Cool Problem

Limitless • 7 Jan 2026

Cool Problem

During the summer, I came across a specific flavor of problem that I could not solve initially, but with some help, I eventually became comfortable with it. **Fun fact:** this question was also asked in a final-round interview at **Squarepoint**.

What made this problem particularly interesting to me was that it *looked* simple on the surface, but quickly forced you to think carefully about stopping times, conditioning, and how to structure expectations. At first glance, the most natural instinct is to brute force it—and that's exactly where I started.

Problem Statement

Let $X_1, X_2, X_3, \cdots$ be i.i.d. $U[0,1]$ random variables and let

$$N = min\{n : X_1 + X_2 + \cdots + X_n > x\}, \quad x \in [0,1]$$

Find $\mathbb{E}(N)$.

In words, we keep drawing uniform random variables until their cumulative sum exceeds x, and we want to know how many draws this takes *on average*.

First Approach: Brute Force via Tail Probabilities

Since N is a discrete random variable taking positive integer values, a natural starting point is to express its expectation in terms of tail probabilities:

$\mathbb{E}[N] = \sum_{n=1}^{\infty} \mathbb{P}(N \geq n)$.

To make progress, we need to understand what the event $N \geq n$ actually means. Observing the definition of N, we see that

$N \geq n \iff \sum_{i=1}^{n-1} X_i \leq x$.

So the problem reduces to computing probabilities involving partial sums of uniform random variables.

Define the event

$A_n = \{\sum_{i=1}^{n} X_i \leq x\}$.

Our goal is to compute $\mathbb{P}(A_n)$ in a form that is easy to work with.

Recursive Formulation

We can write this probability as a volume of a region in $\mathbb{R}^n$:

$\mathbb{P}(A_n) = \int \cdots \int_{A_n} 1 dx_1 \cdots dx_n$.

Let $f_n(x) = \mathbb{P}(A_n)$. Rather than computing this integral directly for each n, we look for a recursive structure.

Conditioning on the last variable $X_n = x_n$, the event A_n requires

$\sum_{i=1}^{n-1} X_i \leq x - x_n$.

This gives the recursion

$f_n(x) = \int_0^x f_{n-1}(x - x_n) dx_n$.

To anchor the recursion, note that since $X_1 \sim U[0, 1]$,

$f_1(x) = \mathbb{P}(X_1 \leq x) = x$.

Using this,

$f_2(x) = \int_0^x (x - x_2) dx_2 = \frac{x^2}{2}$.

At this point, a clear pattern emerges. By induction,

$f_n(x) = \frac{x^n}{n!}$.

Computing $\mathbb{E}(N)$

We can now return to the expression we started with. Using the fact that $\mathbb{P}(N \geq 1) = 1$,

$\mathbb{E}[N] = 1 + \sum_{n=1}^{\infty} f_n(x)$.

Substituting the closed-form expression for $f_n(x)$,

$\mathbb{E}[N] = 1 + \sum_{n=1}^{\infty} \frac{x^n}{n!}$.

This is nothing but the Taylor expansion of e^x, and hence

$\mathbb{E}(N) = \mathrm{e}^x$.

While this approach works nicely, it does rely on explicitly understanding the geometry of these probability regions. Fortunately, there is a cleaner way.

Second Approach: Conditioning on the First Variable

Instead of conditioning on the *last* variable, we now condition on the *first*. This perspective turns out to be much more powerful.

Using the law of total expectation,

$\mathbb{E}[N] = \mathbb{E}[\mathbb{E}[N \mid X_1]]$.

The intuition is simple: if the very first draw already exceeds x, then the process stops immediately.

Define

$f(t) = \mathbb{E}(N \mid X_1 = t), \quad g(x) = \mathbb{E}(N)$.

Since $X_1 \sim U[0,1]$,

$g(x) = \int_0^1 f(t)dt$.

Case Analysis

There are two cases to consider:

- If $t \geq x$, then $N = 1$.
- If $t < x$, then after drawing t, we need to exceed $x - t$ using fresh uniform variables.

In the second case,

$N = min\{n : X_2 + \cdots + X_n > x - t\} + 1 = g(x - t) + 1$.

