A Simple Multiple Integral Solution to the Broken Stick Problem

Vivek Kaushik

December 14, 2021

Abstract

Fix an integer \( n \geq 2 \). Partition the interval \([0, 1]\) into \( n + 1 \) different intervals \( I_1, \ldots, I_{n+1} \). We solve the Broken Stick problem, which is to find the probability that the lengths of these \( n + 1 \) intervals are valid side lengths of a polygon with \( n + 1 \) sides. We will show that this probability is equal to \( 1 - \frac{n+1}{2^n} \) using multiple integration.

1 Introduction

Partition the interval \([0, 1]\) into \( n + 1 \) different intervals \( I_1, \ldots, I_{n+1} \), where \( n \geq 2 \) is an integer. The Broken Stick problem asks to find the probability these \( n + 1 \) intervals have lengths that are valid side lengths of a polygon with \( n + 1 \) sides. By classical geometry, this is equivalent to finding the probability that for each \( 1 \leq i \leq n + 1 \), the length of \( I_i \) is less than the sum of the lengths of \( I_j \) for all \( j \neq i \).

The answer to the Broken Stick problem is:

\[
1 - \frac{n+1}{2^n}. \tag{1.1}
\]

The particular case \( n = 2 \) is well-known. In this case, the Broken Stick problem asks to find the probability that the lengths of \( I_1, I_2, \) and \( I_3 \) are valid side lengths of a triangle. The answer is \( 1/4 \), and multiple solutions from different areas of mathematics have appeared in the literature (see all of [1–5]). D’Andrea and Gomez [6] published a solution to the general case using combinatorial topology.

In this paper, we will give an alternative solution to the Broken Stick problem using multiple integration. We let \( X_1, \ldots, X_n \) be \( n \) independent and identically distributed uniform random variables on \([0, 1]\). Regarding \([0, 1]\) as a stick, these random variables represent the \( n \) breaking points of the stick. We make the ordering assumption that \( X_1 \leq \ldots \leq X_n \) and show that the solution to the Broken Stick problem with this ordering assumption is

\[
\frac{2^n - (n + 1)}{2^n n!}. \tag{1.2}
\]

To do this, we formulate the necessary conditions on these random variables for them to satisfy the Broken Stick problem. Then, for each \( 1 \leq k < n \), we further impose the constraint that exactly \( k \) of these random variables must be less than \( 1/2 \), and we solve the Broken Stick problem with this additional constraint by explicitly evaluating a multiple integral. Using a well-known combinatorial formula, we sum the results over all possible aforementioned \( k \) and obtain the solution to the Broken Stick problem with only our original ordering assumption, namely (1.2). Finally, since there are \( n! \) total orderings of \( X_1, \ldots, X_n \), multiplying the result in (1.2) by \( n! \) recovers the solution to the original Broken Stick problem without any ordering assumptions, namely (1.1).

2 Preliminaries

In this section, we prove two formulas that are crucial for solving the Broken Stick problem.

The first formula is an integration identity of a certain monomial function over an \( n \)-dimensional simplex.

**Theorem 2.0.1 (Integration Identity of Monomial Function Over a Simplex).** Suppose \( p \in \mathbb{N} \cup \{0\} \) and suppose \( a, b \in \mathbb{R} \) with \( a \leq b \). For each \( n \in \mathbb{N} \), we have

\[
\int_{a}^{b} \int_{0}^{x_1} \ldots \int_{0}^{x_{n-1}} x_n^p \, dx_n \ldots \, dx_2 \, dx_1 = \frac{b^{p+n} - a^{p+n}}{(p+1) \ldots (p+n)}. \tag{2.1}
\]
Proof. We proceed by induction on \( n \).

In the case \( n = 1 \), the left hand side of (2.1) becomes the integral

\[
\int_a^b x_1^p \, dx_1. \tag{2.2}
\]

By the elementary power rule of integration, the integral in (2.2) is equal to

\[
\frac{b^{p+1} - a^{p+1}}{p+1},
\]

which clearly agrees with the right hand side of (2.1) upon substituting \( n = 1 \).

Now, suppose the claim holds for the case \( n = k \) for some \( k > 1 \). We show that it must also hold for the case \( n = k + 1 \).

