STATS 217: Introduction to Stochastic Processes I

Lecture 3

The Ballot Problem

- Consider an election with two candidates A and B.
- Suppose that a votes have been cast for A and b votes have been cast for b where $a > b$.
- After the votes have been cast, they are counted in a uniformly random order.
- Since $a > b$, after all the votes are counted, A emerges as the winner.
- What is the probability that A leads B throughout the count?

The Ballot Problem

- For $0 \le i \le a + b$, let S_i denote the lead of A after i votes have been counted.
- Hence, $S_0 = 0$ and $S_{a+b} = a - b$.

The Ballot Problem

- For $0 \leq i \leq a + b$, let S_i denote the lead of A after i votes have been counted.
- Hence, $S_0 = 0$ and $S_{a+b} = a - b$.
- Since the votes are counted in a uniformly random order, the sequence $S_0, S_1, \ldots, S_{a+b}$ is a uniformly random path from 0 to $a - b$.

The Ballot Problem

- For $0 \leq i \leq a + b$, let S_i denote the lead of A after i votes have been counted.
- Hence, $S_0 = 0$ and $S_{a+b} = a - b$.
- Since the votes are counted in a uniformly random order, the sequence $S_0, S_1, \ldots, S_{a+b}$ is a uniformly random path from 0 to $a - b$.
- Therefore,

$$\mathbb{P}[A \text{ leads throughout}] = \frac{N_{a+b}^{\neq 0}(0, a-b)}{N_{a+b}(0, a-b)}.$$

The Ballot Problem

- For $0 \leq i \leq a+b$, let S_i denote the lead of A after i votes have been counted.
- Hence, $S_0 = 0$ and $S_{a+b} = a - b$.
- Since the votes are counted in a uniformly random order, the sequence $S_0, S_1, \ldots, S_{a+b}$ is a uniformly random path from 0 to $a - b$.
- Therefore,

$$\mathbb{P}[A \text{ leads throughout}] = \frac{N^{\neq 0}_{a+b}(0, a-b)}{N_{a+b}(0, a-b)}.$$

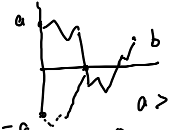
↳ we have a formula for this

$$N_n(a, b) = \binom{n}{\frac{n+b-a}{2}}$$

- So, it only remains to compute $N^{\neq 0}_{a+b}(0, a-b)$.

The Ballot Problem

We need to compute $N^{\neq 0}_{a+b}(0, a-b)$.



$a > 0,\ b > 0$

$N^{0}_{n}(a, b) = N_{n}(-a, b).$

$$N^{\neq 0}_{a+b}(0, \underbrace{a-b}_{>0}) = N^{\neq 0}_{a+b-1}(1, a-b)$$

$$= N_{a+b-1}(1, a-b) - N^{0}_{a+b-1}(1, a-b)$$

$$= N_{a+b-1}(1, a-b) - N_{a+b-1}(-1, a-b)$$

The Ballot Problem

We need to compute $N^{\neq 0}_{a+b}(0, a-b)$.

$$\begin{aligned} N^{\neq 0}_{a+b}(0, a-b) &= N^{\neq 0}_{a+b-1}(1, a-b) \\ &= N_{a+b-1}(1, a-b) - N^{0}_{a+b-1}(1, a-b) \end{aligned}$$

The Ballot Problem

We need to compute $N^{\neq 0}_{a+b}(0, a-b)$.

$$\begin{aligned} N^{\neq 0}_{a+b}(0, a-b) &= N^{\neq 0}_{a+b-1}(1, a-b) \\ &= N_{a+b-1}(1, a-b) - N^{0}_{a+b-1}(1, a-b) \\ &= N_{a+b-1}(1, a-b) - N_{a+b-1}(-1, a-b) \end{aligned}$$

The Ballot Problem

We need to compute $N^{\neq 0}_{a+b}(0, a-b)$.

$$
\begin{aligned}
N^{\neq 0}_{a+b}(0, a-b) &= N^{\neq 0}_{a+b-1}(1, a-b) \\
&= N_{a+b-1}(1, a-b) - N^{0}_{a+b-1}(1, a-b) \\
&= N_{a+b-1}(1, a-b) - N_{a+b-1}(-1, a-b) \\
&= \binom{a+b-1}{a-1} - \binom{a+b-1}{a} \\
&= \frac{a-b}{a+b} \cdot N_{a+b}(0, a-b).
\end{aligned}
$$