Putting everything together,

$g(x) = \int_0^x (1 + g(x - t))dt + \int_x^1 1dt = 1 + \int_0^x g(t)dt$.

Differentiating both sides gives

$g'(x) = g(x)$.

With the boundary condition $g(0) = 1$, we again obtain

$g(x) = \mathrm{e}^x$.

This method avoids multidimensional integrals entirely and already hints at how to extend the result.

Extension to $x \in [1, 2]$

Now suppose $x \in [1, 2]$. The same conditioning idea still applies, but the structure of the recursion changes slightly because the remaining threshold $x - t$ can now lie either in $[0, 1]$ or in $[1, 2]$.

Define

$l(t) = \mathbb{E}(N \mid X_1 = t), \quad h(x) = \mathbb{E}(N)$.

As before,

$h(x) = \int_0^1 l(t)dt$.

We analyze $l(t)$ by splitting into cases depending on the value of $x - t$.

Case Analysis

- **If** $x - t \in (1, 2)$, i.e. $t \in (0, x - 1)$, then after observing $X_1 = t$, the remaining problem is still in the regime $[1, 2]$. Hence,

  $N = h(x - t) + 1$.

- **If** $x - t \in (0, 1)$, i.e. $t \in (x - 1, 1)$, then the remaining problem falls back into the original case. From the previous section,

  $N = g(x - t) + 1 = \mathrm{e}^{x-t} + 1$.

Integral Equation for $h(x)$

Combining both cases, we obtain

$h(x) = \int_0^{x-1}(1 + h(x - t))dt + \int_{x-1}^1(1 + \mathrm{e}^{x-t})dt$.

We now simplify each term.

First,

$\int_0^{x-1} 1dt = x - 1,$

and by a change of variables,

$\int_0^{x-1} h(x-t)dt = \int_1^x h(u)du.$

Next,

$\int_{x-1}^1 1dt = 2 - x,$

and

$\int_{x-1}^1 \mathrm{e}^{x-t}dt = \mathrm{e}^x \int_{x-1}^1 \mathrm{e}^{-t}dt = \mathrm{e} - \mathrm{e}^{x-1}.$

Putting everything together,

$h(x) = 1 + \int_1^x h(t)dt + \mathrm{e} - \mathrm{e}^{x-1}.$

Solving the ODE

Differentiating both sides with respect to x gives

$h'(x) = h(x) - \mathrm{e}^{x-1}.$

This is a first-order linear ODE. Using the boundary condition $h(1) = g(1) = \mathrm{e}$, we solve to obtain

$h(x) = \mathrm{e}^x \left(1 + \frac{1}{\mathrm{e}} - \frac{x}{\mathrm{e}}\right).$

In particular, at $x = 2, h(2) = \mathrm{e}^2 - \mathrm{e}.$

Expected Overshoot Sum

We now switch gears and look at a closely related quantity.

Define

$S_N = X_1 + X_2 + \cdots + X_N,$

the sum at the stopping time. Our goal is to compute $\mathbb{E}(S_N)$.

Let

$k(x) = \mathbb{E}(S_N), \quad q(t) = \mathbb{E}(S_N \mid X_1 = t).$

As before,

$k(x) = \int_0^1 q(t)dt.$

Case Analysis

- If $t > x$, then the process stops immediately and $S_N = t$.
- If $t \leq x$, then $S_N = t + k(x - t)$.

Thus,

$k(x) = \int_0^x (t + k(x - t))dt + \int_x^1 tdt = \frac{1}{2} + \int_0^x k(t)dt.$

Differentiating,

$k'(x) = k(x).$

Using $k(0) = 0.5$, we get

$k(x) = \frac{e^x}{2}.$

Final Insight

Notice something elegant:

$\mathbb{E}(S_N) = \frac{1}{2}e^x = \mathbb{E}(N)\mathbb{E}(X_1).$

This is exactly what Wald's lemma predicts, since N is a stopping time.

Closing Thought

The key lesson here is that conditioning—when done thoughtfully—can turn a high-dimensional integration problem into a one-line differential equation.

The next time you see a chain of uniform random variables, try conditioning early. It will almost always save you from visualizing multidimensional simplices.

Pretty cool, right?