

# Discrete Random Walks 07-03-16

N. T. Gladd

**Initialization:** Be sure the files *NTGStylesheet2.nb* and *NTGUtilityFunctions.m* is are in the same directory as that from which this notebook was loaded. Then execute the cell immediately below by mousing left on the cell bar to the right of that cell and then typing “shift” + “enter”. Respond “Yes” in response to the query to evaluate initialization cells.

In[101]:=

```
SetDirectory[NotebookDirectory[]];
(* set directory where source files are located *)
SetOptions[EvaluationNotebook[], (* load the StyleSheet *)
  StyleDefinitions -> Get["NTGStylesheet2.nb"]];
Get["NTGUtilityFunctions.m"]; (* Load utilities package *)
```

## Purpose

Motivated by reading the very interesting *Safe Leads and Lead Changes in Competitive Sports*, Clauset, Kogan and Redner, I have spent time on issues related to the arcsine distributions encountered in random walk phenomena. The paper points out that these distributions are relevant to NBA basket ball campaigns. It seemed to me, these distributions are also relevant to the PL history of option trading campaigns. I have previously written of applications probability theory to trading campaigns and wanted to add to this corpus. There are arcsine distributions for the amount of time a trade campaign is profitable (or losing), the time in the campaign at which the last return to zero PL would occur, and the time at which the maximum PL would occur. I struggled with some of the texts and articles relevant to this topic, so there was a motivation to review and improve my capabilities in discrete mathematics.

I looked at various internet articles and lectures as well as the texts --  
 Probability and Random Processes, Grimmett and Stirzaker  
 Introduction to Probability, Grinstead and Snell (downloaded copy)  
 Probability Theory and Stochastic Processes with Applications, Knill (downloaded copy)  
 Probability Notes, Weber (downloaded copy)  
 An Introduction to Probability Theory and its Applications, Vols. I and II, Feller (downloaded copy)

As is often the case for me, mathematical notation and formal presentation of topics was a challenge. My background in physics predisposes me toward derivations and problem solving as opposed to the lemmas, theorems and formal proofs. After a few forays, I settled on the classical text by Feller. This is a good book and also provided me some comfort by discussing the history of some results. Some random walk proofs are nonintuitive and were actually developed much later than I would have

expected. I don't feel as bad that my standard physicsy procedure of starting from first principles and moving relatively straightforwardly toward some result ran into obstacles. This notebook consists of working through the details of some of Feller's results on discrete random walks. To facilitate my understanding, I constructed several visualizations of random walk phenomena.

## I Basics

I follow the development of some selected topics in Feller Volume 1, Chapter III. Consistent with the trading campaign motivation discussed above I try to cast the discussion toward the cumulative profit loss path of a series of trades, using wealth  $W$  as the dependent variable. In the following the analogy of the trade campaign with a random walk is that each winning trade occurs with probability  $1/2$  and results in a gain of 1 wealth unit. Some results immediately generalize to arbitrary probability of success.

I specifically work toward the arcsine law for lead changes in a random walk. Feller establishes several intermediate results that figure in the proof of the arcsine law, as well as being of interest in their own right.

### I.1 Ballot theorem (number of ways a streak of winning trades can occur)

This is a fundamental result for random walks that is useful in many contexts.

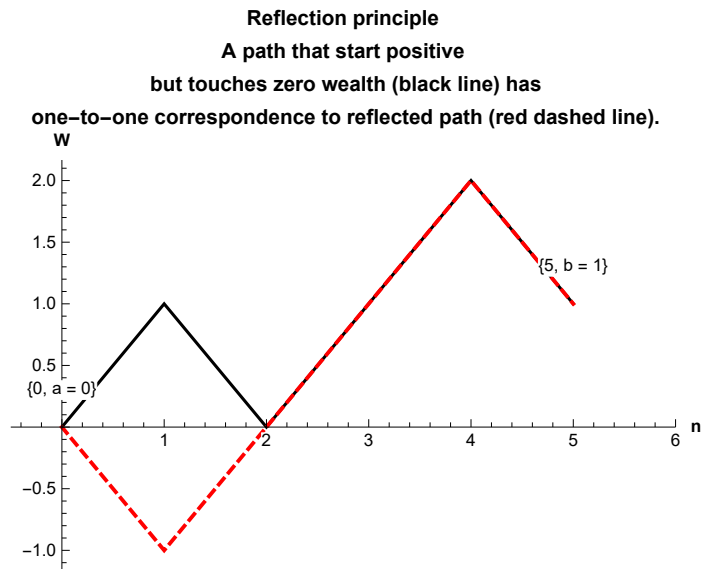
**Theorem 1.** (*Ballot theorem.*) *Let  $x > 0, y > 0$ ; the number of paths  $\{s_1, s_2, \dots, s_x = y\}$  from the origin to  $(x, y)$  such that  $s_1 > 0, s_2 > 0, \dots, s_x > 0$  equals  $(y/x)N_{x,y}$ .*

I follow Feller's arguments but use my own notation. The reflection principle is important.

In[103]=

```
Module[{offset = {0, 0.3}, a = 0, b = 1, n = 5, path, reflectedPath, text, lab},
  path = {{a, 0}, {1, 1}, {2, 0}, {3, 1}, {4, 2}, {5, b}};
  reflectedPath = {{a, 0}, {1, -1}, {2, 0}, {3, 1}, {4, 2}, {5, b}};
  lab =
    Stl["Reflection principle\nA path that start positive\n but touches zero wealth
      (black line) has\n one-to-one correspondence
      to reflected path (red dashed line)."];
  text["a"] = Text["{0, a = 0}", offset];
  text["b"] = Text["{5, b = 1}", {5, 1} + offset];
  ListPlot[{path, reflectedPath}, Joined → True,
    Epilog → {text["a"], text["b"]},
    PlotStyle → {Black, Directive[Thick, Red, Dashed]},
    AxesLabel → {Stl["n"], Stl["W"]}, PlotLabel → lab,
    PlotRange → {{-0.5, 6}, Automatic}]]
```

Out[103]=



A key question is how many paths lead from  $a$  to  $b$ . In order to reach  $b$  in  $n$  trades, there must be  $b - a$  winning trades or  $u$  steps upward.

In[104]=

```
Solve[{u + d == n, u - d == b - a}, {u, d}] [[1]]
```

Out[104]=

```
{u →  $\frac{1}{2}(-a + b + n)$ , d →  $\frac{1}{2}(a - b + n)$ }
```

The number of paths ending at  $b$  is the number of ways  $u$  things can be chosen from  $n$  things  $\binom{n}{u}$

The ballot theorem concerns the number of paths for which the wealth does not return to zero.

In the diagram the total number of paths from  $\{0, a = 0\}$  to  $\{n, b\}$  is

$$\mathcal{N}_n(a, b) = \binom{n}{\frac{1}{2}(-a + b + n)}$$

The number  $\mathcal{N}_n(a, b)$  includes paths that return to the origin. The first step of any path that doesn't return to the origin must be  $\{0, a = 0\} \rightarrow \{1, 1\}$ . By the reflection principle, each path that starts at  $\{1, 1\}$  and touches the origin and ends at  $\{n, b\}$  must have an analog "reflected" path that starts at  $\{-1, 1\}$  and ends at  $\{n, b\}$ . The number of such reflections of the paths that return to the origin is

$$\mathcal{N}_{n-1}(-1, b) = \binom{n-1}{\frac{b+n}{2}}$$

Thus the number of paths that don't return to the origin is

$$\begin{aligned} \mathcal{N}_n^{\text{dont return to } a}(a, b) &= \mathcal{N}_n(1, b) - \mathcal{N}_n(-1, b) \\ &= \binom{n-1}{\frac{1}{2}(b+n-2)} - \binom{n-1}{\frac{b+n}{2}} \\ &= \binom{n-1}{k-1} - \binom{n-1}{k} \\ &= \frac{(n-1)!}{(k-1)!(n-1-(k-1))!} - \frac{(n-1)!}{k!(n-1-k)!} \\ &\text{after some algebra} \\ &= \frac{b}{n} \binom{n}{\frac{b+n}{2}} \end{aligned}$$

which is the ballot theorem path count.

## 1.2 Number of paths of length $2n$ that connect $a = 0$ and $b = 0$ while always being positive

In developing the properties of random walks an important quantity is the number of paths that returns to an initial wealth level of 0, i.e., return to the starting level  $a$ ).

A return to 0 can only occur after an even number of steps, so the number of paths is usually denoted  $2n$  with  $n$  an integer.

