Probability in high dimensions

Aidan Rocke
aidanrocke@gmail.com

April 21, 2019

Abstract

While running a computer simulation that involved the ratio of two random variables whose denominator was a symmetric random variable centred at zero, I observed that this fact never caused the simulation to crash. This made me curious and I discovered that this interesting numerical observation turned out to be a theorem. While this is not a fundamental discovery, it is one of many experiences that convinced me of the increasingly important role computers will play in the discovery of mathematical theorems in the future.

1 Analysis of a special case

In order to make progress, I decided to start by analysing the special case of $a_i \sim \mathcal{U}(\{-1,1\})$ where:

$$\forall n \in \mathbb{N}, P(a_n = 1) = P(a_n = -1) = \frac{1}{2} \tag{1}$$

$$S_0 = \{(a_n)_{n=1}^N \in -1, 1 : \sum_n a_n = 0\} \tag{2}$$

Knowing that S_0 is non-empty only if we have parity of positive and negative terms, we may deduce that:

$$S_0 \neq \emptyset \iff N \in 2\mathbb{N} \tag{3}$$

For the above reason, I focused my analysis on the following sequence:

$$u_N = P(\sum_{n=1}^{2N} a_n = 0) = \frac{\binom{2N}{N}}{2^{2N}} = \frac{(2N)!}{2^{2N}(N!)^2} \tag{4}$$

1.1 Proof that u_N is decreasing

We can demonstrate that u_N is strictly decreasing by considering the ratio:

$$\frac{u_{N+1}}{u_N} = \frac{\frac{(2N+2)!}{2^{2N+2}((N+1)!)^2}}{\frac{(2N)!}{2^{2N}(N!)^2}} = \frac{(2N+2)(2N+1)}{4(N+1)^2} = \frac{2N+1}{2N+2} < 1 \tag{5}$$

Now, with (5) we have what is necessary to show that:

$$\lim_{n\to\infty} u_N = 0 \tag{6}$$

1.2 Analysis of the limit $\lim_{N \to \infty} u_N$

Using (5) we may also derive a recursive definition of u_N:

$$u_{N+1} = \frac{2N+1}{2N+2} \cdot u_N \tag{7}$$

and given that $u_0 = 1$ we have:

$$u_N = \prod_{n=0}^{N-1} \frac{2n+1}{2n+2} = \frac{1}{3} \cdot \frac{3}{4} \cdot \frac{5}{6} \cdot ... \tag{8}$$

At this point we can make the useful observation:

$$\lim_{N \to \infty} u_N = 0 \implies \lim_{N \to \infty} -\ln u_N = \infty \tag{9}$$

1.3 Proof that $\lim_{N \to \infty} u_N = 0$

By combining (7) and (9) we find that:

$$-\ln u_N = -\ln \prod_{n=0}^{N-1} \frac{2n+1}{2n+2} = \sum_{n=0}^{N-1} \ln \frac{2n+2}{2n+1} = \sum_{n=0}^{N-1} \ln \big(1 + \frac{1}{2n+1}\big) \tag{10}$$

and we note that when $n \in \mathbb{N}$ is large:

$$\ln \big(1 + \frac{1}{n}\big) \approx \frac{1}{n} \tag{11}$$

Now, from (11) it follows that:

$$\sum_{n=1}^{\infty} \frac{1}{2n+1} = \infty \implies \sum_{n=0}^{\infty} \ln \big(1 + \frac{1}{2n+1}\big) = \infty \tag{12}$$

So we may conclude that (6) is indeed true. In some sense, when n is large we may expect to observe the expected value with vanishing probability. This does not contradict the weak law of large numbers however since the weak law only states that the sample average should converge to the expected value whereas in our analysis we are comparing the sample sum to the expected value.

1.4 Discussion

A natural question that follows is whether the above method may be used to handle other cases. Let's consider $a_i \sim \mathcal{U}(\{-1, 0, 1\})$ where:

$$\forall n \in \mathbb{N}, P(a_n = 1) = P(a_n = 0) = P(a_n = -1) = \frac{1}{3} \tag{13}$$

so we may define:

$$u_N = P(\sum_{n=1}^{4N} a_n = 0) = \frac{(4N)!}{3^{4N}} \sum_{k=1}^{N} \frac{1}{(2k)!^2 (4N-4k)!} \tag{14}$$

I actually tried to analyse the combinatorics of this sequence but realised that even if I managed to show that this sequence converged to zero, it wasn't clear how this method would manage to handle the most general setting, the case of all integer dimensions $N \in \mathbb{N}$, and it didn't appear to be very effective in terms of the number of calculations per case. For the more general case, a different approach would be necessary.