In the case \( n = k + 1 \), the left hand side of (2.1) is equal to

\[
\int_a^b \left( \int_0^{x_1} \ldots \int_0^{x_k} x_{k+1}^p \, dx_{k+1} \ldots dx_2 \right) \, dx_1 = \int_a^b x_1^{p+k} \frac{dx_1}{(p + 1) \ldots (p + k)} = \frac{b^{p+k+1} - a^{p+k+1}}{(p + 1) \ldots (p + k + 1)}, \tag{2.3}
\]

where we obtained the right hand side of (2.3) by using the induction hypothesis to evaluate the \( k \)-dimensional integral within the parentheses on the left hand side of (2.3).

The next formula is a well-known combinatorial formula.

**Theorem 2.0.2** (Combinatorial Identity). For all \( n \in \mathbb{N} \), we have

\[
\sum_{p=0}^n \binom{n}{p} = 2^n, \tag{2.4}
\]

where

\[
\binom{n}{p} = \frac{n!}{(n-p)! \, p!} \tag{2.5}
\]

is the binomial coefficient.

**Proof.** The left hand side of (2.4) is precisely the expansion of \((1 + 1)^n\) using the Binomial Theorem. On the other hand, we have \((1 + 1)^n = 2^n\), which is the right hand side of (2.4).

**3 Multiple Integration Solution to the Broken Stick Problem**

We now are ready to solve the Broken Stick problem. As stated in the introduction, we let \( n \geq 2 \) be an integer and let \( X_1, \ldots, X_n \) be \( n \) independent and identically distributed uniform random variables on \([0, 1]\). Regarding \([0, 1]\) as a stick, these random variables represent the breaking points of \([0, 1]\) so that we obtain \( n + 1 \) different intervals \( I_1, \ldots, I_{n+1} \) that partition \([0, 1]\). The Broken Stick problem asks us to find the probability that the length of each \( I_i \) is less than the sum of the lengths of \( I_j \) for all \( j \neq i \).

We first explicitly rewrite this probability in terms of conditions on \( X_1, \ldots, X_n \) assuming an ordering of these random variables.

**Theorem 3.0.1** (Broken Stick Problem Assuming Ordering). Assume that \( X_1 \leq \ldots \leq X_n \). Then, the Broken Stick problem is equivalent to computing the probability that these random variables satisfy the following conditions:

\[
X_1 < \frac{1}{2}, \quad X_1 \leq X_2 < \frac{1}{2} + X_1, \quad \ldots, \quad X_{n-1} \leq X_n < \frac{1}{2} + X_{n-1}, \quad X_n > \frac{1}{2}. \tag{3.1}
\]

**Proof.** Since \( X_1 \leq \ldots \leq X_n \), the interval \([0, 1]\) is partitioned into \( I_1, \ldots, I_{n+1} \), where

\[
I_1 = [0, X_1], \\
I_i = [X_{i-1}, X_i], \quad 1 < i \leq n, \\
I_{n+1} = [X_n, 1].
\]
We let $\ell(I)$ denote the length of the interval $I$.

The condition that the length of $I_1$ is less than the sum of the lengths of all $I_j$ for $j \neq 1$ is equivalent to the condition

$$X_1 = \ell(I_1) < \sum_{j \neq 1}^{n+1} \ell(I_j)$$

$$= \ell(I_{n+1}) + \sum_{j \neq 1}^{n} \ell(I_j)$$

$$= 1 - X_n + \sum_{j=2}^{n} (X_j - X_{j-1})$$

$$= 1 - X_1.$$

Rearranging yields $X_1 < 1/2$, which is precisely the first condition listed in (3.1).

Let $1 < i \leq n$. The condition that the length of $I_i$ is less than the sum of the lengths of all $I_j$ for $j \neq i$ is equivalent to the condition

$$0 \leq X_i - X_{i-1} = \ell(I_i) < \sum_{j \neq i}^{n+1} \ell(I_j)$$

$$= \ell(I_1) + \ell(I_{n+1}) + \sum_{j \neq 1,n+1,i}^{n} \ell(I_j)$$

$$= X_1 + 1 - X_n + \sum_{j \neq 1,n+1,i}^{n} (X_j - X_{j-1})$$

$$= 1 - X_i + X_{i-1}.$$

Rearranging yields $X_i < 1/2 + X_{i-1}$. On the other hand, by the ordering assumption, we have $X_{i-1} \leq X_i$. These two inequalities are the same as the $i$-th condition listed in (3.1).

