs based on context.

Random walks play a vital central role in various fields of study, including network study, statistical analysis, biology, and mathematics. In the network science field, the random walks are the primary methods used to derive information from networked systems since they offer a diffusion processes model. Being linear processes, RWs are subject to analysis. For instance, linear algebra approaches may be used to describe dynamics regarding modes relaxing on various time scales. In this research, this paper has explored the numerous models of random walks that exist and the fields they apply in. Study and research indicate that RW models are applied differently in each field of study, and the approaches used to quantify the probabilities and distributions in each model vary from one field to another. However, the principles and the basis of them providing a framework for the stochastic representation of entities and modeling of phenomena is outstanding in each of them. Generally, they are used to model stochastic processes in various domains and fields of study.

References

[1] Masuda, N., Porter, M. A., & Lambiotte, R., "Random walks and diffusion on networks," Physics Reports, 716–717, 1–58 (2017).

[2] Hurter, A. P., & Kaminsky, F. C., "An Application of Regenerative Stochastic Processes to a Problem in Inventory Control," *Operations Research*, 15(3), 467–472 (1967).

[3] Teixeira, A., & ČErný, J., "From random walk trajectories to random interlacements," Ensaios Matemáticos, 23(1) (2013).

[4] Hornikova, A., "Statistics for the Quality Control Chemistry Laboratory," Technometrics, 48(1), 151 (2006).

[5] Slade, G., "Self-avoiding walk, spin systems and renormalization," Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 475(2221), 20180549 (2019).

[6] Benson, A. R., Gleich, D. F., & Lim, L. H., "The Spacey Random Walk: A Stochastic Process for Higher-Order Data," SIAM Review, 59(2), 321–345 (2017).

[7] Takács, L., "A Bernoulli excursion and its various applications," Advances in Applied Probability, 23(03), 557–585 (1991).

[8] Rieger, J., "Crossover from random walk to self-avoiding walk," Physical Review A, 38(10), 5448–5449 (1988).

[9] Baudoin, F., Demni, N., & Wang, J., "Quaternionic stochastic areas," Stochastic Processes and Their Applications, 131, 311–339 (2021).

[10] Bertacchi, D., Coletti, C. F., & Zucca, F., "Global survival of branching random walks and tree-like branching random walks," Latin American Journal of Probability and Mathematical Statistics, 14(1), 381 (2017).