2 A more general case

2.1 A random walk on $\mathbb{Z}$

Let's suppose $a_i \sim \mathcal{U}([-N, N])$ where $[-N, N] \subset \mathbb{Z}$. We may then define:

$$S_n = \sum_{i=1}^{n} a_i \tag{15}$$

Due to the i.i.d. assumption we have:

$$\mathbb{E}\big[S_n\big] = n \cdot \mathbb{E}\big[a_i\big] = 0 \tag{16}$$

We may now define:

$$u_n = P(S_n = 0) \tag{17}$$

and ask whether u_n is decreasing. In other words, what is the probability that we observe the expected value as n becomes large?

2.2 Small and Large deviations

It's useful to observe the following nested structure:

$$\forall k \in [0, N], \{|S_n| \leq k\} \subset \{|S_n| \leq k+1\} \tag{18}$$

From (18), we may deduce that:

$$P(|S_n| \leq N) + P(|S_n| > N) = 1 \tag{19}$$

So we are now ready to define the probability of a 'small' deviation:

$$\alpha_n = P(|S_n| \leq N) \tag{20}$$

as well as the probability of 'large' deviations:

$$\beta_n = P(|S_n| > N) \tag{21}$$

Additional motivation for analysing α_n and β_n arises from:

$$P(S_{n+1} = 0 | |S_n| > N) = 0 \tag{22}$$

$$P(S_{n+1} = 0 | |S_n| \leq N) = \frac{1}{2N+1} \tag{23}$$

Furthermore, by the law of total probability we have:

$$\begin{aligned} P(S_{n+1} = 0) &= P(S_{n+1} = 0 | |S_n| \leq N) \cdot P(|S_n| \leq N) + P(S_{n+1} = 0 | |S_n| > N) \cdot P(|S_n| > N) \\ &= P(S_{n+1} = 0 | |S_n| \leq N) \cdot P(|S_n| \leq N) \\ &= \frac{P(|S_n| \leq N)}{2N+1} \end{aligned} \tag{24}$$

2.3 A remark on symmetry

It's useful to note the following alternative definitions of α_n and β_n that emerge due to symmetries intrinsic to the problem:

$$\beta_n = P(|S_n| > N) = 2 \cdot P(S_n > N) = 2 \cdot P(S_n < -N) \tag{25}$$

$$\alpha_n = P(|S_n| \leq N) = 1 - 2 \cdot P(S_n > N) = 1 - 2 \cdot P(S_n < -N) \tag{26}$$

2.4 The case of $n = 1$ and $n = 2$

Given that $S_0 = 0$:

$$P(S_1 = 0) = \frac{P(|S_0| \leq N)}{2N+1} = \frac{1}{2N+1} \tag{27}$$

As for the case of $n = 2$:

$$P(|S_2| \leq N) = 1 \implies P(S_2 = 0) = \frac{1}{2N+1} \tag{28}$$

2.5 The case of $n = 3$

The case of $n = 3$ requires that we calculate:

$$\begin{aligned} P(S_2 > N) &= \sum_{i=1}^{N} P(S_2 > N | S_1 = i) \cdot P(S_1 = i) \\ &= \frac{1}{2N+1} \sum_{i=1}^{N} (\frac{1}{2N+1} + ... + \frac{N}{2N+1}) \\ &= \frac{N \cdot (N-1)}{2 \cdot (2N+1)^2} \end{aligned} \tag{29}$$

and using (29) we may derive $P(S_2 \leq N)$:

$$\begin{aligned} P(S_2 \leq N) &= 1 - 2 \cdot P(S_2 > N) \\ &= 1 - \frac{N \cdot (N-1)}{(2N+1)^2} \\ &= \frac{3N^2 + 5N + 1}{(2N+1)^2} \sim \frac{3}{4} \end{aligned} \tag{30}$$

and so for $n = 3$ we have:

$$\begin{aligned} P(S_3 = 0) &= P(S_3 = 0 || S_2| \leq N) \cdot P(|S_2| \leq N) \\ &= \frac{3N^2 + 5N + 1}{(2N+1)^3} \sim \frac{3}{8N} \end{aligned} \tag{31}$$

2.6 Average drift or why $P(S_n = k) > P(S_n = k+1)$

It's useful to note that we may decompose n into:

$$n = \hat{n} + n_z \tag{32}$$

where $\hat{n}$ represents the total number of positive and negative terms, allowing us to ignore the trivial contribution of zero terms n_z.

For the above reason, it's convenient to decompose S_n into:

$$S_n = S_n^+ + S_n^- \tag{33}$$

where S_n^+ defines the sum of the positive terms and S_n^- defines the sum of the negative terms.