### Criterion 1.2 Feller 2.4

$$\begin{aligned} W_0 = 0, W_1 > 0, W_2 > 0, \dots, W_{2n-1} > 0, W_{2n} = 0 \\ \text{or, more explicitly,} \\ W_0 = 0, W_1 = 1, W_2 = 2, \dots, W_{2n-2} = 2, W_{2n-1} = 1, W_{2n} = 0 \end{aligned}$$

Each of the paths that returns to  $\{2n, 0\}$  must pass through  $\{2n-1, 1\}$ . Applying the ballot theorem

In[105]:=

```
TraditionalForm@With[{nSteps = 2 n - 1, b = 1},
  Ballot[nSteps, 1]]
```

Out[105]//TraditionalForm=

$$\frac{\binom{2n-1}{n}}{2n-1}$$

Feller introduces a convenient quantity which related to the ballot theorem number of paths.

$$L_{2n} = \frac{1}{n+1} \binom{2n}{n}$$

I initially tried to avoid the use of quantities such as  $L_{2n}$  (and the  $u_{2n}$  and  $f_{2n}$  to follow) but they really do simplify the exposition.

Specifically, the ballot theorem path count can be represented as  $L_{2n-2}$ .

In[106]:=

```
Module[{results, info},
  results = Table[{n, 2n, Ballot[2n-1, 1], L[2n-2]}, {n, 1, 5}];
  info = PrependTo[results, {"n", "2n (length of path)", "ballot", "L_{2n-2}"}];
  LGrid[info, "Number paths with\nW_0=0, W_1>0, ..., W_{2n-1}>0, W_{2n}=0"]]
```

**Number paths with**

**$W_0=0, W_1>0, \dots, W_{2n-1}>0, W_{2n}=0$**

n	2n (length of path)	ballot	$L_{2n-2}$
1	2	1	1
2	4	1	1
3	6	2	2
4	8	5	5
5	10	14	14

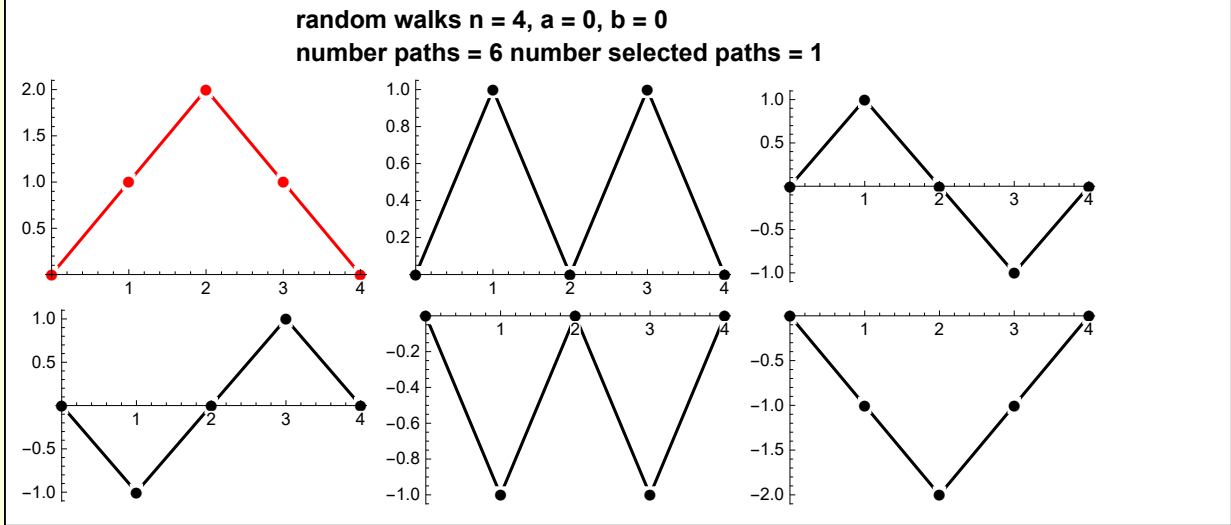
Out[106]=

I develop some visualizations of the  $n=2, 3$  cases. The black paths denote all possible paths, the red paths are the first return paths with all intermediate values being positive.

In[112]=

```
ShowSelectedPaths[4, 0, 0, (Min[#[[2 ;; -2]]] > 0) &]
```

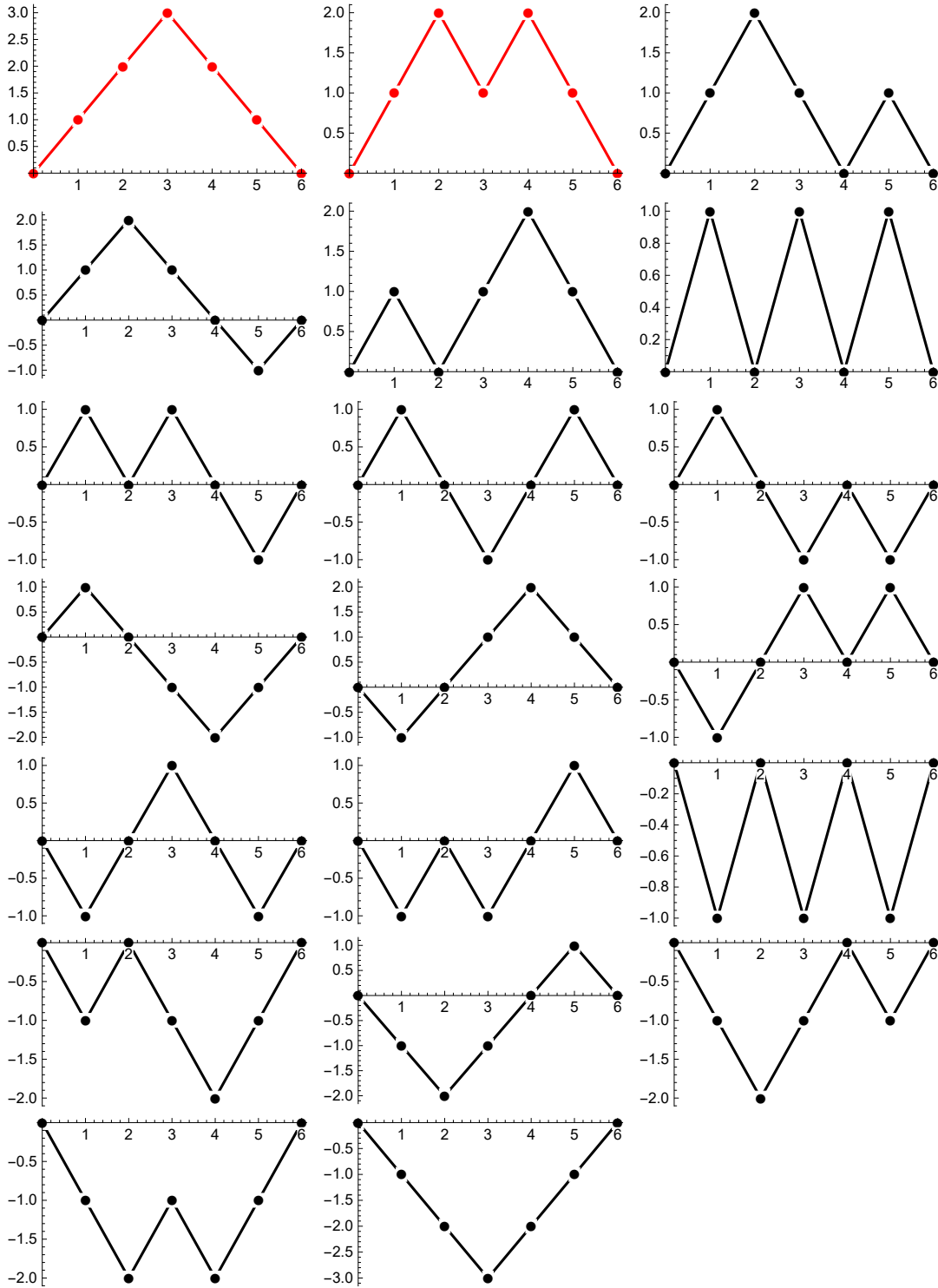
Out[112]=



In[113]:=

```
ShowSelectedPaths[6, 0, 0, (Min[#[[2 ;; -2]]] > 0) &]
```

random walks n = 6, a = 0, b = 0  
number paths = 20 number selected paths = 2



Out[113]=

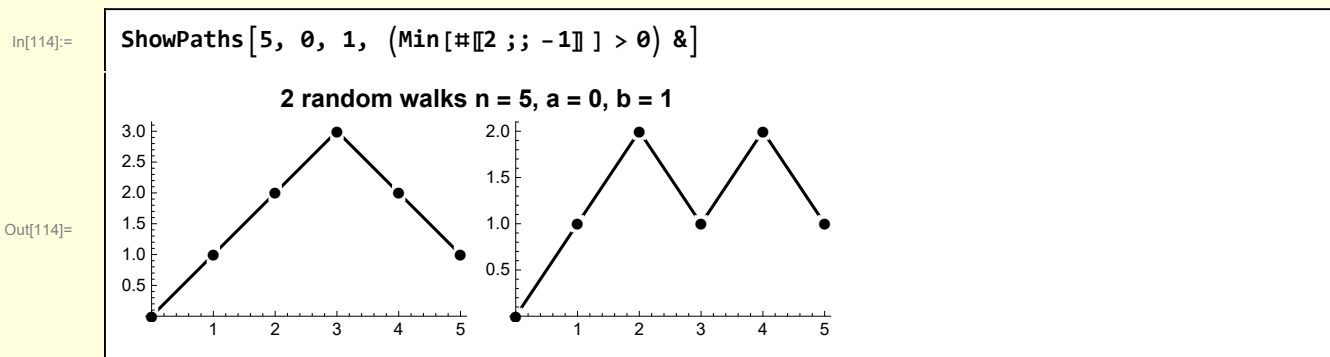
### 1.3 Number of paths of length $2n$ with end points $a = 0$ $b = 0$ with all intermediate steps being $\geq 0$