Lastly, the condition that the length of $I_{n+1}$ is less than the sum of the lengths of all $I_j$ for $j \neq n + 1$ is equivalent to the condition

$$1 - X_n = \ell(I_{n+1}) < \sum_{j \neq n+1}^{n+1} \ell(I_j)$$

$$= \sum_{j \neq 1}^{n} \ell(I_j)$$

$$= \sum_{j=2}^{n} (X_j - X_{j-1})$$

$$= X_n.$$

Rearranging yields $X_n > 1/2$, which is precisely the last condition listed in (3.1).

In addition to the ordering assumption we made on the random variables in Theorem 3.0.1, we further impose a condition that a certain number of these random variables must be less than or equal to 1/2. This will allow us to solve the Broken Stick problem by explicitly evaluating a multiple integral.

**Theorem 3.0.2** (Broken Stick Solution With Ordering and Extra Bound Assumptions). As before, assume $X_1 \leq \ldots \leq X_n$. Fix $1 \leq k < n$, and let $p_{n,k}$ denote the probability that $X_1, \ldots, X_n$ satisfy all the conditions listed in (3.1) and the additional conditions $X_1, \ldots, X_k < 1/2$, $X_{k+1} \geq 1/2$. Then,

$$p_{n,k} = \frac{n!}{k! (n-k)!} - 1 - \frac{1}{2^n \cdot n!}.$$  \hfill (3.2)

**Proof.** To prove this, we set up and evaluate a multiple integral corresponding to $p_{n,k}$.

The integrand for this multiple integral will be the joint density function of $X_1, \ldots, X_n$, which is 1.
We wish to set up the multiple integral so that we integrate in the order of \( x_n, \ldots, x_1 \). We determine the integral bounds corresponding to \( x_n, \ldots, x_1 \) in that order. We obtain each integral bound for \( x_i \) by obtaining a bound for the corresponding random variable \( X_i \).

For each \( k+1 < i \leq n \), we have \( X_i \geq X_{i-1} \) and \( X_i \geq 1/2 \). On the other hand, we have \( X_i < 1/2 + X_{i-1} \) and \( X_i \leq 1 \). Therefore, we have the bound

\[
X_{i-1} = \max \left( X_{i-1}, \frac{1}{2} \right) \leq X_i < \min \left( \frac{1}{2} + X_{i-1}, 1 \right) = 1.
\]

Furthermore, we have \( X_{k+1} \geq 1/2 \) and \( X_{k+1} \geq X_k \). On the other hand, we have \( X_{k+1} \leq \frac{1}{2} + X_k \) and \( X_{k+1} \leq 1 \). Therefore, we have the bound

\[
\frac{1}{2} = \max \left( X_k, \frac{1}{2} \right) \leq X_{k+1} < \min \left( \frac{1}{2} + X_k, 1 \right) = \frac{1}{2} + X_k.
\]

Moreover, for \( 1 < i \leq k \), we have \( X_i \geq X_{i-1} \). On the other hand, we have \( X_i \leq 1/2 + X_{i-1} \) and \( X_i < 1/2 \). Therefore, we have the bound

\[
X_{i-1} = \max(X_{i-1}, 0) \leq X_i \leq \min \left( \frac{1}{2} + X_{i-1}, \frac{1}{2} \right) = \frac{1}{2}.
\]

Putting everything together, we see that

\[
p_{n,k} = \int_0^1 \int_{x_1}^{u_1} \cdots \int_{x_{k-1}}^{u_{k-1}} \int_{x_{k+1}}^{u_{k+1}} \cdots \int_{x_{n-1}}^{u_{n-1}} \left( \int_{x_k}^{u_k} \left( \frac{1}{2} \right)^k du_n \cdots du_{k+2} \right) du_k \cdots du_2 du_1.
\]

