Sums of Random Numbers
and its Hidden Geometric Connections

JONAH SUSSMAN
9 MAY 2020

Preface

This paper is largely inspired by the Numberphile video "Darts in Higher Dimensions"[1]. Please check out the video, it's quite good. Additionally, this paper will try and explain the concepts shown piece-by-piece in an attempt to make the concepts more digestible. And with that...

1 The Problem

Suppose you had some variables $x_1, x_2, x_3, ...$ that all fell within the range $[0, 1]$ uniformly. Now, let $Q(n) = x_1 + x_2 + x_3 + \cdots + x_n$. Let m be the smallest value of n such that $Q(n) > 1$. As all of the variables of x are randomly chosen, ask the following question: *"What is the average, or expected value of m, $E[m]$?"*

2 Unpacking the Question

Let's unpack the question into a more digestible format. Hopefully the first sentence is fairly clear. We have a group of variables, labeled x_1, x_2, x_3 on and on and on. We also know that the variables are randomly and *uniformly* distributed in such a way that any x obeys $0 \leq x \leq 1$.

Next, imagine that all of the values of x line up in a queue, with x_1 at the front, x_2 behind it, x_3 behind x_2 and so on. $Q(n)$ takes the first n xs in the queue and sums up their values. For instance, $Q(1)$ would result in the expression x_1, and $Q(4) = x_1 + x_2 + x_3 + x_4$.

Now, imagine that you are evaluating $Q(n)$. You start with $Q(1)$, then move to $Q(2)$, then $Q(3)$ and so on. Taking note of the value of $Q(n)$, you stop once the sum is greater than one, and record whatever value of n you plugged into the function Q as the value m. For instance, if while evaluating you got $Q(1) = 0.5$, $Q(2) = 1$, and $Q(3) = 1.7$, you would record an m value of 3.

Finally, looking at the *distribution* of possibilities, of all the possible sums Q resulting from the randomly chosen variables x, we can associate an m value with each of them. Thus, we can ask the question, probabilistically, what is the average m value if we were to perform this process?

With this question now unpacked, we can move on to solving the problem.

3 Preliminary Observations

As an aside, let's refresh our memory as to how probability works. Suppose you were at a betting house in Las Vegas (Note: this paper does not condone gambling). A shady individual comes up to you, pulls out a twenty-sided die and offers you a bet. If you can roll a "20" his die, you get $100, otherwise you owe him $10. Assuming he makes good on his deal, should you take the bet? Or a more mathematically rigorous question would be, *what's the expected amount of money you would make by taking the bet?*

A good approach would to find out would be something like the following:

$$E[money] = -\$10 \cdot P(roll \neq 20) + \$100 \cdot P(roll = 20) \tag{1}$$

Substituting in the probabilities, we find out that the expected amount of money, $E[money]$ is

$$E[money] = -\$10 \cdot \frac{19}{20} + \$100 \cdot \frac{1}{20} \tag{2}$$

$$= -\$4.50 \tag{3}$$

Oh no! If you took the shady man's bet, on average you'd be $4.50 in the hole! Bad deal.

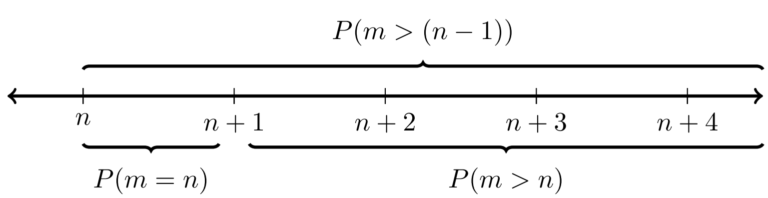
Taking this idea over to our problem, we can apply this line of reasoning to $E[m]$. If we take each value of m and multiply it by the probability of that m occurring, we can arrive at $E[m]$. Therefore, $E[m]$ should look something like this:

$$E[m] = 1 \cdot P(m = 1) + 2 \cdot P(m = 2) + 3 \cdot P(m = 3) + \dots \tag{4}$$

Now, we don't know how to calculate $P(m = somevalue)$. However, note that

$$P(m > (s - 1)) = P(m = s) + P(m > s) \tag{5}$$

As m and s are both integers, this can be explained visually like so:



Using a bit of algebra, we arrive at the following expression for $E[m]$:

$$\begin{aligned} E[m] &= 1 \cdot P(m=1) + 2 \cdot P(m=2) + 3 \cdot P(m=3) + \dots \\ &= 1 \cdot (P(m>0) - P(m>1)) + 2 \cdot (P(m>1) - P(m>2)) + \dots \end{aligned} \tag{6}$$

Notice that each of the subtracting terms gets replaced by one more of itself from the next term. Ex. $-P(m > 1)$ gets added together with $2 \cdot P(m > 1)$. Using this fact, $E[m]$ clearly equals

$$E[m] = P(m>0) + P(m>1) + P(m>2) + \dots \tag{7}$$

4 First Calculations

In the previous section, we deduced that we need to find all of the probabilities $P(m > q)$, where $q \in \mathbb{Z}_0^+$. Let's try and find out what these probabilities are.