The Ballot Problem

We need to compute $N^{\neq 0}_{a+b}(0, a-b)$.

$$\begin{aligned}
N^{\neq 0}_{a+b}(0, a-b) &= N^{\neq 0}_{a+b-1}(1, a-b) \\
&= N_{a+b-1}(1, a-b) - N^{0}_{a+b-1}(1, a-b) \\
&= N_{a+b-1}(1, a-b) - N_{a+b-1}(-1, a-b) \\
&= \binom{a+b-1}{a-1} - \binom{a+b-1}{a} \\
&= \frac{a-b}{a+b} \cdot N_{a+b}(0, a-b).
\end{aligned}$$

Hence,

$$\mathbb{P}[A \text{ leads throughout}] = \frac{a-b}{a+b}.$$

The Ballot Problem

- One way to reinterpret the conclusion of the Ballot problem is that for any $a > b \geq 0$ and for a simple symmetric random walk starting from $S_0 = 0$,

$$\mathbb{P}[S_i > 0 \quad \forall i = 1, \ldots, a+b-1 \mid S_{a+b} = a-b] = \frac{a-b}{a+b}.$$

$$\mathbb{P}[S_1 > 0, \ldots, S_{a+b-1} > 0, S_{a+b} = a-b]$$

$$= \frac{a-b}{a+b} \cdot \mathbb{P}[S_{a+b} = a-b].$$

The Ballot Problem

- One way to reinterpret the conclusion of the Ballot problem is that for any $a > b \geq 0$ and for a simple symmetric random walk starting from $S_0 = 0$,

$$\mathbb{P}[S_i > 0 \quad \forall i = 1, \dots, a+b-1 \mid S_{a+b} = a-b] = \frac{a-b}{a+b}.$$

- Rewritten in more convenient notation, for any integers $k, n > 0$,

$$\mathbb{P}[S_1 > 0, \dots, S_{n-1} > 0, S_n = k] = \frac{k}{n} \cdot \mathbb{P}[S_n = k]. \quad (*)$$

- On the homework, you will explore variants of this.

Path counting and applications

So far, we used path counting to answer the following questions:

- What is the distribution of the first time that a symmetric simple random walk returns to its starting point?
- What is the probability that in a uniformly random counting of the votes, the candidate with more votes stays ahead throughout?

Path counting and applications

So far, we used path counting to answer the following questions:

- What is the distribution of the first time that a symmetric simple random walk returns to its starting point?
- What is the probability that in a uniformly random counting of the votes, the candidate with more votes stays ahead throughout?

Today we will study the following questions:

- What is the distribution of the first time that a symmetric simple random walk hits 1?

τ_1 = first time to hit 1

$\mathbb{P}(\tau_1 < \infty) = 1$ pmf(τ_1)

$\mathbb{E}[\tau_1] = \infty$ prob$(\tau_1 = k)$

Path counting and applications

So far, we used path counting to answer the following questions:

- What is the distribution of the first time that a symmetric simple random walk returns to its starting point?
- What is the probability that in a uniformly random counting of the votes, the candidate with more votes stays ahead throughout?

Today we will study the following questions:

- What is the distribution of the first time that a symmetric simple random walk hits 1?
- If the random walk is run for $2n$ steps, what is the distribution of the *last* time that the random walk visits its starting point?

Path counting and applications

So far, we used path counting to answer the following questions:

- What is the distribution of the first time that a symmetric simple random walk returns to its starting point?
- What is the probability that in a uniformly random counting of the votes, the candidate with more votes stays ahead throughout?

Today we will study the following questions:

- What is the distribution of the first time that a symmetric simple random walk hits 1?
- If the random walk is run for $2n$ steps, what is the distribution of the *last* time that the random walk visits its starting point?
- What is the distribution of the fraction of the time that the random walk is positive?...
- and more!

Hitting time

Let $(S_n)_{n \geq 0}$ be a simple, symmetric random walk starting from 0. For $b > 0$, let $\tau_b = \inf\{n \geq 1 : S_n = b\}$.

