A Second Chance For Random Varaibles

Ali Fele Paranj

February 26, 2025

Abstract

To me, the notion of random variables is something that I learn a new aspect of it as I encounter it in different situations. When I first learned the notion of random variable, it was already been abused by some physicists by the way that taught it to us (and it is most likely to be so to most of the people who learned it from non-mathematicians!). In this note I will explain some aspects of random variables through a running example. After this note you will have a clear understanding of the notion of random variables, martingales, filtration, and you will have an intuitive feeling about the statements like "σ-algebra can be thought of as information".

Random Variables

We start with an example.

Example 1. Consider an experiment in which we toss a coin infinitely many times. The sample space Ω for such experiment will be the $\mathbb{F}_2^{\mathbb{N}}$, i.e. the set of all sequences composed of zeros and ones. We want to show that Ω is "similar" to the interval $(0, 1]$ in interesting ways. Consider the partition $\Omega = A_0 \cup A_1$, where A_0 is the set of all outcomes that start with 0 and A_1 is the set of all outcomes that start with 1. Similarly we can have the partition $\Omega = A_{00} \cup A_{01} \cup A_{10} \cup A_{11}$ where A_{01}, for instance, is the set of all outcomes that start with 01. We can associate with, say, A_{01} a subset of $(0, 1]$ whose the dyadic expansion of all numbers in that interval starts with 01. For instance

$$A_{00} \mapsto (0, 1/4], \qquad A_{01} \mapsto (1/4, 1/2], \qquad A_{10} \mapsto (1/2, 3/4], \qquad A_{11} \mapsto (3/4, 1].$$

The idea of probability theory is to remove all kind of randomness by considering as big state space as possible. For instance, assume you want to do an infinite coin toss experiment. From the probability theory point of view, the moment you start to do your experiment you will be some $\omega \in \Omega$, for which every step of your coin tossing is known. In other words, any combination of outcomes that you might encounter in your experiment is already included in the sample space, thus any particular outcome you might get is a concrete element of the probability space. Now imagine the following story!

Example 2. You are talking with one of your friends over phone and he says while he was speaking he just started random coin toss experiment (with a fail coin!). Then you feel peculiar as you look at the real line on the black board of your office and you know that he is now assigned to some $\omega \in (0, 1]$. But which one? You wish you know the exact point, then you could bet with your friend that you can predict the outcome of all of his coin tosses and you get a lot of money.

So far, you should have believed that anybody anywhere in this universe, once they start a random coin toss experiment, then they are automatically assigned to one of the points $\omega \in (0, 1]$, as by definition Ω is the set of all possible outcomes that when anybody anywhere in this universe, once they start a random coin toss experiment, their result is also included in Ω. But in probability theory, do not care about the outcome of the particular experiment of your friend, but we instead are interested in looking at all of the possible experiments all at once, and make important observations. This is very similar to the idea of ensembles in statistical physics.

Example 3. You friend from previous example tells you that since he started doing the infinite coin tossing experiment, he wants to do something with it. He decides to choose some direction (the same direction for the entire infinite experiment) and then start moving back and forth on that line based on the result of each coin toss. 1 step forward for a heads and 1 step backward for a tails. Once you notice his strange decision you get nervous, because you don't know where he will end up to be in. So you decide to do something and calculate how far one can possibly go? For instance if your friends comes up with an outcome like $(H, T, H, T, H, \cdots)$, then there is no danger since he wont be able to go very far. But what about all other cases?

The answer of the question above might be complicated at first, but it is very straightforward given the identification we had before. E.g. you know that all the experiments that start with a patters line $(H, \cdots)$ should belong to $A_1 \mapsto (1/2, 1]$, and etc. So you start drawing the following *dyadic functions*.
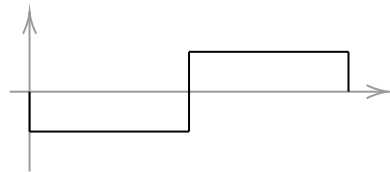

The function above (a random variable!) is very simple, yet tells us a very important fact. This function "reports" the result of the very first step for all the experiments. The experiments in $(0, 1/2]$ will go one step backwards as their first outcome is a tails, and the experiments in $(1/2, 1]$ will go one step forward, as their first outcome is a heads. Cool! Then you continued doing this for the second, and the third step.
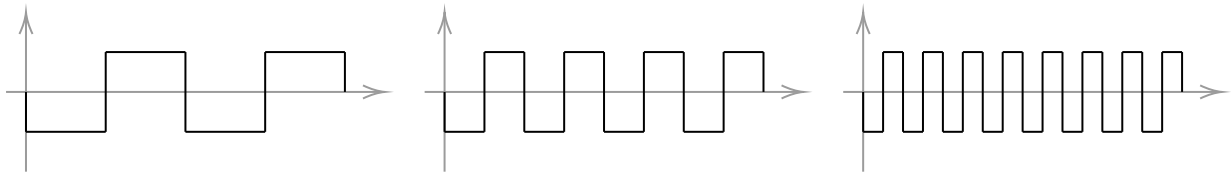

We call these functions as d_1, d_2, and d_3 respectively. The left most function (yet another random variable) is showing the result of the second step for different experiments, i.e. for those experiments in $(0, 1/4]$, that their second toss was a tails will move backwards one step. The other dyadic functions above have a similar interpretation. Very cool! So what if ...