**Criterion 1.3 Feller 2.5**

$W_0 = 0, W_1 \geq 0, W_2 \geq 0, \dots, W_{2n-1} \geq 0, W_{2n} = 0$   
 or, more explicitly,  
 $W_0 = 0, W_1 = 1, W_2 \geq 0, \dots, W_{2n-2} \geq 0, W_{2n-1} = 1, W_{2n} = 0$

These paths have a one-to-one correspondence with the case just considered.

For example, consider the case for  $n = 6$  in section 1.3 where there were two paths out of 20 possible paths that satisfied the criterion.



Consider the portion of these paths immediately above from  $\{1,1\}$  to  $\{n-1,1\}$  and note that they satisfy the conditions that all values are  $\geq 1$ . If these paths are translated to  $\{0,0\}$  to  $\{n-2, 0\}$  it satisfies the condition that all values are  $\geq 0$ .

For example, consider paths that are only 4 steps long. From Feller's formula we see the number of paths satisfying Criterion 1.3 is  $L_{2n}$

In[115]:= `With[{n = 2},  
 L[2 n]]`

Out[115]= 2

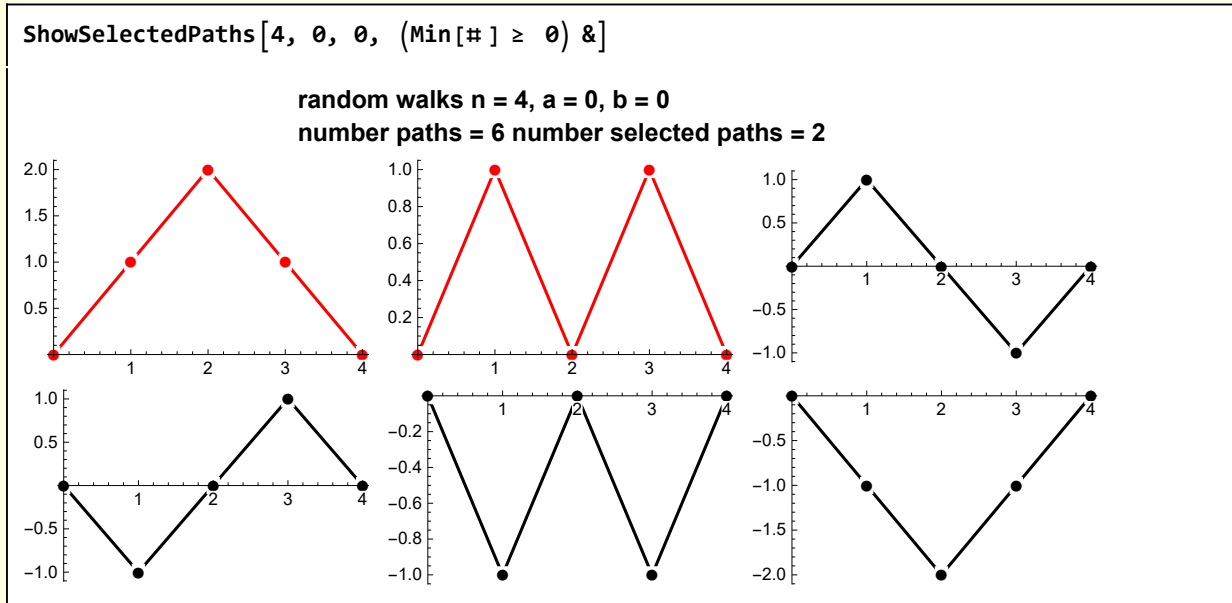
I illustrate those two paths



In[116]:=

ShowSelectedPaths[4, 0, 0, (Min[#] ≥ 0) &amp;]

Out[116]=



## 2 Feller's formulas related to first returns for flips of a fair coin

Feller develops some results for fair coin flips,  $p = q = 1/2$ .

Some terminology

**First return is the event**

$$W_0 = 0, W_1 \neq 0, W_2 \neq 0, \dots, W_{n-1} \neq 0, W_n = 0$$

**First passage through  $k > 0$  is the event**

$$W_0 = 0, W_1 < k, W_2 < k, \dots, W_{n-1} < k, W_n = k$$

Define some quantities  $u_{2n}$  and  $f_{2n}$  convenient for discussing the probability of events associated with fair coin random walks

$$u_{2n} \equiv \binom{2n}{n} \left(\frac{1}{2}\right)^{2n}$$

$$f_{2n} \equiv \frac{1}{2n} u_{2n-2} \quad f_0 \equiv 0$$

with

$$f_{2n} = u_{2n-2} - u_{2n}$$

In[117]=

```
Module[{results, info},
  results = Table[{n, 2 n, IPu[2 n], IPu[2 n] 22 n}, {n, 1, 5}];
  info = PrependTo[results, {"n", "2 n", "u2 n", "u2 n22 n"}];
  LGrid[info, "Examples of u2 n"]]
```

Examples of  $u_{2n}$ 

n	2 n	$u_{2n}$	$u_{2n}2^{2n}$
1	2	$\frac{1}{2}$	2
2	4	$\frac{3}{8}$	6
3	6	$\frac{5}{16}$	20
4	8	$\frac{35}{128}$	70
5	10	$\frac{63}{256}$	252

Out[117]=

The quantities  $u_{2n}$  and  $f_{2n}$  are related to random walks in the following theorem.

**Theorem 1.** For each  $n \geq 1$ :

$$(4.4) \quad u_{2n} = \mathbf{P}\{s_{2n} = 0\}$$

III.4]

REFORMULATION

75

$$(4.5) \quad u_{2n} = \mathbf{P}\{s_1 \neq 0, s_2 \neq 0, \dots, s_{2n} \neq 0\}$$

$$(4.6) \quad u_{2n} = \mathbf{P}\{s_1 \geq 0, s_2 \geq 0, \dots, s_{2n} \geq 0\}$$

or in words: The three events, (a) a return to the origin takes place at time  $2n$ , (b) no return occurs up to and including time  $2n$ , and (c) the path is non-negative between 0 and  $2n$ , have the common probability  $u_{2n}$ .

Furthermore,

$$(4.7) \quad f_{2n} = \mathbf{P}\{s_1 \neq 0, s_2 \neq 0, \dots, s_{2n-1} \neq 0, s_{2n} = 0\}$$

$$(4.8) \quad f_{2n} = \mathbf{P}\{s_1 \geq 0, s_2 \geq 0, \dots, s_{2n-2} \geq 0, s_{2n-1} < 0\}$$

that is: the two events (a) the first return to the origin takes place at time  $2n$ , and (b) the first passage through  $-1$  occurs at time  $2n - 1$ , have the common probability  $f_{2n}$ .

**Criterion 2.1a Feller 4.7**

$$W_0 = 0, W_1 \neq 0, W_2 \neq 0, \dots, W_{2n-1} \neq 0, W_{2n} = 0$$

First return occurs at  $2n$ .