Hitting time

Let $(S_n)_{n\geq 0}$ be a simple, symmetric random walk starting from 0. For $b > 0$, let $\tau_b = \inf\{n \geq 1 : S_n = b\}$.

- We already saw that $\mathbb{P}[\tau_b < \infty] = 1$ and $\mathbb{E}[\tau_b] = \infty$.
- What is the pmf of τ_b?

Hitting time

Let $(S_n)_{n\geq 0}$ be a simple, symmetric random walk starting from 0. For $b > 0$, let $\tau_b = \inf\{n \geq 1 : S_n = b\}$.

- We already saw that $\mathbb{P}[\tau_b < \infty] = 1$ and $\mathbb{E}[\tau_b] = \infty$.
- What is the pmf of τ_b?
- We will make use of the trick of looking at the **time reversal** of the walk.

Hitting time

Let $(S_n)_{n\geq 0}$ be a simple, symmetric random walk starting from 0. For $b > 0$, let $\tau_b = \inf\{n \geq 1 : S_n = b\}$.

- We already saw that $\mathbb{P}[\tau_b < \infty] = 1$ and $\mathbb{E}[\tau_b] = \infty$.
- What is the pmf of τ_b?
- We will make use of the trick of looking at the **time reversal** of the walk.
- Recall that $S_j = \sum_{i=1}^{j} X_j$ where $X_1, X_2, \dots$ are i.i.d. Rademacher random variables.

$S_0 = 0$
$S_1 = X_1$
$S_2 = X_1 + X_2$ .....

each X_i is a coin flip
$+1$ w.p. $\frac{1}{2}$
-1 w.p. $\frac{1}{2}$

$$(X_1, \dots, X_n) \sim (X_n, \dots, X_1)$$
$$S_1 = X_1 \sim X_n = S_n - S_{n-1}$$
$$S_2 = X_1 + X_2 \sim X_n + X_{n-1} = S_n - S_{n-2}$$

Hitting time

Let $(S_n)_{n\geq 0}$ be a simple, symmetric random walk starting from 0. For $b > 0$, let $\tau_b = \inf\{n \geq 1 : S_n = b\}$.

- We already saw that $\mathbb{P}[\tau_b < \infty] = 1$ and $\mathbb{E}[\tau_b] = \infty$.
- What is the pmf of τ_b?
- We will make use of the trick of looking at the **time reversal** of the walk.
- Recall that $S_j = \sum_{i=1}^{j} X_j$ where $X_1, X_2, \ldots$ are i.i.d. Rademacher random variables.
- Since $(X_1, \ldots, X_n)$ has the same distribution as $(X_n, \ldots, X_1)$, it follows that

Hitting time

Let $(S_n)_{n\geq 0}$ be a simple, symmetric random walk starting from 0. For $b > 0$, let $\tau_b = \inf\{n \geq 1 : S_n = b\}$.

- We already saw that $\mathbb{P}[\tau_b < \infty] = 1$ and $\mathbb{E}[\tau_b] = \infty$.
- What is the pmf of τ_b?
- We will make use of the trick of looking at the **time reversal** of the walk.
- Recall that $S_j = \sum_{i=1}^{j} X_j$ where $X_1, X_2, \dots$ are i.i.d. Rademacher random variables.
- Since $(X_1, \dots, X_n)$ has the same distribution as $(X_n, \dots, X_1)$, it follows that

$$\begin{aligned} S_j &\sim X_n + X_{n-1} + \cdots + X_{n-j} \\ &\sim S_n - S_{n-j}, \end{aligned}$$

Hitting time

Let $(S_n)_{n\geq 0}$ be a simple, symmetric random walk starting from 0. For $b > 0$, let $\tau_b = \inf\{n \geq 1 : S_n = b\}$.

