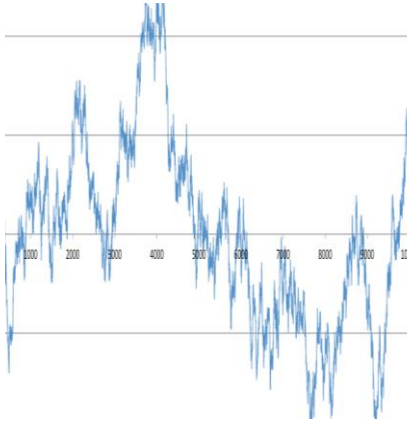
Distribution independent statistics for symmetric random walks: an intuitive proof

Robert Cordery, Fairfield University

Claude Zeller, C. Zeller Consulting LLC

Applications of random walks

- Diffuse light scattering:
  - random media (paper, skin)
  - atmospheres
- Brownian motion
- Critical phenomena
- Games of chance (like the stock market and options)
- Markov Chain Monte Carlo

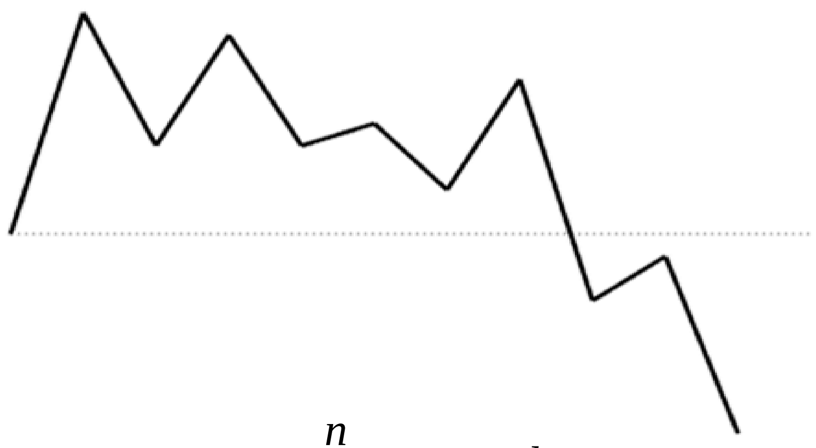


Fluctuation theory: partial sum statistics

- Fluctuation theory addresses walks with independent, identically distributed (IID) random steps
  - Developed 1949 – 1960 by Andersen, Baxter, Darling, Feller, Spitzer, Wendel and others algebraic meth
  - Proved results with a variety of combinatoric and algebraic methods
  - Showed that many statistics are (surprisingly) independent of the step size probability distribution
- We found a simple pairing of walks that yields intuitive proofs of many fluctuation theory results
  - Explains why order statistics and first passage statistics for continuous step distributions look “combinatoric”