From 1.2, there were  $L_{2n-2}$  paths that satisfied

$$W_0 = 0, W_1 > 0, W_2 > 0, \dots, W_{2n-1} > 0, W_{2n} = 0$$

By symmetry there are another  $L_{2n-2}$  paths that satisfy

$$W_0 = 0, W_1 < 0, W_2 < 0, \dots, W_{2n-1} < 0, W_{2n} = 0$$

Number paths with  
 $W_0=0, W_1 \neq 0, \dots, W_{2n-1} \neq 0, W_{2n}=0$

n	(length of path)	$2 L_{2n-2}$	$\mathbb{P}f_{2n} 2^{2n}$
1	2	2	2
2	4	2	2
3	6	4	4
4	8	10	10
5	10	28	28

The probability of first return at  $2n$  is

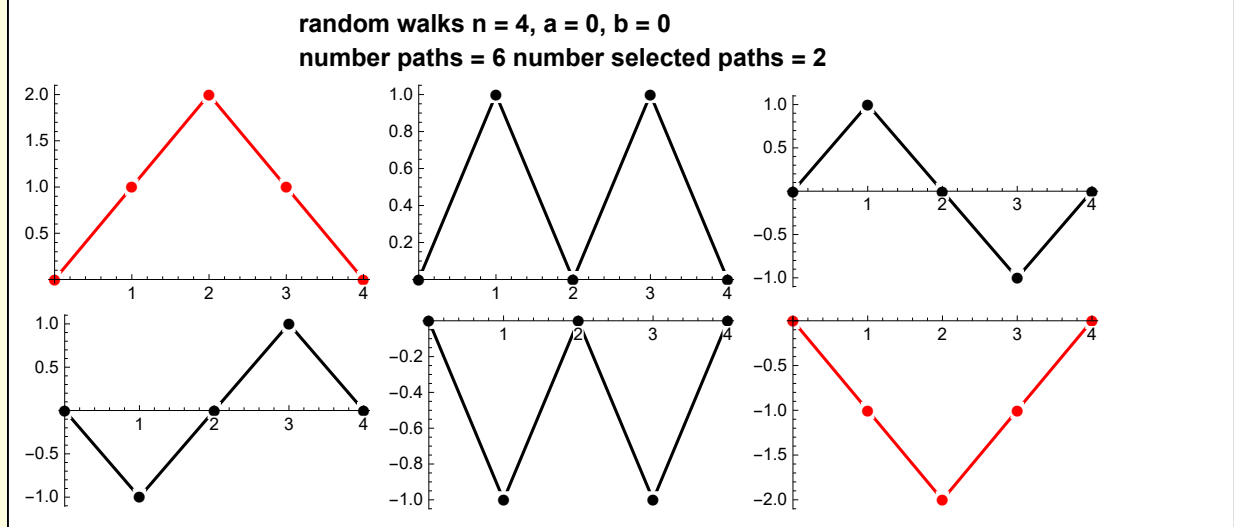
$$(L_{2n-2} + L_{2n-2}) \left(\frac{1}{2}\right)^{2n} = f_{2n}$$

Example first return paths (red) for the case  $n = 2$

In[118]:=

```
ShowSelectedPaths[4, 0, 0, (FreeQ[#, 2 ;; -2], 0) &]
```

Out[118]=



**Criterion 2.1b Feller 4.8**

$$W_0 = 0, W_1 \geq 0, W_2 \geq 0, \dots, W_{2n-1} \geq 0, W_{2n} < 0$$

This is a straightforward analog of Criterion 1.3. See above

**Criterion 2.1c Feller 4.5**

$$W_0 = 0, W_1 \neq 0, W_2 \neq 0, \dots, W_{2n-1} \neq 0, W_{2n} \neq 0$$

Paths that don't touch zero. This is  $1 - \mathbb{P}[\text{no returns}]$  at  $i = 2, 4, \dots, 2n$

$$1 - f_2 - f_4 - \dots - f_{2n}$$

$$= 1 - (u_0 - u_2) - (u_2 - u_4) \dots - (u_{2n-2} - u_{2n})$$

$$= u_{2n}$$

For example

Number paths with  $W_0=0, W_1 \neq 0, \dots, W_{2n-1} \neq 0, W_{2n} \neq 0$

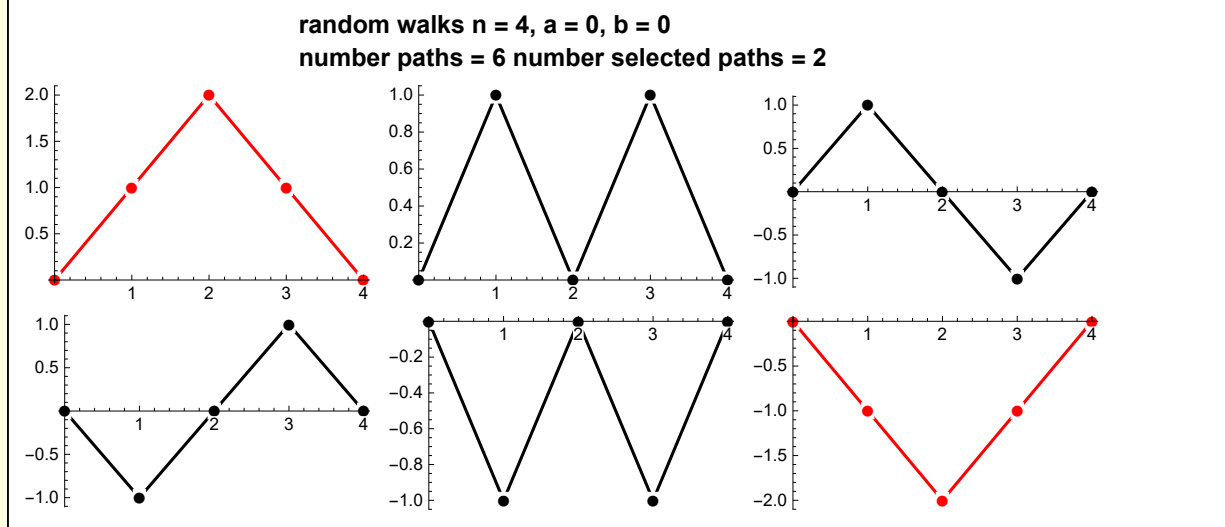
n	(length of path)	$P_{u_{2n}} 2^{2n}$
1	2	2
2	4	6
3	6	20
4	8	70
5	10	252

The following show examples of no return paths (red) for  $n = 4$

In[128]:=

```
ShowSelectedPaths[4, 0, 0, (FreeQ[#, [2 ;; -2]], 0) &]
```

Out[128]=



**Criterion 2.1C Feller 4.6**

Paths that are always positive or zero

$$W_0 = 0, W_1 \geq 0, W_2 \geq 0, \dots, W_{2n-1} \geq 0, W_{2n} \geq 0$$

Number paths with  $W_0=0, W_1 \geq 0, \dots, W_{2n-1} \geq 0, W_{2n} \geq 0$

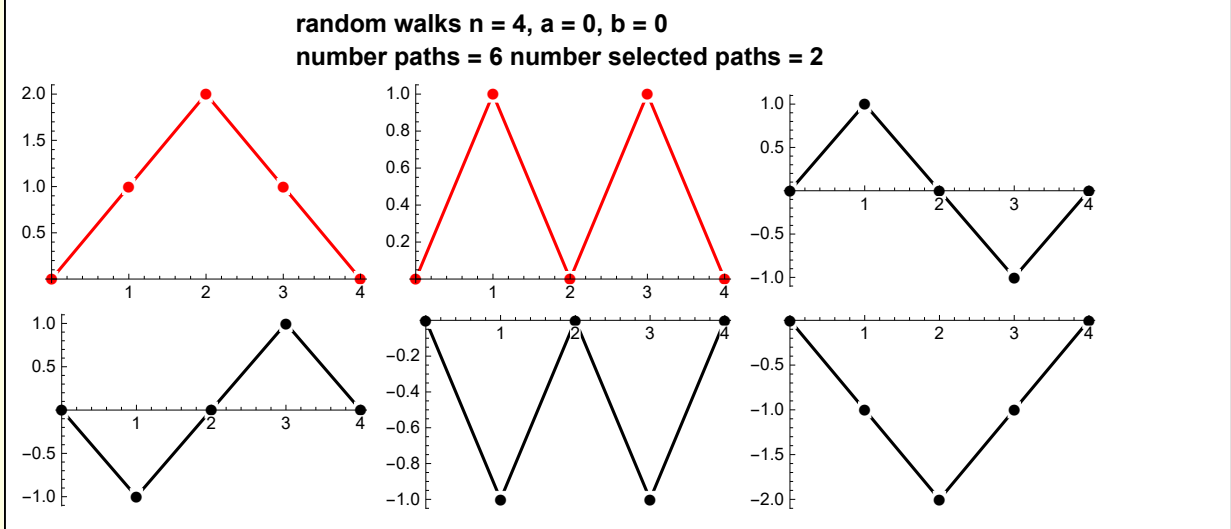
n	(length of path)	$Pu_{2n}2^{2n}$
1	2	2
2	4	6
3	6	20
4	8	70
5	10	252

The following show examples of always positive paths for n = 4

In[130]:=

ShowSelectedPaths[4, 0, 0, (Min[#[[2 ;; -1]]] >= 0) &]

Out[130]=



### 3 Probability that a random walk of length 2n has 2k positive intervals and n-2k negative intervals

Feller introduces the quantity  $p_{2k, 2n-2k}$

$$p_{2k, 2n-2k} = u_{2k} u_{2n-2k}$$

Examples of  $p_{2k, 2n}$

k	2k	n	2n	$p_{2k, 2n}$	$p_{2k, 2n}2^{2n}$
1	2	2	4	$\frac{1}{4}$	4
2	4	2	4	$\frac{3}{8}$	6

It is useful to visualize random walks for which 2k steps are positive and 2n-2k are negative.

