Random Walk Properties and Visualization in Mathematica

Sean Pitman

December 5, 2018

1 Introduction

Random walks have been applied to many real-world visualizations including stock prices over time, particle motion, bacterial interaction, and online social network interaction. The primary goal of this project is the development of code that will allow users to visualize their own multi-dimensional random walks based on user choice of starting value, number of steps, number of simulations, and step distribution. Definitions and properties of random walks detailed by Dr. Marek Biskup in [1] will guide the implementation and testing of this applet.

2 Background

A pertinent place to start is the definition of a random walk.

Definition 1 (Random walk)**.** *Suppose that $X_1, X_2, \ldots$ is a sequence of $\mathbb{R}^d$-valued independent and identically distributed random variables. A* random walk *started at $z \in \mathbb{R}^d$ is the sequence $(S_n)_{n\geq 0}$ where $S_0 = z$ and*

$$S_n = S_{n-1} + X_n, \quad n \geq 1. \tag{2.1}$$

The quantities (X_n) are referred to as steps *of the random walk.*

That is, S_n describes the position of the random walk at time n, choosing its next step at random. Equation 2.1 could also be written as

$$S_n = z + X_1 + \cdots + X_n \tag{2.2}$$

for each $n \geq 1$. While $X_1, X_2, \ldots$ are independent as random variables, $S_0, S_1, \ldots$ are not independent [1].

3 Mathematica Implementation

The properties of different step distributions manifest visually in different ways depending on the distribution chosen. As such it's useful to be able to consider cases for multiple dimensions. In Mathematica, this visualization is accomplished slightly differently for each dimension, so each implementation is tackled separately before consolidation. All methods make use of Mathematica's various `Distribution[]` functions along with `RandomVariate[]` to sample from them. These values then perform the gradual summation of a random walk using the `Accumulate[]` function.

3.1 One-dimensional Case

The most straightforward task is in the one dimensional case. A list of values using `Table[]` and `Accumulate[]` are directly plotted with the `ListLinePlot[]` function. This is tabled for the desired number of simulations so that multiple runs can be plotted on the same graph by the function.

3.2 Two-dimensional Case

The two dimensional case requires more finesse when using `ListLinePlot[]`. First, two lists of walk values are generated in the same way as the one-dimensional case, but then `Table[]` is used to pair these values into points that `ListLinePlot[]` can use to plot in the Cartesian plane.

3.3 Three-dimensional Case

One more change is made to allow the three dimensional case to draw from the work in the previous dimensions, and that is the use of `Graphics3D[]` to perform the task of `ListLinePlot[]` in three dimensions. Similar to the change in the two-dimensional case, three lists of random walk values are generated and then sequentially paired with `Table[]` before being read by the 3D graphic function `Line[]` with `Sphere[]` being used to indicate the beginning (noted by a blandly colored sphere) and end (noted by a sphere of matching color) of each simulation. The spheres each have radius 0.5.

3.4 Additional Information

When initial coordinate values are calculated, raw values are computed up to 1000 steps, to be used by the `Manipulate[]` function responsible for visualization. Except in the one dimensional case, the internal tables of point values are organized as lists of lists containing each simulation's list of points. This allows the `Graphics3D[]` and `ListLinePlot[]` functions to easily plot sets of points and visually distinguish simulations by color.

4 Results

The success of the applet is gauged on how effectively and accurately it is able to provide visualizations of random walks. Specifically, random walks of different step distributions will have striking visual characteristics and properties that are laid out by Dr. Biskup. A variety of these were tested and compared with the applet's performance.

4.1 Gaussian Random Walk

Referring back to 2.2, the *gaussian random walk* has steps X_i that take any value in $\mathbb{R}$, where $\{X_i\}_{i=1}^n \sim$ *iid* $\mathrm{N}(0,1)$. That is, each step is distributed normal (or Gaussian) with mean zero and variance 1. Explicitly,

$$P(X_i \leq x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx \tag{4.1}$$

Thanks to the Central Limit Theorem, we have some specifics to think about when viewing this sample of random walks in Figure 2.