Example 4. You do the calculation above and you suddenly realize if you add all of the dyadic functions above you will be able to look at the result of all of the experiments up to the froth coin toss! (based on the dyadic functions calculate above). Then you will get the following results.
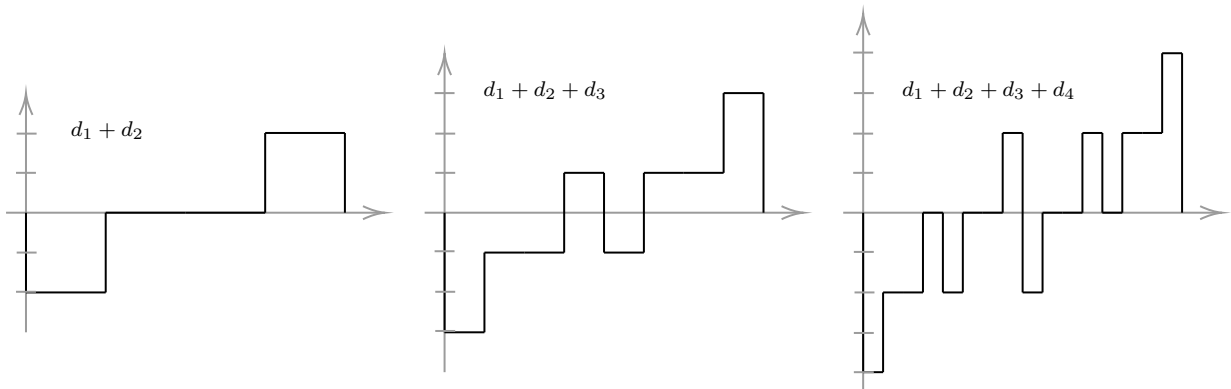
This is very interesting! In the last diagram above, you can see the end point of the people doing experiments for all possible experiments. The most far distance they can travel is four steps (forward or backward) and that happens only in 2/16 of experiments. The second most far that they can get is 2 steps (forward of backward) that happens in 8/16 of the experiments. Lastly, in 6/16 of the experiments the person will remain still after the four steps.

The graph below is a bonus! Observe that after 7 steps, there is a zero chance that one can return to the origin!

$$d_1 + d_2 + d_3 + d_4 + d_5 + d_6 + d_7$$
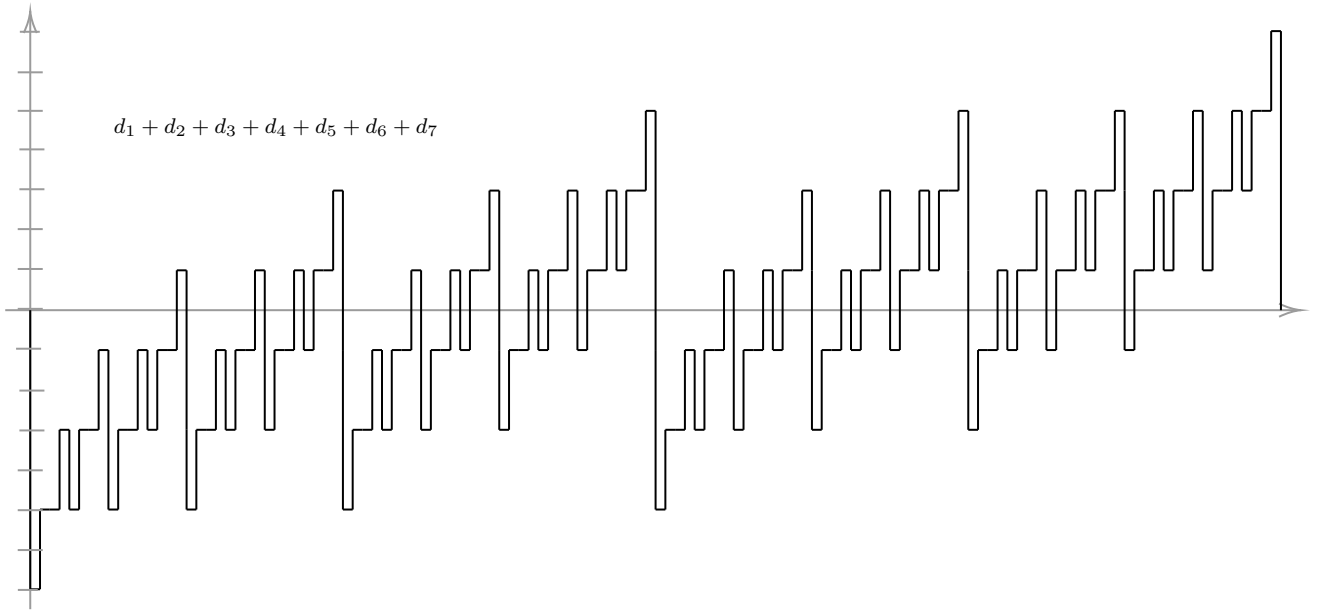
The example demonstrates how we can make more and more complicated spaces, as complex as that it contains the outcome of all experiments. Then random variables will be just functions that are reporting "something" about that particular outcome. For instance, rank k dyadic functions d_k simply reports if the kth coin toss in an experiment is heads or tails. We can also make more complicated random variables that "reports" more sophisticated properties of a particular outcome. For instance, the same as we did above, we can define $S(\omega) = \sum_{i=1}^{n} d_i(\omega)$, which is the sum of values of the tossed coins until the nth toss. With this setting, knowing a probability measure on the space (that determines how likely is for an even $A \in \mathcal{F}$ to happen), we can study the likelihood of different values that can be assume by our random variables.

In the example below we can see the similar random variables as above (position of the walker after n steps), but this time we have also included the stopping time. The walker will stop if they reach -2 or $+3$, which ever happens first. Then we will have

$d_1 + d_2$

$d_1 + d_2 + d_3$

$d_1 + d_2 + d_3 + d_4$

$d_1 + d_2 + d_3 + d_4 + d_5$

$d_1 + d_2 + d_3 + d_4 + d_5 + d_6$