I introduce some functions for identifying patterns in random walks

In[81]:=

```
Clear[PositiveIntervalQ, NumberPositiveNegativeIntervals];
PositiveIntervalQ[{Wi_, Wip1_}] :=
  If[Or[Wi > 0, Wip1 > 0], 1, 0];

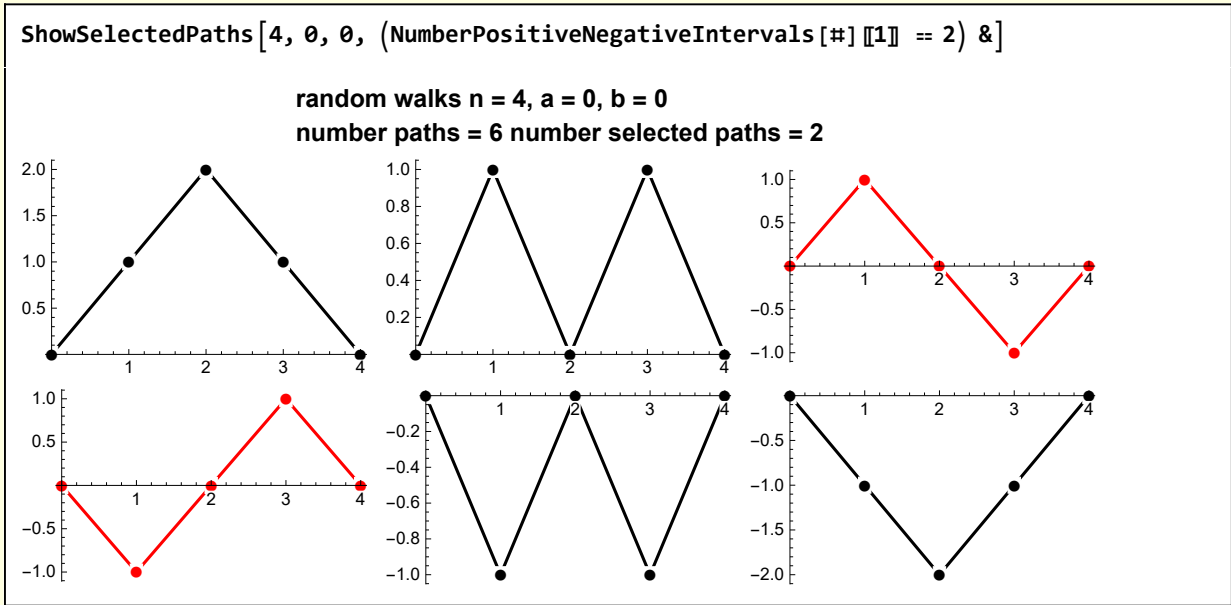
NumberPositiveNegativeIntervals[path_] :=
  Module[{nPositive},
    nPositive = PositiveIntervalQ /@ Partition[path, 2, 1] // Total;
    {nPositive, Length[path] - 1 - nPositive}
```

Examples for  $n = 2, k = 1$ . The red paths occur with probability  $p_{2,2}$

In[132]:=

```
ShowSelectedPaths[4, 0, 0, (NumberPositiveNegativeIntervals[#] [[1]] == 2) &]
```

Out[132]=



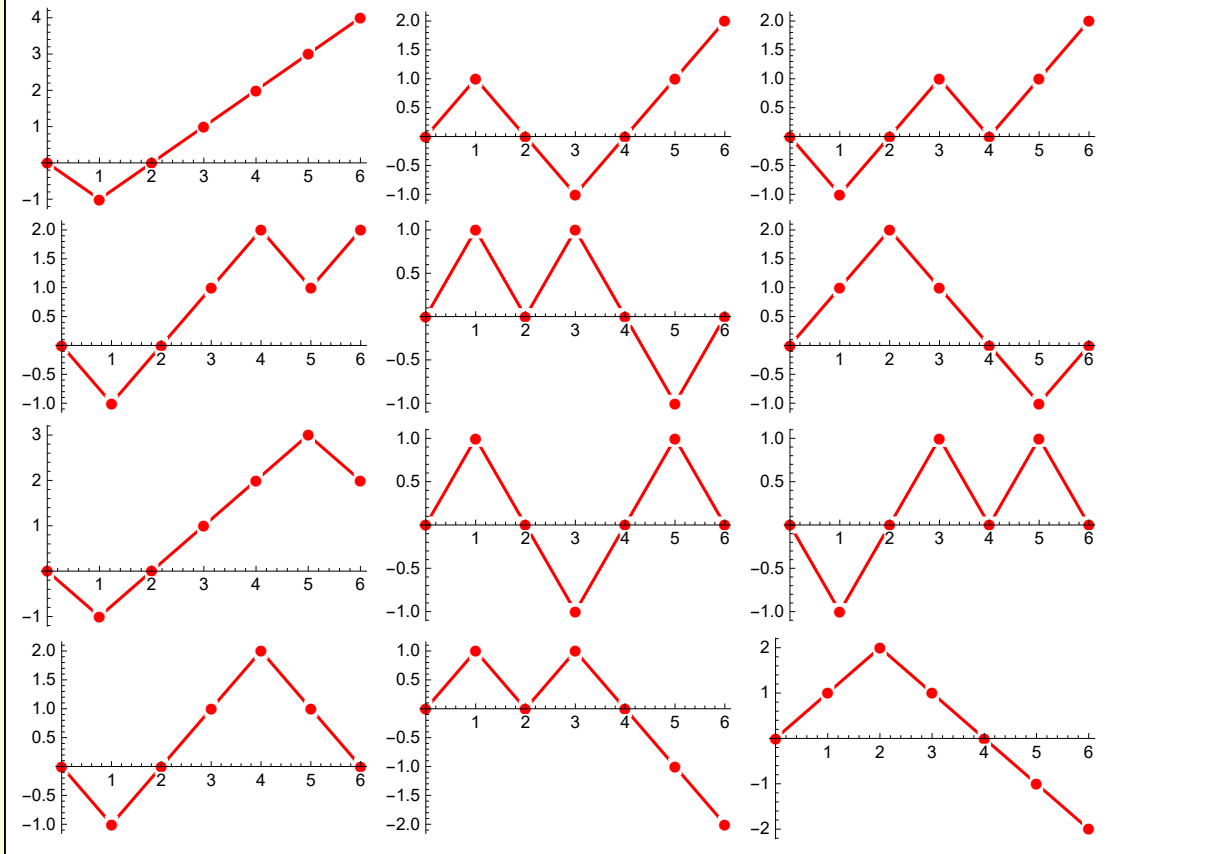
Examples for  $n = 3, k = 1$ . There are  $2^6$  paths in all. I show only the paths that occur with probability  $p_{2,4}$

In[133]=

```
ShowSelectedPathsOnly[6, And[(NumberPositiveNegativeIntervals[#][[1]] == 4),
(NumberPositiveNegativeIntervals[#][[2]] == 2) &]
```

**n = 6 generates 64 random walks  
12 selected paths**

Out[133]=



## 4 Proof by induction that $p_{2k,2n-2k}$ corresponds to the probability that a rw will have $2k$ positive intervals and $2n - 2k$ negative intervals