Random walks on the reals

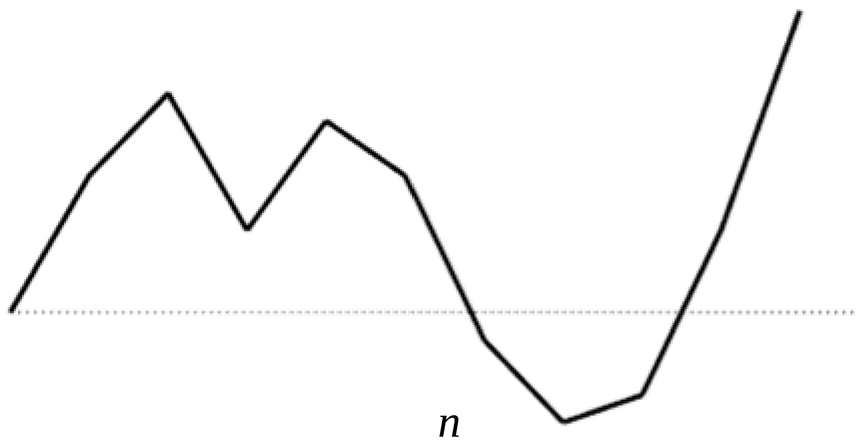
Alternating sign walk



$$z_n = \sum_{k=1}^{n} \left(-1\right)^k s_k$$

Step lengths > 0 from a continuous pdf

Symmetric step distribution

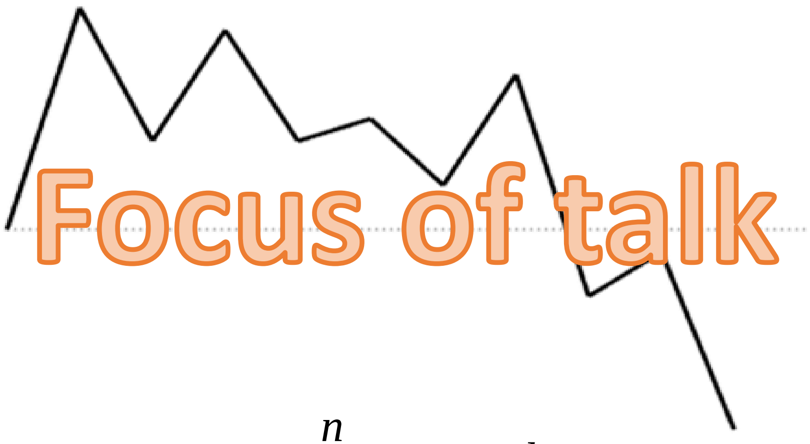


$$z_n = \sum_{k=1}^{n} s_k$$

Step lengths from a symmetric continuous pdf

Random walks on the reals

Alternating sign walk



$$z_n = \sum_{k=1}^{n} (-1)^k s_k$$

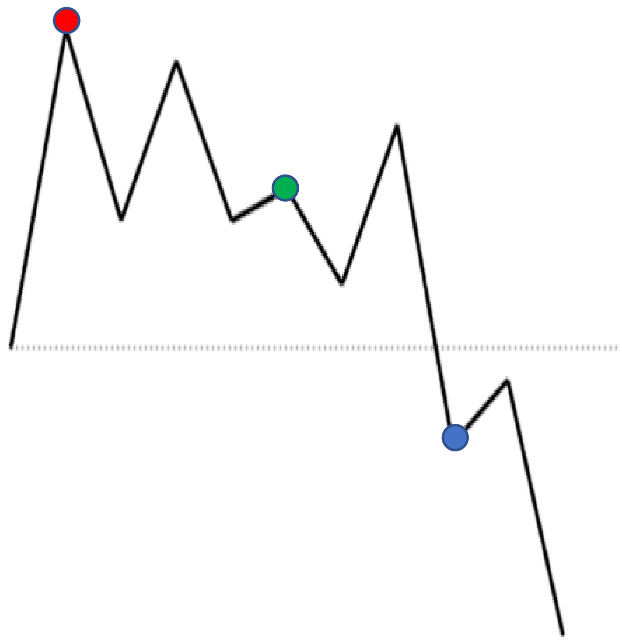
Step lengths > 0 from a continuous pdf

Symmetric step distribution

$$z_n = \sum_{k=1}^{n} s_k$$

Step lengths from a symmetric continuous pdf

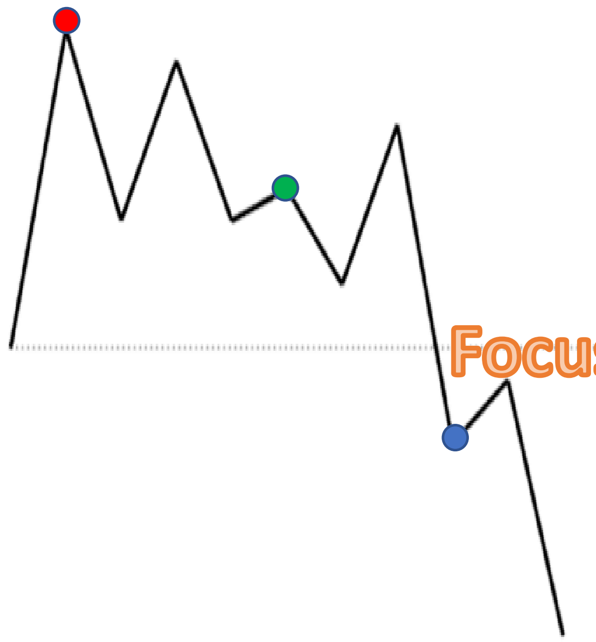
Distribution independent statistics



For an N-step walk with real IID steps (either alternating or random sign)

- Statistics independent of the step size distribution:
    - Step with first passage to a negative value
    - Step with highest peak (lowest valley)
    - Step with k^{th} highest peak (k^{th} lowest valley)
    - Number of positive steps
- But not distribution of the height of highest peak

Distribution independent statistics