Theorem 1 (Central Limit Theorem)**.** *Consider a one-dimensional random walk with $\mathbb{E}(X_i^2) < \infty$. Then, as $n \to \infty$,*

$$\frac{S_n - n\mathbb{E}(X_i)}{\sqrt{n}} \tag{4.2}$$

has asymptotically normal distribution with mean zero and variance $\sigma^2 = Var(X_i)$.

As a result, a property of the standard normal walk is that the distribution of S_n is also normal with mean zero but variance $\sqrt{n}$. This means that after n steps, the typical displacement of the random walk from its starting value is "order-$\sqrt{n}$" [1]. This is something the applet can help investigate.

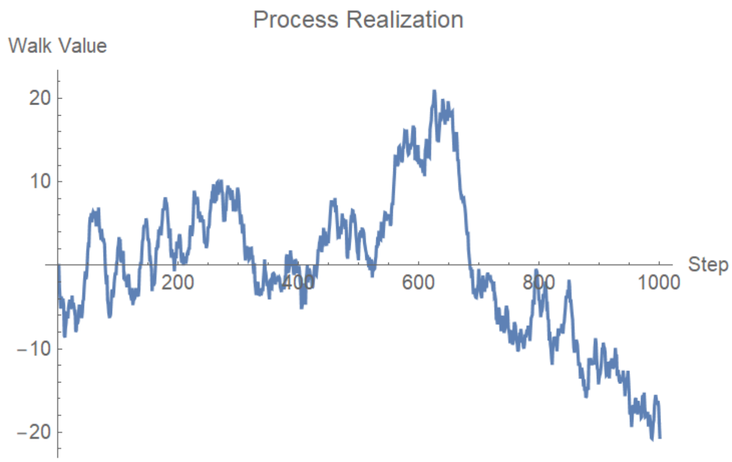

Figure 1: A 1000 step, one dimensional random walk with standard normally distributed steps.

Using the applet, a full run of a 1000 step Gaussian random walk is simulated. Figure 1 displays some of these characteristics. Starting from 0, we see values of S_n wander in a manner fairly centered on 0. Additionally, the largest peak of this run, around 20 near step 600, coincides with our expectation of the typical displacement of S_n being of order-$\sqrt{n}$, in this case being $\sqrt{600} \approx 25$. Of course it makes sense to test this idea with multiple simulations. Using the applet we can look at any desired number of simulations, say 50.

Figure 2: 50 simulations of 1000 step, one dimensional random walks with standard normally distributed steps.

We can see this normal behavior of the endpoint of S_n in play, with most of the values of S_{1000} falling between about -32 and 32 ($\approx \sqrt{1000}$, one standard deviation of S_{1000}) in Figure 2. Higher dimensions of the Gaussian random walk retain specific quirks, mainly of a stumbling path characteristic of the relatively smaller central values the steps tend to take between -4 and 4. The applet allows visualization of these two and three dimensional cases seen in Figures 3 and 4 respectively.

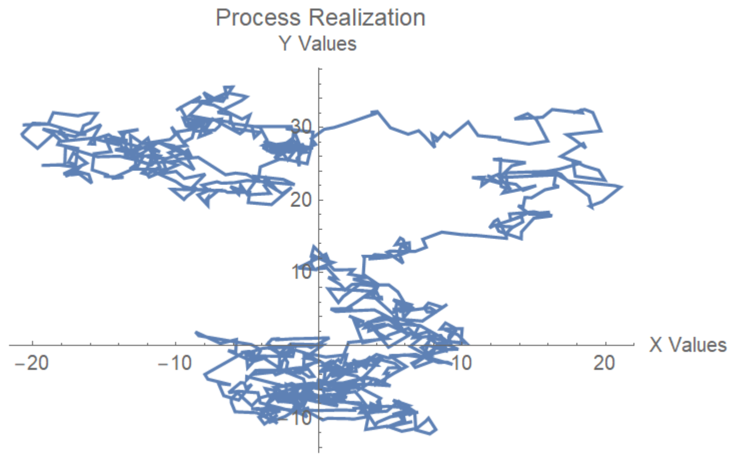

Figure 3: A 1000 step, two dimensional random walk with standard normally distributed steps.