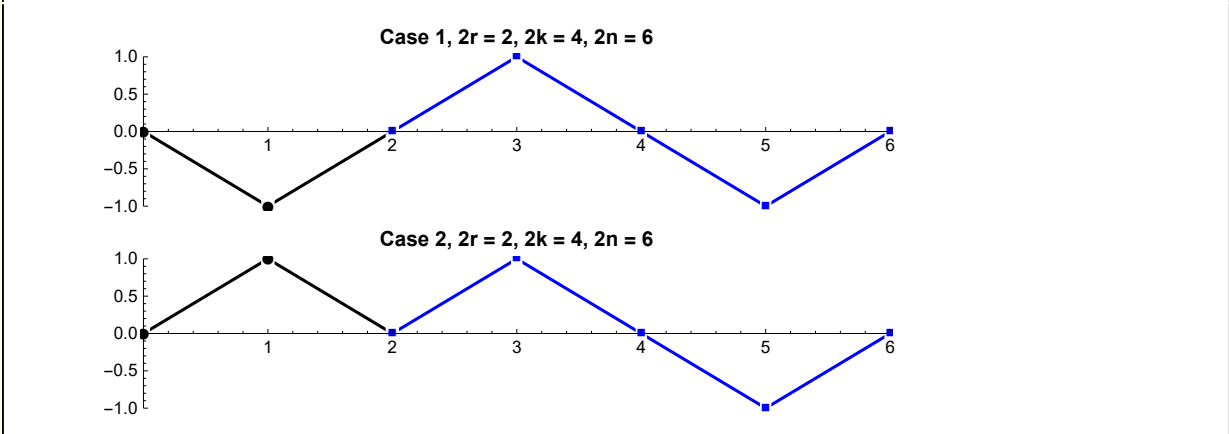
Consider a rw of length  $2n$  for which the first return to the origin occurs at  $2r$ . Two examples of such path for  $r = 2$  and  $n = 6$  are

In[134]=

```
Module[{examplePath, lab, g},
  examplePath[1] = Transpose[{Range[0, 6], {0, -1, 0, 1, 0, -1, 0}}];
  examplePath[2] = Transpose[{Range[0, 6], {0, 1, 0, 1, 0, -1, 0}}];
  lab = Stl@StringForm["Case 1, 2r = 2, 2k = 4, 2n = 6"];
  g[1] = ListLinePlot[{examplePath[1][[1 ;; 3]], examplePath[1][[3 ;; 7]]},
    PlotMarkers -> Automatic, PlotStyle -> {Black, Blue},
    PlotLabel -> lab, ImageSize -> {500, 100}, PlotRangePadding -> None,
    AspectRatio -> 0.20, Frame -> {False, False, False, False}];

  lab = Stl@StringForm["Case 2, 2r = 2, 2k = 4, 2n = 6"];
  g[2] = ListLinePlot[{examplePath[2][[1 ;; 3]], examplePath[2][[3 ;; 7]]},
    PlotMarkers -> Automatic, PlotStyle -> {Black, Blue},
    PlotLabel -> lab, ImageSize -> {500, 100}, PlotRangePadding -> None,
    AspectRatio -> 0.20, Frame -> {False, False, False, False}];
  Grid[{{g[1]}, {g[2]}}]]
```

Out[134]=



There are two cases with equal probability — a first return (0, 2r) with positive intervals and a first return with negative intervals — the black lines in the diagram above. Each of these two cases have probability  $\frac{1}{2} f_{2r}$ . The corresponding number of such first return paths is  $\frac{1}{2} f_{2r} 2^{2r}$ .

The objective is to determine the probability of paths for which 2k intervals are positive and 2n - 2k are negative over the entire path —  $p_{2k, 2n-2k}$ .

Consider case 1 first. The black part of the path, which has only negative intervals, has probability  $\frac{1}{2} f_{2r}$ . The second part of the path (the blue part) has both positive and negative intervals with probability  $p_{2k, 2n-2r}$ . Thus, the number of paths for which there are 2k positive intervals between 0 and 2n is

$$\left(\frac{1}{2} f_{2r} 2^{2r}\right) (p_{2k, 2n-2r}) 2^{2n-2r} = \left(\frac{1}{2} f_{2r}\right) (p_{2k, 2n-2r}) 2^{2n}$$

$$= \left(\frac{1}{2} f_{2r}\right) u_{2k} u_{2n-2r-2k} 2^{2n}$$

Consider case 2 for which the first return path is positive but there are additional positive intervals in the blue part of the path. As before, the number of black return paths is  $\frac{1}{2} f_{2r} 2^{2r}$ . For the second part of



the path we have the number of paths is  $\mathbb{P}_{2k-2r, 2n-2r} 2^{2n-2r}$ .

Thus, the number of paths for which there are  $2k$  positive intervals for the path 0 to  $2n$  is

$$\begin{aligned} & \left(\frac{1}{2} f_{2r} 2^{2r}\right) (\mathbb{P}_{2k-2r, 2n-2r}) 2^{2n-2r} = \left(\frac{1}{2} f_{2r}\right) (\mathbb{P}_{2k-2r, 2n-2r}) 2^{2n} \\ & = \left(\frac{1}{2} f_{2r}\right) u_{2k-2r} u_{2n-2r-(2k-2r)} 2^{2n} \\ & = \left(\frac{1}{2} f_{2r}\right) u_{2k-2r} u_{2n-2k} 2^{2n} \end{aligned}$$

Combining the two cases

$$\mathbb{P}_{2k, 2n-2k} 2^{2n} = \sum_{r=1}^{n-k} \left(\frac{1}{2} f_{2r}\right) u_{2k} u_{2n-2r-2k} 2^{2n} + \sum_{r=1}^k \left(\frac{1}{2} f_{2r}\right) u_{2k-2r} u_{2n-2k} 2^{2n}$$

or

$$\mathbb{P}_{2k, 2n-2k} = u_{2k} \sum_{r=1}^{n-k} \left(\frac{1}{2} f_{2r}\right) u_{2n-2r-2k} + u_{2n-2k} \sum_{r=1}^k \left(\frac{1}{2} f_{2r}\right) u_{2k-2r}$$

Recall the relationship

$$u_{2n} = \sum_{i=1}^{2n} f_{2i} u_{2n-2i}$$

But, by hypothesis, the over probability was supposed to be  $p_{2k, 2n-2k}$  so

$$\begin{aligned} \mathbb{P}_{2k, 2n-2k} & = u_{2k} \sum_{r=1}^{n-k} \left(\frac{1}{2} f_{2r}\right) u_{2n-2r-2k} + u_{2n-2k} \sum_{r=1}^k \left(\frac{1}{2} f_{2r}\right) u_{2k-2r} \\ & = \frac{1}{2} u_{2k} u_{2n-2k} + \frac{1}{2} u_{2n-2k} u_{2k} \\ & = u_{2k} u_{2n-2k} \end{aligned}$$

as expected

With the probability  $p_{2k, 2n-2k}$  established, I can calculate the probability that path remains positive after  $2k$  steps. In the trading context I am considering, this is the probability that a sequence of  $2n$  trades will be profitable after  $2k$  trades

$$\mathbb{P}_{2n}^{2k} = \sum_{i=1}^k p_{2i, 2n-2i} = \sum_{i=1}^k u_{2i} u_{2n-2i}$$

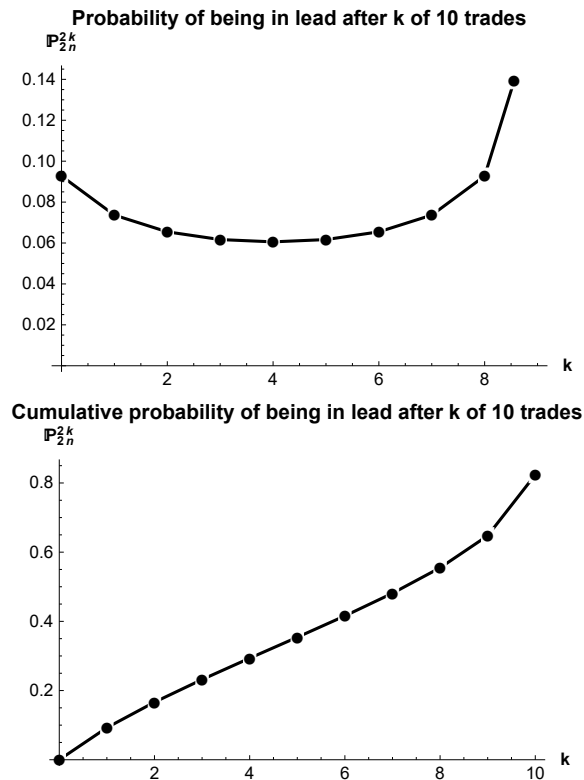
In[135]=

```

Module[{n = 10, results, cumResults, lab, DisplayPath, g},
  DisplayPath[path_] := Thread[List[Range[0, Length[path] - 1], path]];
  results = Table[IPu[2 k] IPu[2 n - 2 k], {k, 1, n}];
  cumResults = FoldList[Plus, 0, results];
  lab = Stl@StringForm["Probability of being in lead after k of `` trades", n];
  g[1] =
    ListLinePlot[{DisplayPath[results]}, PlotMarkers -> Automatic, PlotStyle -> Black,
      AxesLabel -> {Stl["k"], Stl["P_{2n}^{2k}"]}, PlotLabel -> lab, ImageSize -> {500, 200}];
  lab = Stl@StringForm["Cumulative probability of being in lead after k of `` trades",
    n];
  g[2] = ListLinePlot[{DisplayPath[cumResults]}, PlotMarkers -> Automatic,
    PlotStyle -> Black, AxesLabel -> {Stl["k"], Stl["P_{2n}^{2k}"]},
    PlotLabel -> lab, ImageSize -> {500, 200}];
  Grid[{{g[1]}, {g[2]}}]]