We now evaluate the multiple integral on the right hand side of (3.3). First, we make the affine change of variables \( x_i = \frac{1-u_i}{2} \) for \( 1 \leq i \leq k \) and \( x_i = 1-u_i \) for \( k < i \leq n \), which has Jacobian Determinant

\[
\left| \frac{\partial (x_1, \ldots, x_n)}{\partial (u_1, \ldots, u_n)} \right| = \left( \frac{1}{2} \right)^k.
\]

By the Change of Variables formula, we obtain

\[
p_{n,k} = \int_0^1 \int_0^{u_1} \cdots \int_0^{u_{k-1}} \int_{x_k}^{u_k} \left( \frac{1}{2} \right)^k du_n \cdots du_{k+2} \frac{u_{n-k}}{(n-k)!} \frac{1}{(n-k)!} du_k \cdots du_2 du_1.
\]

\[
= \frac{1}{2^n} \frac{n}{(n-k)!} \frac{1}{(n-k)!} \cdots \frac{1}{(n-k)!} \int_0^1 \int_0^{u_1} \cdots \int_0^{u_{k-1}} \left( \frac{1}{2} \right)^k du_n \cdots du_{k+2} \frac{1}{n!} du_k \cdots du_2 du_1
\]

\[
= \frac{1}{2^n} \frac{n}{(n-k)!} \frac{1}{(n-k)!} \cdots \frac{1}{(n-k)!} \int_0^1 \int_0^{u_1} \cdots \int_0^{u_{k-1}} \left( \frac{1}{2} \right)^k du_n \cdots du_{k+2} \frac{1}{n!} du_k \cdots du_2 du_1.
\]

where we obtained (3.5) by evaluating the \( n-k \) dimensional integral within the parentheses in (3.4) using our integration identity from Theorem 2.0.1 and obtained (3.6) by evaluating the integral in (3.5) using Theorem 2.0.1 once again.

We now can solve the Broken Stick problem assuming the ordering \( X_1 \leq \ldots \leq X_n \) with no extra conditions imposed. By the law of total probability, the probability that \( X_1, \ldots, X_n \) satisfy the
conditions listed in (3.1) is equal to

\[ \sum_{k=1}^{n-1} p_{n,k} = \sum_{k=1}^{n-1} \left( \frac{n}{k} - 1 \right) \]

\[ = \frac{1}{2^n n!} \sum_{k=1}^{n-1} \binom{n}{k} - \frac{1}{2^n n!} \sum_{k=1}^{n-1} 1 \]

\[ = \frac{1}{2^n n!} \left( 2^n - \binom{n}{0} - \binom{n}{n} \right) - \frac{1}{2^n n!} (n-1) \]

(3.7)

\[ = \frac{1}{2^n n!} (2^n - 2) - \frac{1}{2^n n!} (n-1) \]

\[ = \frac{2^n - (n+1)}{2^n n!} \]

where we simplified the summation term in (3.7) using our combinatorial identity from Theorem 2.0.2.

Finally, since there are \( n! \) different orderings of the random variables \( X_1, \ldots, X_n \), the solution to the original Broken Stick problem without any assumptions on ordering is

\[ n! \left( \frac{2^{n} - (n+1)}{2^n n!} \right) = 1 - \frac{n+1}{2^n}. \]

References

[1] Gerald Goodman. The problem of the broken stick reconsidered. The Mathematical Intelligencer, 30(3):43–49, Jun 2008.
[2] Émile Lemoine. Sur une question de probabilités. Bulletin de la Société Mathématique de France, 1:39–40, 1872-1873.
[3] Émile Lemoine. Quelques questions de probabilités résolues géométriquement. Bulletin de la Société Mathématique de France, 11:13–25, 1883.
[4] Eugene Seneta and François Jongmans. The problem of the broken rod and ernesto cesaro’s early work in probability. Mathematical Scientist, 30(2):67–76, 2005.
[5] Henri Poincaré. Calcul des probabilités. Gauthier-Villiers, 1912.
[6] Carlos D’Andrea and Emiliano Gómez. The broken spaghetti noodle. The American Mathematical Monthly, 113(6):555–557, 2006.