For $P(m > 0)$ and $P(m > 1)$ it is trivial to find out what values they are. By the very nature of the problem, m will always be at least 2. This is because when $n = 0$, $Q(0) = 0$ which is less than 1. Additionally, since the process involves finding the first n that is *greater* than 1, and $Q(1) \in [0, 1]$, the only n value that even *could* be greater than 1 is $n = 2$. Therefore, $P(m > 0) = P(m > 1) = 1$.

Now, we move on to $P(m > 2)$. This is the first expression that requires a change in perspective. First, let's try and wrap our minds around what this expression really represents. It's saying, *"What is the probability that m is greater than 2?* In other words, every sum that comes before $n = 3$ must be less than or equal to one. This makes sense, as if the sum were greater than 1 before that point, then m would have been recorded earlier in the process.

Written in a more concrete "mathematical" way

$$0 \leq x_1 + x_2 \leq 1 \tag{8}$$

Now, we can use some geometry and geometric intuition to find out this probability. Graphing the expression yields Fig. (1)

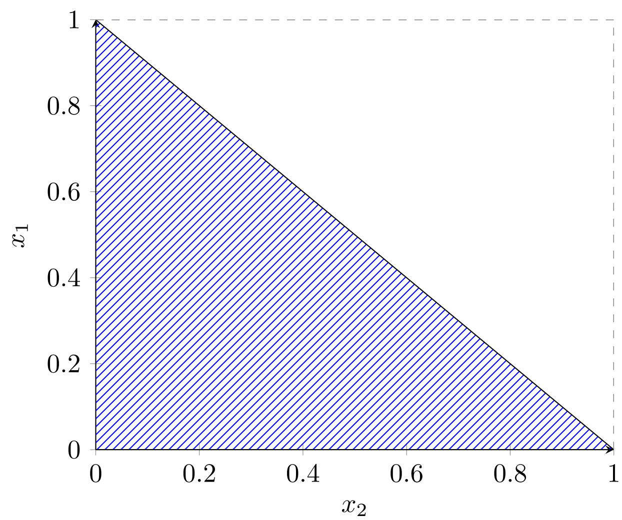


Figure 1: Plot of $0 \leq x_1 + x_2 \leq 1$

Choosing x_1 and x_2 randomly within the range $[0, 1]$, it can be easily seen that the probability that $0 \leq x_1 + x_2 \leq 1$ is simply the area of the whole square over the area of the blue shaded region. One way to "gut check" ourselves is to envision the graph in another light.

Say for instance, you were dropping a ball randomly down a 1x1 square chute. If we wanted to find the probability that the ball would land in the blue triangular region, we could simply take the area of the of that region over the area of the square, $P(blueregion) = \frac{A_{blue}}{A_{square}}$.

Applying this back to our original situation, and using the formulas for the area of a triangle and square, we arrive at

$$P(m > 2) = \frac{\frac{1}{2}bh}{s^2} = \frac{\frac{1}{2} \cdot 1 \cdot 1}{1^2} = \frac{1}{2} \tag{9}$$

5 A Different View of Triangles

What we just did was very subtle. We converted geometry into probability. Additionally, we converted the equality $0 \leq x_1 + x_2 \leq 1$ into a triangular area. Rearranging the equality and keeping in mind that $x_1 \in [0, 1]$, we can come up with this equation:

$$0 \leq x_1 \leq 1 - x_2 \tag{10}$$

Borrowing a concept from vector calculus, we can find the area of this inequality via the following double integral

$$\int_0^1 \int_0^{1-x_1} dx_2 dx_1 \tag{11}$$

First we integrate along x_2 from 0 until we reach the line $1 - x_1$ over the constant function 1. Then, we integrate over x_1 from 0 to 1. Indeed, let's check and make sure that this is correct

$$\begin{aligned}\int_0^1 \int_0^{1-x_1} dx_2 dx_1 &= \int_0^1 (1-x_1) dx_1 \\ &= (x_1 - \frac{1}{2}x_2^2)\Big|_0^1 \\ &= 1 - \frac{1}{2} \\ &= \frac{1}{2}\end{aligned}$$

To really solidify this idea in our heads, lets perform one more specific example.