```

Out[135]=



## 5 Arcsine law

The arcsine functional form emerges in the limit that  $n \gg 1$ . Recall

$$P_{2n}^{2k} = \sum_{i=1}^k p_{2i, 2n-2i} = \sum_{i=1}^k u_{2i} u_{2n-2i}$$

The term  $u_{2j}$  is

In[143]=

```
Clear[i];
w5[1] = Pu[2 i]
```

Out[144]=

$$2^{-2i} \text{Binomial}[2i, i]$$

In[145]=

```
w5[2] = w5[1] /. Binomial[a_, b_] ->  $\frac{a!}{b! (a-b)!}$ 
```

Out[145]=

$$\frac{2^{-2i} (2i)!}{(i!)^2}$$

Invoke Sterling's approximation

In[146]=

```
w5[3] = w5[2] /. (a_)! -> a^a Exp[-a] Sqrt[2 pi a]
```

Out[146]=

$$\frac{1}{\sqrt{i} \sqrt{\pi}}$$

Performing these operations on the entire summand

In[147]=

```
w5[4] = Pu[2 i] Pu[2 n - 2 i] /. Binomial[a_, b_] ->  $\frac{a!}{b! (a-b)!}$  /.
(a_)! -> a^a Exp[-a] Sqrt[2 pi a] // Simplify
```

Out[147]=

$$\frac{1}{\sqrt{i} \sqrt{-i+n} \pi}$$

It is convenient to cast the problem into that of determining the probability that a given fraction of intervals of the rw are positive.

$$\mathbb{P}\left[\frac{k}{n} \leq x\right] = \sum_{\substack{i=1 \\ n=n}}^x \frac{1}{\sqrt{i} \sqrt{-i+n} \pi}$$

Let  $v = i/n$ 

$$\mathbb{P}\left[\frac{k}{n} \leq x\right] = \sum_{v=\frac{1}{n}}^x \frac{1}{\sqrt{in} \sqrt{-in+n} \pi} = \sum_{v=\frac{1}{n}}^x \frac{1}{\sqrt{in} \sqrt{-in+n} \pi} = \sum_{v=\frac{1}{n}}^x \frac{(1/n)}{\sqrt{v} \sqrt{1-v} \pi}$$

For  $n \gg 1$  the sum is approximated by an integral over  $v$  with  $dv = 1/n$

$$\mathbb{P}\left[\frac{k}{n} \leq x\right] = \frac{1}{\pi} \int_0^x \frac{1}{\sqrt{v} \sqrt{1-v}} dv$$

Note

In[148]:=  $\text{Integrate}\left[\frac{1}{\pi \sqrt{v(1-v)}}, \{v, 0, x\}, \text{Assumptions} \rightarrow \{x \in \text{Reals}, 0 < x < 1\}\right]$

Out[148]=  $\frac{2 \text{ArcSin}[\sqrt{x}]}{\pi}$

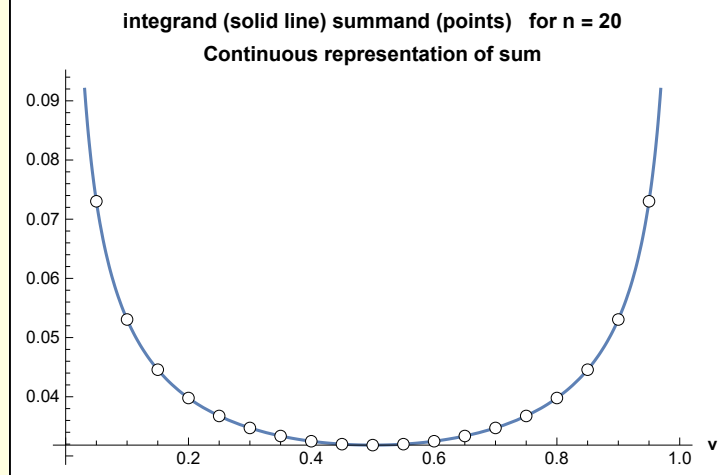
The desired probability for  $n \gg 1$  is

$$\mathbb{P}\left[\frac{k}{n} \leq x\right] = \frac{2}{\pi} \arcsin(\sqrt{x})$$

I compare the summand and integrand for  $n = 20$

In[149]:= `Module[{n = 20, values, points, lab},  
 values = Table[{v,  $\frac{1/n}{\sqrt{v} \sqrt{1-v} \pi}$ }, {v,  $\frac{1}{n}$ ,  $1 - \frac{1}{n}$ ,  $\frac{1}{n}$ }] // N;  
 points = (OC[#, Black]) & /@ values;  
 lab =  
 Stl@StringForm["integrand (solid line) summand (points) for n = ``\nContinuous  
 representation of sum", n];  
 Plot[ $\frac{1/n}{\sqrt{v} \sqrt{1-v} \pi}$ , {v, 0, 1}, AxesLabel -> {Stl["v"], Automatic},  
 PlotLabel -> lab, Epilog -> points]]`

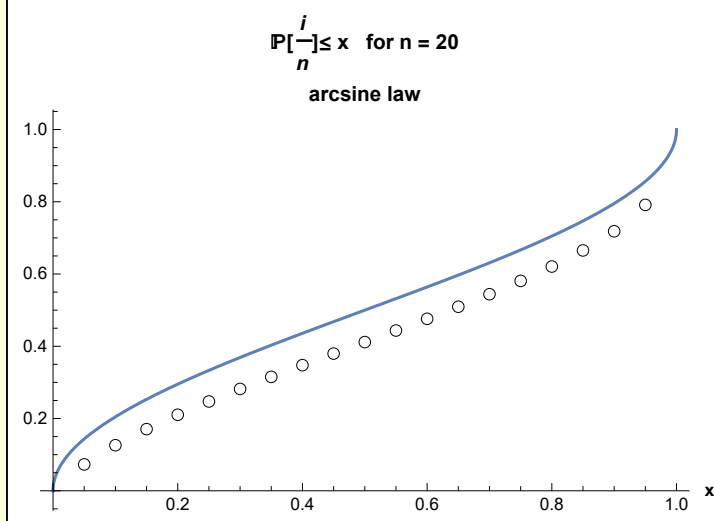
Out[149]=



In[150]=

```
Module[{n = 20, values, points, lab},
  values = Table[{v,  $\frac{1/n}{\sqrt{v} \sqrt{1-v} \pi}$ }, {v,  $\frac{1}{n}$ ,  $1 - \frac{1}{n}$ ,  $\frac{1}{n}$ }] // N;
  values = FoldList[Plus, 0, values][[All, 2]][[2 ;; -1]];
  values = Transpose[{Range[1, n - 1] / n, values}];
  points = (OC[#, Black]) & /@ values;
  lab = Stl@StringForm[" $\mathbb{P}[\frac{i}{n} \leq x$  for n = ``\n arcsine law", n];
  Plot[ $\frac{2 \text{ArcSin}[\sqrt{x}]}{\pi}$ , {x, 0, 1},
  AxesLabel -> {Stl["x"], Automatic}, PlotLabel -> lab, Epilog -> points]]
```

Out[150]=



The approximation is quite good, even for  $n = 20$ .

The fact that the continuous arcsine form is accurate even for modest values led me to forego the the derivation of the arcsine distributions for last returns and maxima using discrete random walks. Actually, they are more straightforward that the derivation of the arcsine distribution for the fraction of time in the lead. Instead I made the connect with trading for Brownian motions and focused on numerical methods.