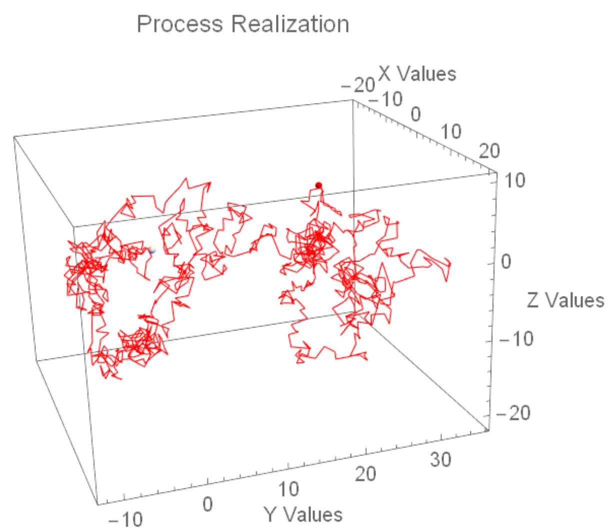

Figure 4: A 1000 step, three dimensional random walk with standard normally distributed steps.

4.2 Heavy tailed random walk

In stark contrast to the Gaussian random walk is the *Cauchy random walk* which retains the symmetry property but has "heavy tails." This property actually contributes to one of the peculiar properties of the Cauchy distribution, in that both its mean and variance are undefined specifically because of this tail behavior. In such a walk, steps will follow a "standard" Cauchy distribution with location parameter 0 and scale parameter 1. That is,

$$P(X_i \leq x) = \int_{-\infty}^{x} \frac{1}{\pi} \frac{1}{1+x^2} \tag{4.3}$$

A defining feature of such walks visually is that of "macroscopic" jumps in a single step that are comparable to the typical distance of the walks from the starting value at the step the jump occurs on [1].

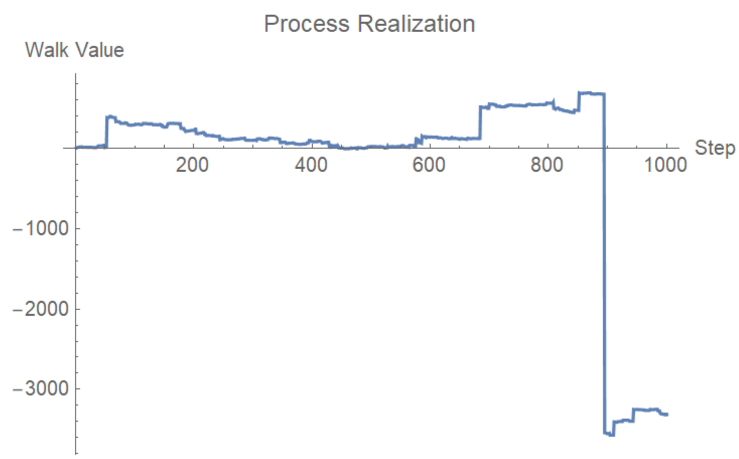

Figure 5: A 1000 step, one dimensional random walk with "standard" Cauchy distributed steps.

We see a characteristic "macroscopic" jump around step 900. These are qualities that make the Cauchy random walk so recognizable regardless of dimension.