- We already saw that $\mathbb{P}[\tau_b < \infty] = 1$ and $\mathbb{E}[\tau_b] = \infty$.
- What is the pmf of τ_b?
- We will make use of the trick of looking at the **time reversal** of the walk.
- Recall that $S_j = \sum_{i=1}^{j} X_j$ where $X_1, X_2, \ldots$ are i.i.d. Rademacher random variables.
- Since $(X_1, \ldots, X_n)$ has the same distribution as $(X_n, \ldots, X_1)$, it follows that

$$\begin{aligned} S_j &\sim X_n + X_{n-1} + \cdots + X_{n-j} \\ &\sim S_n - S_{n-j}, \end{aligned}$$

so that

$$(S_1, S_2, \ldots, S_n) \sim (S_n - S_{n-1}, S_n - S_{n-2}, \ldots, S_n).$$

Hitting time

- We want to compute $\mathbb{P}[\tau_b = k]$.

Hitting time

Recall that
$S_0 = 0$
$b > 0$

- We want to compute $\mathbb{P}[\tau_b = k]$.
- We have

$$\mathbb{P}[\tau_b = k] = \mathbb{P}[S_1 < b, \ldots, S_{k-1} < b, S_k = b] \quad .$$

Hitting time

- We want to compute $\mathbb{P}[\tau_b = k]$.
- We have

$$\begin{aligned}\mathbb{P}[\tau_b = k] &= \mathbb{P}[S_1 < b, \ldots, S_{k-1} < b, S_k = b] \\ &= \mathbb{P}[\underbrace{S_1 < S_k}_{S_k - S_1 > 0}, \ldots, S_{k-1} < S_k, S_k = b]\end{aligned}$$

Hitting time

- We want to compute $\mathbb{P}[\tau_b = k]$.
- We have

$$\begin{aligned}\mathbb{P}[\tau_b = k] &= \mathbb{P}[S_1 < b, \ldots, S_{k-1} < b, S_k = b] \\ &= \mathbb{P}[S_1 < S_k, \ldots, S_{k-1} < S_k, S_k = b] \\ &= \mathbb{P}[\underbrace{S_k - S_{k-1}}_{S_1} > 0, \ldots, \underbrace{S_k - S_1}_{S_{k-1}} > 0, S_k = b]\end{aligned}$$

$$(S_1 \ldots S_k) \sim (S_k - S_{k-1}, \ldots, S_k - S_1, S_k)$$

Hitting time

$(S_1 \cdots S_k)$

- We want to compute $\mathbb{P}[\tau_b = k]$. $\sim (S_k - S_{k-1}, \dots, S_k - S_1, S_k)$
- We have

$$
\begin{aligned}
\mathbb{P}[\tau_b = k] &= \mathbb{P}[S_1 < b, \dots, S_{k-1} < b, S_k = b] \\
&= \mathbb{P}[S_1 < S_k, \dots, S_{k-1} < S_k, S_k = b] \\
&= \mathbb{P}[S_k - S_{k-1} > 0, \dots, S_k - S_1 > 0, S_k = b] \\
&= \mathbb{P}[S_1 > 0, \dots, S_{k-1} > 0, S_k = b]
\end{aligned}
$$

$S_1 = X_1$ / $S_k - S_{k-1} = X_k$

↓ ballot problem

Hitting time

- We want to compute $\mathbb{P}[\tau_b = k]$.
- We have

$$\begin{aligned}\mathbb{P}[\tau_b = k] &= \mathbb{P}[S_1 < b, \ldots, S_{k-1} < b, S_k = b] \\ &= \mathbb{P}[S_1 < S_k, \ldots, S_{k-1} < S_k, S_k = b] \\ &= \mathbb{P}[S_k - S_{k-1} > 0, \ldots, S_k - S_1 > 0, S_k = b] \\ &= \mathbb{P}[S_1 > 0, \ldots, S_{k-1} > 0, S_k = b] \\ &= \frac{b}{k} \cdot \mathbb{P}[S_k = b].\end{aligned}$$

$\mathbb{P}[S_k = 1]$

$k = 2\ell + 1$

$\Rightarrow$ must have ℓ steps -1

$\ell + 1$ steps $+1$

$2^{-(2\ell+1)} \binom{2\ell+1}{\ell}$

$b = 1$

$\frac{b}{k} = \frac{1}{k}$

$\mathbb{P}[\tau_1 = k] \sim \frac{1}{k^{3/2}}$

$\mathbb{E}[\tau_1] \approx \sum k \cdot \frac{1}{k^{3/2}} = \sum \frac{1}{\sqrt{k}}$ $\quad \mathbb{P}[S_k = 1] = \frac{c}{\sqrt{k}}$

Last return to 0

- Let $(S_n)_{n\geq 0}$ be a symmetric simple random walk starting from $S_0 = 0$.
- Let
$$\hat{\tau}_{2n} := \max\{0 \leq i \leq 2n : S_i = 0\}$$
denote the last time before $2n$ that the random walk visits 0.