For $P(m > 3)$, we first create the region $0 \leq x_1 + x_2 + x_3 \leq 0$. Rearranging, we find that $P(m > 3)$ is equal to the region defined By

$$0 \leq x_1 \leq 1 - x_2 - x_3 \tag{12}$$

We can graph this region like so

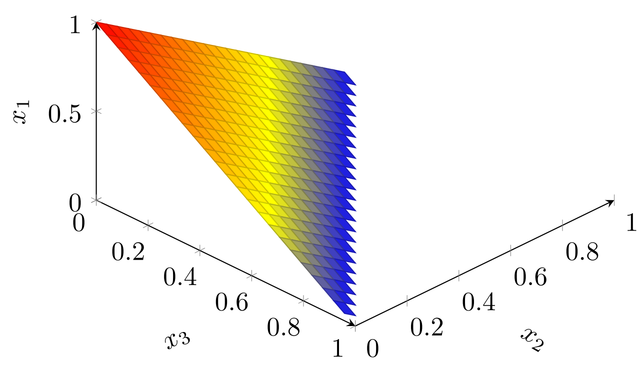


Figure 2: Graph of the pyramid representing the region of Eq.(12)

Using vector calculus, we can find the volume under the pyramid with the given integral

$$\int_0^1 \int_0^{1-x_1} \int_0^{1-x_1-x_2} dx_3 dx_2 dx_1$$
$$\int_0^1 \int_0^{1-x_1} (1-x_1-x_2) dx_2 dx_1$$
$$\frac{1}{2} \int_0^1 (x_1-1)^2 dx_1$$
$$\frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6} \tag{13}$$

Thus $P(m > 3) = \frac{1}{6}$

Since we know that we can relate probability to area or volume, we can now start to generalize this fact.

If we want to find $P(m > q)$, we have to find the area (for higher dimensions, hyper-volume) contained withing the inequality $0 \leq x_1 + x_2 + \cdots + x_q \leq 1$. Rearranging, treating x_1 as an output variable of sorts, and reminding ourselves that $x_1 \in [0, 1]$, we arrive at the inequality $0 \leq x_1 \leq 1 - x_2 - x_3 - \cdots - x_q$. Using vector calculus, we integrate over x_q from 0 until it reaches the hyper plane $1 - x_1 - x_2 - \cdots - x_{q-1}$. Then we integrate over x_{q-1} from 0 until it reaches the lower dimension hyperplane $1 - x_1 - x_2 - \cdots - x_{q-2}$. This continues on until we finally integrate over x_1 from 0 to 1. Much like with the blue shaded triangle, we then take this hyper-volume and divide it by the hyper-volume of the hyper-cube (which will always be 1 as all xs range from $[0, 1]$. Thus we arrive at the final expression

$$P(m > q) = \int_0^1 \int_0^{1-x_1} \cdots \int_0^{1-x_1-x_2-\cdots-x_{q-1}} (1) dx_q dx_{q-1} \ldots dx_1 \tag{14}$$

6 Hyper-Pyramids and Deformation

This expression is fantastic, as it allows us to calculate and value of $P(m > q)$ that we could ever want. However, it is extremely unwieldy. It would be nice to get a closed form solution for the expression.

Let us think about what the expression is actually telling us. In 2 dimensions, the expression results in the area of a triangle with sides equal to 1. In 3 dimensions, it results in a pyramid that lays against the corner in the origin, with sides equal to 1. Let's generalize this concept into higher dimensions with a shape we will call a "hyper-pyramid".

Let $\hat{\boldsymbol{e}}_1, \hat{\boldsymbol{e}}_2, \hat{\boldsymbol{e}}_3, \ldots$ be the standard euclidean basis vectors. For example $\hat{\boldsymbol{e}}_1 =< 1, 0, 0, \cdots >$, $\hat{\boldsymbol{e}}_2 =< 0, 1, 0, \cdots >$, etc... A "hyper-pyramid of dimension n and radius r" exists in the space with the basis vectors $\hat{\boldsymbol{e}}_1, \hat{\boldsymbol{e}}_2, \hat{\boldsymbol{e}}_3, \ldots \hat{\boldsymbol{e}}_n$. A vector $\boldsymbol{v} = x_1\hat{\boldsymbol{e}}_1 + x_2\hat{\boldsymbol{e}}_2 + \cdots + x_n\hat{\boldsymbol{e}}_n$ is considered to be inside the hyper-pyramid if $0 \leq x_1 + x_2 + \cdots + x_n \leq r$.

Let $H(n)$ be the hyper-volume of a hyper-pyramid of dimension n and radius 1. We can use the previous information that we have collected to state that $P(m > q) = H(q)$. We need to find a closed form solution of $H(n)$.

We know that $H(1) = 1$. We also know that $H(2) = \frac{1}{2}$, because $H(2)$ is simply the area of a triangle with a base and height of 1. However, let's reframe how we think of the area of a triangle.

When we evaluate $H(n)$, we can think about it "operating" on the dimensions below it. As when we increase n by one, we get a new dimension orthogonal to every other dimension. Additionally, When taking the cross-section of the hyper-pyramid along this new dimension at 0, the cross section's area is simply equal to $H(n-1)$. For the specific case of the hyper-pyramid, we notice that this value "squishes" the other dimensions linearly down until the cross section evaluated at the new orthogonal dimension at 0 is equal to 0.