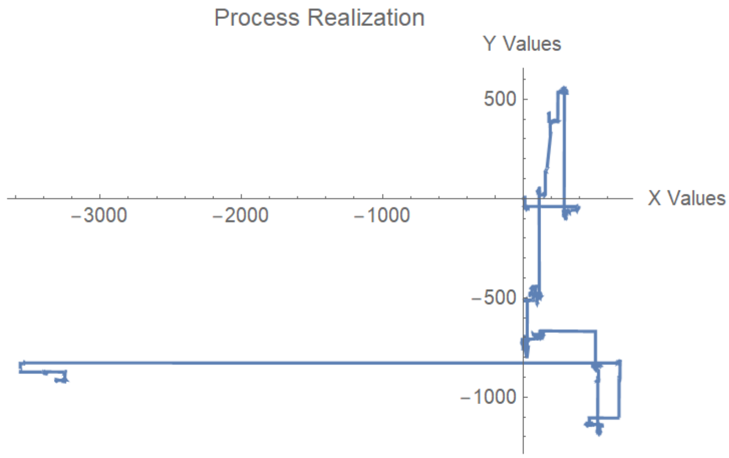

Figure 6: A 1000 step, two dimensional random walk with "standard" Cauchy distributed steps.

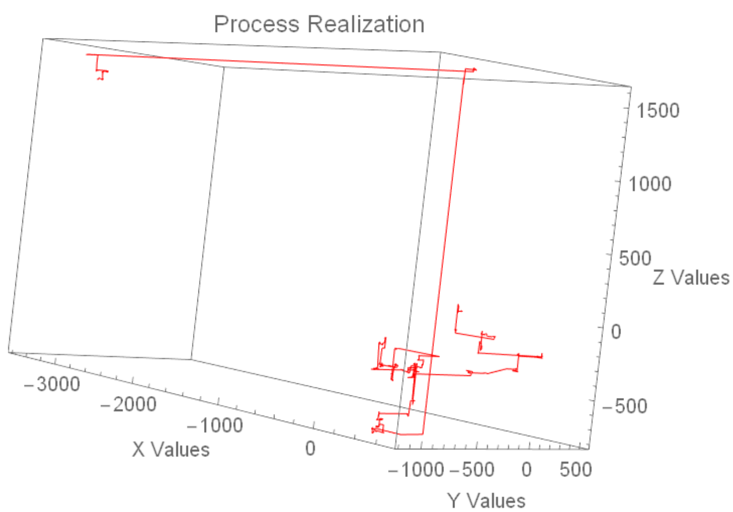

Figure 7: A 1000 step, three dimensional random walk with "standard" Cauchy distributed steps.

Additionally, because of the Cauchy distribution's lack of a second moment, the Central Limit theorem does not apply. The applet can provide more of an experience when looking at many simulations, say 50, as in Figure 8.

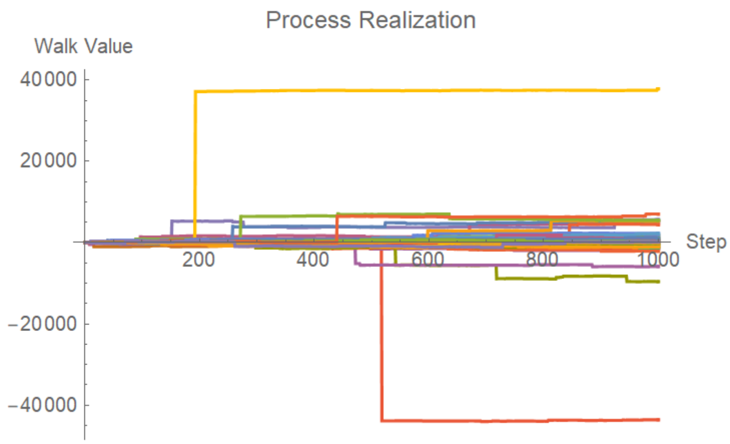

Figure 8: 50 simulations of 1000 step, three dimensional random walks with "standard" Cauchy distributed steps.

We're able to place far less rules on the Cauchy random walk compared to the work on the Gaussian random walk.

5 Conclusions

Striking differences between standard Normal and Cauchy distributions are apparent and visually confirm-able thanks to the Mathematica applet. The Mathematica applet successfully simulates and visualizes random walks in dimensions one, two, and three, providing an excellent way to observe and convey the telling characteristics of probability distributions through random walks.

References

[1] B. Marek. Random walks. In *PCMI Undergraduate Summer School Lecture Notes*, pages 15–20, Park City, Utah, July 2007. University of California, Los Angeles. As part of the Park City Mathematics Institute Outreach Program.