Last return to 0

- Let $(S_n)_{n \geq 0}$ be a symmetric simple random walk starting from $S_0 = 0$.
- Let
$$\hat{\tau}_{2n} := \max\{0 \leq i \leq 2n : S_i = 0\}$$
denote the last time before $2n$ that the random walk visits 0.
- What is the distribution of $\hat{\tau}_{2n}$?

Last return to 0

- On the homework, you will use the reflection principle to show that for any integer $n > 0$,

$$\mathbb{P}[S_1 \neq 0, \ldots, S_{2n} \neq 0 \mid S_0 = 0] = \mathbb{P}[S_{2n} = 0].$$

Last return to 0

- On the homework, you will use the reflection principle to show that for any integer $n > 0$,

$$\mathbb{P}[S_1 \neq 0, \ldots, S_{2n} \neq 0 \mid S_0 = 0] = \mathbb{P}[S_{2n} = 0].$$

- Given this, we have

$$\mathbb{P}[\hat{\tau}_{2n} = 2k] = \mathbb{P}[S_{2k} = 0, S_{2k+1} \neq 0, \ldots, S_{2n} \neq 0]$$

Last return to 0

- On the homework, you will use the reflection principle to show that for any integer $n > 0$,

$$\mathbb{P}[S_1 \neq 0, \dots, S_{2n} \neq 0 \mid S_0 = 0] = \mathbb{P}[S_{2n} = 0]. \quad (*)$$

- Given this, we have

$$\begin{aligned}\mathbb{P}[\hat{\tau}_{2n} = 2k] &= \mathbb{P}[S_{2k} = 0, S_{2k+1} \neq 0, \dots, S_{2n} \neq 0] \\ &= \mathbb{P}[S_{2k} = 0]\mathbb{P}[S_{2k+1} \neq 0, \dots, S_{2n} \neq 0 \mid S_{2k} = 0]\end{aligned}$$

$$= \mathbb{P}[S_1 \neq 0 \dots S_{2n-2k} \neq 0 \mid S_0 = 0]$$

$$= \mathbb{P}[S_{2n-2k} = 0].$$

Last return to 0

- On the homework, you will use the reflection principle to show that for any integer $n > 0$,

$$\mathbb{P}[S_1 \neq 0, \ldots, S_{2n} \neq 0 \mid S_0 = 0] = \mathbb{P}[S_{2n} = 0].$$

- Given this, we have

$$\begin{aligned}\mathbb{P}[\hat{\tau}_{2n} = 2k] &= \mathbb{P}[S_{2k} = 0, S_{2k+1} \neq 0, \ldots, S_{2n} \neq 0] \\ &= \mathbb{P}[S_{2k} = 0]\mathbb{P}[S_{2k+1} \neq 0, \ldots, S_{2n} \neq 0 \mid S_{2k} = 0] \\ &= \mathbb{P}[S_{2k} = 0]\mathbb{P}[S_1 \neq 0, \ldots, S_{2n-2k} \neq 0 \mid S_0 = 0]\end{aligned}$$

Last return to 0

- On the homework, you will use the reflection principle to show that for any integer $n > 0$,

$$\mathbb{P}[S_1 \neq 0, \ldots, S_{2n} \neq 0 \mid S_0 = 0] = \mathbb{P}[S_{2n} = 0].$$

- Given this, we have

$$\begin{aligned}\mathbb{P}[\hat{\tau}_{2n} = 2k] &= \mathbb{P}[S_{2k} = 0, S_{2k+1} \neq 0, \ldots, S_{2n} \neq 0] \\ &= \mathbb{P}[S_{2k} = 0]\mathbb{P}[S_{2k+1} \neq 0, \ldots, S_{2n} \neq 0 \mid S_{2k} = 0] \\ &= \mathbb{P}[S_{2k} = 0]\mathbb{P}[S_1 \neq 0, \ldots, S_{2n-2k} \neq 0 \mid S_0 = 0] \\ &= \mathbb{P}[S_{2k} = 0]\mathbb{P}[S_{2n-2k} = 0].\end{aligned}$$