By grouping the terms in the manner of (33) we may observe that when $\hat{n}$ is large the average positive/negative step length is given by:

$$\Delta = \frac{N}{2} \tag{34}$$

so that if τ positive steps and $\hat{n} - \tau$ negative steps are taken:

$$\mathbb{E}[S_n^+] = \tau \cdot \Delta \tag{35}$$

$$\mathbb{E}[S_n^-] = (\hat{n} - \tau) \cdot (-\Delta) \tag{36}$$

$$\mathbb{E}[S_n] = \mathbb{E}[S_n^+] + \mathbb{E}[S_n^-] = \Delta \cdot (2\tau - \hat{n}) \tag{37}$$

and we note that:

$$\mathbb{E}[S_n] \geq 0 \implies \tau \geq \lfloor \frac{\hat{n}}{2} \rfloor \tag{38}$$

Furthermore, due to symmetry:

$$P(|S_n| = k) > P(|S_n| = k+1) \iff P(S_n = k) > P(S_n = k+1) \tag{39}$$

so it suffices to demonstrate $P(S_n = k) > P(S_n = k+1)$.

In order to proceed with our demonstration we choose $\tau \in [\lfloor \frac{\hat{n}}{2} \rfloor + 1, \hat{n}N - 1]$ and find that P has a monotone relationship with the binomial distribution:

$$P(S_n = \lfloor \Delta \cdot (2\tau - \hat{n}) \rfloor) \propto \binom{\hat{n}}{\tau} \frac{1}{2^{\hat{n}}} \tag{40}$$

where $\tau \geq \lfloor \frac{\hat{n}}{2} \rfloor$ implies that:

$$\forall k \geq 0, \frac{P(S_n = k)}{P(S_n = k+1)} \sim \frac{(\tau+1)!(\hat{n} - \tau - 1)!}{\tau!(\hat{n} - \tau)!} = \frac{\tau + 1}{\hat{n} - \tau} > 1 \tag{41}$$

which holds for all $n_z \leq n$.

2.7 Proof that u_n is decreasing

Given (24) we may derive the following ratio:

$$\frac{u_{n+1}}{u_n} = \frac{P(|S_n| \leq N)}{(2N+1) \cdot P(S_n = 0)} \tag{42}$$

So in order to prove that u_n is decreasing we must show that:

$$P(|S_n| \leq N) < (2N+1) \cdot P(S_n = 0) \tag{43}$$

and we note that this follows immediately from (42) since:

$$P(|S_n| \leq N) = 2\sum_{k=1}^{N} P(S_n = k) + P(S_n = 0) < (2N+1) \cdot P(S_n = 0) \tag{44}$$

2.8 Proof that $\lim\limits_{n\to\infty} u_n = \lim\limits_{n\to\infty} \alpha_n = 0$

Now, given (44) we may define:

$$\forall N \in \mathbb{N}, q_n = \frac{P(|S_n| \leq N)}{(2N+1)P(S_n = 0)} < 1 \tag{45}$$

Furthermore, using the concentration phenomenon (40) we may show that when n is large:

$$P(S_{2n} = 0) \sim \frac{1}{2^{2n}}\binom{2n}{n} \sim \frac{1}{2^{2n}} \frac{\sqrt{4\pi n}(\frac{2n}{e})^{2n}}{2\pi n(\frac{n}{e})^{2n}} \sim \frac{1}{\sqrt{\pi n}} \tag{46}$$

From this we may deduce that q_n is decreasing:

$$\prod_{n=1}^{\infty} q_n = \prod_{n=1}^{\infty} \frac{P(S_{n+1} = 0)}{P(S_n = 0)} = \lim_{n\to\infty} \frac{P(S_{n+1} = 0)}{P(S_1 = 0)} = 0 \tag{47}$$

Likewise, given that:

$$\alpha_n = P(|S_n| \leq N) = (2N+1) \cdot P(S_{n+1} = 0) \tag{48}$$

we may conclude that large deviations are exponentially more likely as n becomes large:

$$\lim_{n\to\infty} \alpha_n = \lim_{n\to\infty} (2N+1) \cdot P(S_{n+1} = 0) = 0 \tag{49}$$

$$\lim_{n\to\infty} \beta_n = \lim_{n\to\infty} P(|S_n| > N) = \lim_{n\to\infty} 1 - \alpha_n = 1 \tag{50}$$

A geometric interpretation of the last two limits is that the ratio of the largest hyperplane intersection of the discrete hypercube relative to the total volume of the discrete hypercube shrinks as the dimension of the hypercube tends to infinity.

3 Discussion

One interesting lesson from this experience besides the growing importance of experimental mathematics is that random structures, in this case a random walk, are useful for analysing high-dimensional objects. To what degree this is because our intuition for high dimensional structures isn't much better than random, I don't yet have the answer.