## Functions

In[84]=

```
Clear[Ballot];
Ballot[n_, b_] :=  $\frac{b}{n} \text{Binomial}[n, \frac{1}{2} (b + n)]$ 
```

Feller's formula

$$L_{2n} = \frac{1}{n+1} \binom{2n}{n}$$

In[86]:=

```

Clear[L, Pu, Pf, Pp];
L[twon_] := Module[{n = twon/2},  $\frac{1}{n+1}$  Binomial[2 n, n]];
Pu[twon_] := Module[{n = twon/2},  $\left(\frac{1}{2}\right)^{2n}$  Binomial[2 n, n]];
Pf[twon_] := Module[{n = twon/2}, If[n ≥ 1,  $\left(\frac{1}{2n}\right)$  Pu[2 n - 2], 0]];
Pp[twok_, twon_] := Module[{k = twok/2, n = twon/2}, Pu[2 k] Pu[2 n - 2 k]];

```

$$Pu_{2n} \equiv \binom{2n}{n} \left(\frac{1}{2}\right)^{2n}$$

$$Pf_{2n} \equiv \frac{1}{2n} Pu_{2n-2} \quad f_0 \equiv 0$$

In[121]:=

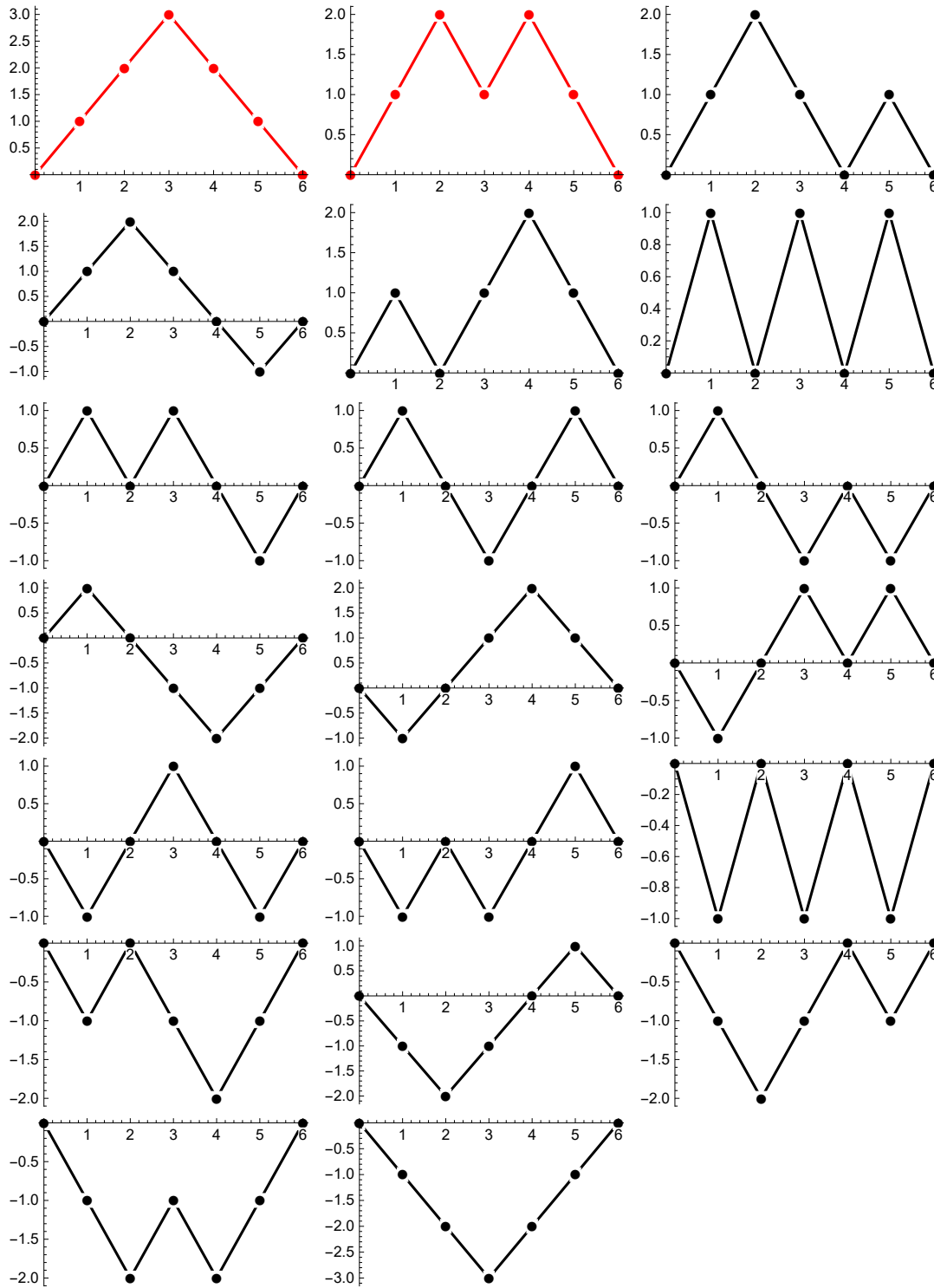
```

Clear[ShowPaths];
ShowPaths[n_, a_, b_, Selector_] :=
Module[
{u, d, steps, permutations, paths, selectedPaths, displayedPaths, plots, lab},
{u, d} =  $\left\{\frac{1}{2}(n + (b - a)), \frac{1}{2}(n - (b - a))\right\}$ ;
steps = Flatten[{Table[1, {u}], Table[-1, {d}]}];
permutations = Permutations[steps];
paths = Map[FoldList[Plus, 0, #] &, permutations];
selectedPaths = Select[paths, Selector];
displayedPaths = Thread[List[Range[0, n], #]] & /@ selectedPaths;
plots =
ListLinePlot[#, PlotMarkers → Automatic, PlotStyle → Black] & /@ displayedPaths;
lab = Stl[StringForm["` random walks n = `", a = `", b = `",
Length[selectedPaths], n, a, b]];
Labeled[Grid[Partition[plots, UpTo[3]]], lab, Top]

```

ShowSelectedPaths[6, 0, 0, (Min[#[[2 ;; -2]]] > 0) &]

random walks n = 6, a = 0, b = 0  
 number selected paths = 2



In[93]:=

```

Clear[ShowSelectedPaths];
ShowSelectedPaths[n_, a_, b_, Selector_] :=
Module[{u, d, steps, permutations,
  paths, selectedPaths, displayedPaths, plots, Color, lab},
  {u, d} = { $\frac{1}{2}(n + (b - a))$ ,  $\frac{1}{2}(n - (b - a))$ };
  steps = Flatten[{Table[1, {u}], Table[-1, {d}]}];
  permutations = Permutations[steps];
  paths = Map[FoldList[Plus, 0, #] &, permutations];
  selectedPaths = Select[paths, Selector];
  displayedPaths = Thread[List[Range[0, n], #] & /@ paths;
  Color[path_] := If[Selector[path][[All, 2]], Red, Black];
  plots = ListLinePlot[#, PlotMarkers -> Automatic, PlotStyle -> Color[#] & /@
    displayedPaths;
  lab = Stl[StringForm["random walks n = ``, a = ``, b =
    ``\nnumber paths = `` number selected paths = ``",
    n, a, b, Length[paths], Length[selectedPaths]]];
  Labeled[Grid[Partition[plots, UpTo[3]]], lab, Top]]

```

In[95]:=

```

Clear[ShowPaths];
ShowPaths[n_, Selector_] :=
Module[{paths, selectedPaths, displayedPaths, plots, Color, lab},
  paths = ConstructPaths[n];
  selectedPaths = Select[paths, Selector];
  displayedPaths = Thread[List[Range[0, n], #] & /@ paths;
  Color[path_] := If[Selector[path][[All, 2]], Red, Black];
  plots = ListLinePlot[#, PlotMarkers -> Automatic, PlotStyle -> Color[#] & /@
    displayedPaths;
  lab = Stl[StringForm["n = `` generates `` random walks\n`` selected paths",
    n, Length[paths], Length[selectedPaths]]];
  Labeled[Grid[Partition[plots, UpTo[3]]], lab, Top]]

```

In[97]:=

```

Clear[ShowSelectedPathsOnly];
ShowSelectedPathsOnly[n_, Selector_] :=
Module[{paths, selectedPaths, displayedPaths, plots, Color, lab},
  paths = ConstructPaths[n];
  selectedPaths = Select[paths, Selector];
  displayedPaths = Thread[List[Range[0, n], #] & /@ selectedPaths;
  Color[path_] := If[Selector[path][[All, 2]], Red, Black];
  plots = ListLinePlot[#, PlotMarkers -> Automatic, PlotStyle -> Color[#] & /@
    displayedPaths;
  lab = Stl[StringForm["n = `` generates `` random walks\n`` selected paths",
    n, Length[paths], Length[selectedPaths]]];
  Labeled[Grid[Partition[plots, UpTo[3]]], lab, Top]]

```



In[99]:=

```
Clear[ConstructPaths];
ConstructPaths[n_] :=
Module[{pathList, ExtendPath},
  ExtendPath[path_, x_] := Append[path, path[[-1]] + x];
  pathList = {{0, 1}, {0, -1}};
  For[i = 2, i ≤ n, i++,
    pathList = Join[ExtendPath[#, 1] & /@ pathList, ExtendPath[#, -1] & /@ pathList];
  pathList]
```