- Notice, in particular, that

$$\mathbb{P}[\hat{\tau}_{2n} = 2k] = \mathbb{P}[\hat{\tau}_{2n} = 2n - 2k](!)$$

Last return to 0

For large k, by Stirling's formula, we have

$$\mathbb{P}[S_{2k} = 0] \sim \frac{1}{\sqrt{\pi k}},$$

$\mathbb{P}[S_{2k+1} = 1] = \frac{c}{\sqrt{k}}$

so that

Last return to 0

For large k, by Stirling's formula, we have

$$\mathbb{P}[S_{2k} = 0] \sim \frac{1}{\sqrt{\pi k}},$$

so that

$$\mathbb{P}[\hat{\tau}_{2n} = 2k] = \mathbb{P}[S_{2k} = 0]\mathbb{P}[S_{2n-2k} = 0]$$

Last return to 0

For large k, by Stirling's formula, we have

$$\mathbb{P}[S_{2k} = 0] \sim \frac{1}{\sqrt{\pi k}},$$

so that

$$\begin{aligned} \mathbb{P}[\hat{\tau}_{2n} = 2k] &= \mathbb{P}[S_{2k} = 0]\mathbb{P}[S_{2n-2k} = 0] \\ &\sim \frac{1}{\sqrt{\pi k}} \frac{1}{\sqrt{\pi(n-k)}} \end{aligned}$$

Last return to 0

For large k, by Stirling's formula, we have

$$\mathbb{P}[S_{2k} = 0] \sim \frac{1}{\sqrt{\pi k}},$$

so that

$$\begin{aligned}\mathbb{P}[\hat{\tau}_{2n} = 2k] &= \mathbb{P}[S_{2k} = 0]\mathbb{P}[S_{2n-2k} = 0] \\ &\sim \frac{1}{\sqrt{\pi k}} \frac{1}{\sqrt{\pi(n-k)}} \\ &= \frac{1}{n} f\left(\frac{k}{n}\right),\end{aligned}$$

where

$$f(x) = \frac{1}{\pi\sqrt{x(1-x)}} \quad 0 < x < 1.$$

Last return to 0

Plot of $f(x) = \frac{1}{\pi\sqrt{x(1-x)}}$ for $0 < x < 1$.

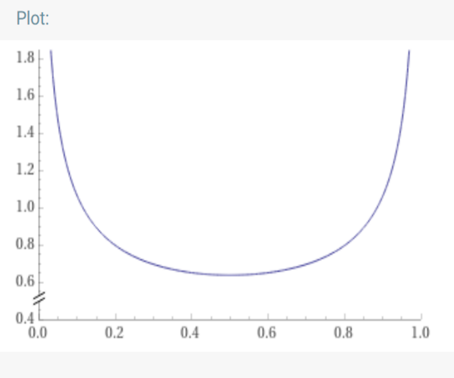


Image generated using Wolfram Alpha

The Arcsine Law

- For large k, n, we have

$$\mathbb{P}[\hat{\tau}_{2n} = 2k] \sim \frac{1}{n} f\left(\frac{k}{n}\right).$$

The Arcsine Law

- For large k, n, we have

$$\mathbb{P}[\hat{\tau}_{2n} = 2k] \sim \frac{1}{n} f\left(\frac{k}{n}\right).$$

- Therefore, for any $0 < x < 1$,

$$\mathbb{P}[\hat{\tau}_{2n} \leq 2nx] \sim \sum_{0 \leq k/n \leq x} \frac{1}{n} f\left(\frac{k}{n}\right)$$

The Arcsine Law

- For large k, n, we have

$$\mathbb{P}[\hat{\tau}_{2n} = 2k] \sim \frac{1}{n} f\left(\frac{k}{n}\right).$$

- Therefore, for any $0 < x < 1$,

$$\mathbb{P}[\hat{\tau}_{2n} \leq 2nx] \sim \sum_{0 \leq k/n \leq x} \frac{1}{n} f\left(\frac{k}{n}\right)$$

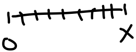
$$\sim \int_0^x f(t)dt$$

$$= \frac{2}{\pi} \cdot \arcsin \sqrt{x}.$$

$x \in (0,1)$

what is the prob that the last 0 is in first 10%

$\rightarrow \frac{2}{\pi} \cdot \arcsin \sqrt{0.1}$



The Arcsine Law

Plot of $g(x) = \frac{2}{\pi} \cdot \arcsin \sqrt{x}$ for $0 \le x \le 1$.