In the context of the triangle ($H(2)$), notice that in Fig. (1), when we evaluate $x_2 = 1$, the "cross section" equals 1. Then, as the the value x_2 the remaining dimensions, in this case just x_1, gets "squished" linearly until $x_2 = 1$, at which point $x_1 = 0$. The "squishification" function for $H(2)$ is $(1 - x_2)$.

In the context of the 3D pyramid ($H(3)$), notice that in Fig. (2), there is a new orthogonal axis, x_3. When we evaluate x_3, the cross section is $H(2)$, or the triangle with base and height of 1. Then, as x_3 increases from 0 to 1, *both the base and the height* decrease in proportion with x_3. Thus, for $H(3)$, the "squishification" function is $(1 - x_3)^2$.

In general, the "squishification" function for any $H(n)$ can be found by $(1 - x_n)^{n-1}$. This make sense, as we can visualize the hyper-pyramid's cross sections being contained within an $n-1$ dimensional hypercube that shrinks as the new orthogonal dimension goes from 0 to 1.

As an aside, we could see what happens with other "squishification" functions. For example, if instead we had the function be $(1 - x_n)^{\frac{n-1}{2}}$, the resulting hyper volumes would be that of an upper slice of a sphere. One could do more research into the various solids that arise out of this process.

We can formulate this result in a more mathematical way. The cross sectional area of $H(n)$ is given by $H(n-1)(1 - x_n)^{n-1}$. Integrating with respect to x_n over the interval 0 to 1 will yield us the formula for the hyper volume!

$$\begin{aligned} H(n) &= \int_0^1 H(n-1)(1 - x_n)^{n-1} dx_n \\ &= H(n-1) \int_0^1 (1 - x_n)^{n-1} dx_n \end{aligned}$$

We let $u = 1 - x_n$. Thus $dx_n = -du$, and the bounds range from 1 to 0.

$$\begin{aligned} H(n) &= H(n-1) \cdot -\int_1^0 u^{n-1} du \\ &= H(n-1) \cdot -\frac{1}{n}(u)^n \bigg|_1^0 \\ &= H(n-1) \cdot \frac{1}{n} \end{aligned}$$

Expanding out, we finally arrive attempt

$$\begin{aligned} H(n) &= H(n-2) \cdot \frac{1}{n-1} \cdot \frac{1}{n} \\ &= \frac{1}{1} \cdot \frac{1}{2} \cdot \frac{1}{3} \cdot \dots \cdot \frac{1}{n} \end{aligned}$$

$$H(n) = \frac{1}{n!} \tag{15}$$

7 Conclusion and Further Research

This result for the volume of a hyper pyramid leads us to the most beautiful Conclusion. Using the fact that $P(m > q) = H(q) = \frac{1}{q!}$, we come up with the following solution

$$\begin{aligned} E[m] &= P(m > 0) + P(m > 1) + P(m > 2) + \dots \\ &= 1 + H(1) + H(2) + H(3) + H(4) + \dots \\ &= 1 + 1 + \frac{1}{2} + \frac{1}{3!} + \frac{1}{4!} + \dots \end{aligned}$$

Thus, the expected value of m, after all of this work is the most beautiful solution

$$E[m] = e \tag{16}$$

The original question dealt with probabilities of independently random numbers ranging from 0 to 1. We wanted to know what the expected number of terms we would have to sum in order for it to be greater than 1. We then visualized each of the numbers as independent axes on plots, and found out that the areas and volumes of the regions that obeyed $0 \leq x_1 + x_2 + x_3 + \dots + x_q \leq 1$ gave us the $P(m > q)$.

Thus, we followed its trail to an analogous question: "What is sum of all volumes of hyper-pyramids for all dimensions $n \in \mathbb{Z}^+$ plus 1?" Because the hyper-volume of a hyper pyramid can be found via $\frac{1}{n!}$ where n is the dimension, the resulting value of $E[m]$ directly follows the taylor series expansion of e^x when $x = 1$.

Some future areas of research could be to find out different solids that are generated with different "squishification" functions. For instance, we know that when the function is $(1-x_n)^{n-1}$, we get a hyper pyramid, and when $(1-x_n)^{\frac{n-1}{2}}$ we get a section of a hypersphere, but what other function yield other solids? What if we let the variables range outside of [0, 1]? What if we let the function vary based on the number of dimensions according to a series of some sort? These are all interesting questions that could be answered in a different paper.

References

[1] Haran, Brady and Grant Sanderson. *Darts in Higher Dimensions. YouTube* Numberphile, 17 Nov. 2019, `https://www.youtube.com/watch?v=6_yU9eJ0NxA.`