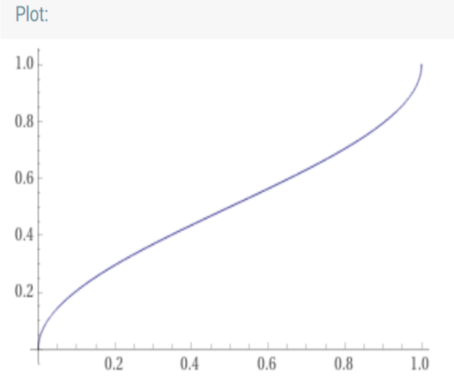


Image generated using Wolfram Alpha

The Arcsine Law

Some values of $g(x) = \frac{2}{\pi} \cdot \arcsin \sqrt{x}$. } arcsine distribution

$$x = 0.1 \quad g(x) \approx 0.204.$$
$$x = 0.2 \quad g(x) \approx 0.295.$$
$$x = 0.3 \quad g(x) \approx 0.369.$$
$$x = 0.4 \quad g(x) \approx 0.435.$$
$$x = 0.5 \quad g(x) = 0.5.$$
$$x = 0.6 \quad g(x) \approx 0.565.$$
$$x = 0.7 \quad g(x) \approx 0.631.$$
$$x = 0.8 \quad g(x) \approx 0.705.$$
$$x = 0.9 \quad g(x) \approx 0.796.$$

The Arcsine Law

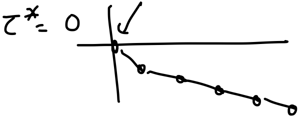
Using similar arguments, one can show that,

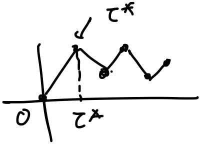
$$\mathbb{P}\left[\frac{|\{1 \le i \le n : S_i > 0\}|}{n} \le x\right] \to \frac{2}{\pi} \cdot \arcsin \sqrt{x}.$$

and also that

$$\mathbb{P}\left[\frac{t_*}{n} \le x\right] \to \frac{2}{\pi} \cdot \arcsin \sqrt{x},$$

where $0 \le t_* \le n$ is the first time when the random walk takes on its maximum value.





Some consequences of the Arc[redacted] Law

- In a (long) sequence of coin flips, the probability that heads leads 90% of the time is roughly 20%.

Some consequences of the Arcsine Law

- In a (long) sequence of coin flips, the probability that heads leads 90% of the time is roughly 20%.
- (Due to S. Dunbar) In other words, there is a 20% chance that a totally random investment fund has positive net fortune 90% of the time.

Some consequences of the Arcsine Law

- In a (long) sequence of coin flips, the probability that heads leads 90% of the time is roughly 20%.
- (Due to S. Dunbar) In other words, there is a 20% chance that a totally random investment fund has positive net fortune 90% of the time.
- (Clauset, Kogan, Redner, 2015) In e.g. professional basketball, the distribution of the last lead change and time of the maximum lead change follow the arcsine law, which is what is predicted by a symmetric simple random walk.

Physics Review E

Some consequences of the Arcsine Law

- In a (long) sequence of coin flips, the probability that heads leads 90% of the time is roughly 20%.
- (Due to S. Dunbar) In other words, there is a 20% chance that a totally random investment fund has positive net fortune 90% of the time.
- (Clauset, Kogan, Redner, 2015) In e.g. professional basketball, the distribution of the last lead change and time of the maximum lead change follow the arcsine law, which is what is predicted by a symmetric simple random walk.
- In particular, lead changes are most likely near the start and the end of the game. Similarly, the time of the largest lead is most likely to be near the start and the end of the game.