For an N-step walk with real IID steps (either alternating or random sign)

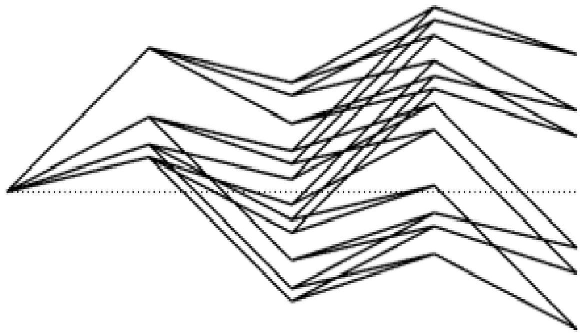
- Statistics independent of the step size distribution:
  Focus of talk
  - Step with first passage to a negative value
  - Step with highest peak (lowest valley)
  - Step with k^{th} highest peak (k^{th} lowest valley)
  - Number of positive steps
- But not distribution of the height of highest peak

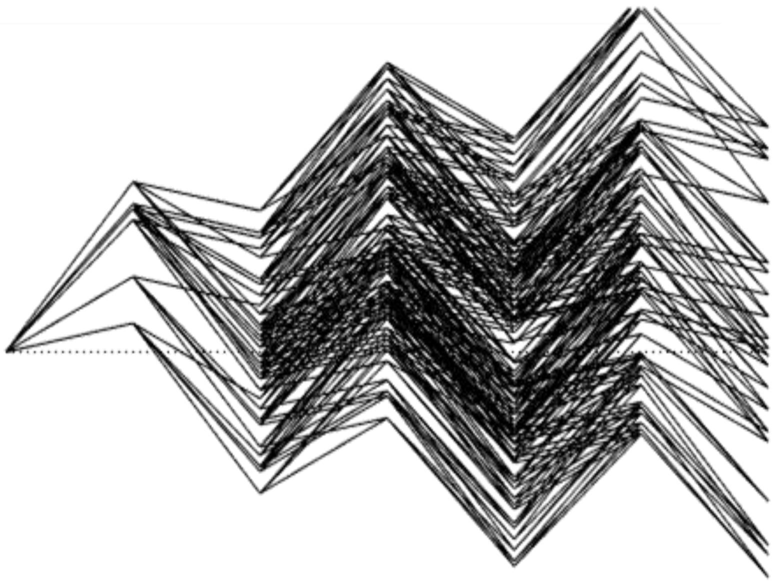
Statistics averaged over permutations of steps are independent of the set of lengths

- IID steps implies the probability of a finite walk is independent of permutations of steps (so called “symmetrically distributed”)
  - Symmetrically distributed is more general than IID,
    - Coordinates near the surface of a hypersphere
- N-step walk → permutations of a finite set of lengths
- For each set of lengths A_N, we average over permutations
  - Our main result is that averaging these statistics over permutations gives a result independent of A_N
  - Except for a constraint on the lengths of measure 0
  - Averaging over sets of lengths is thus average of a constant

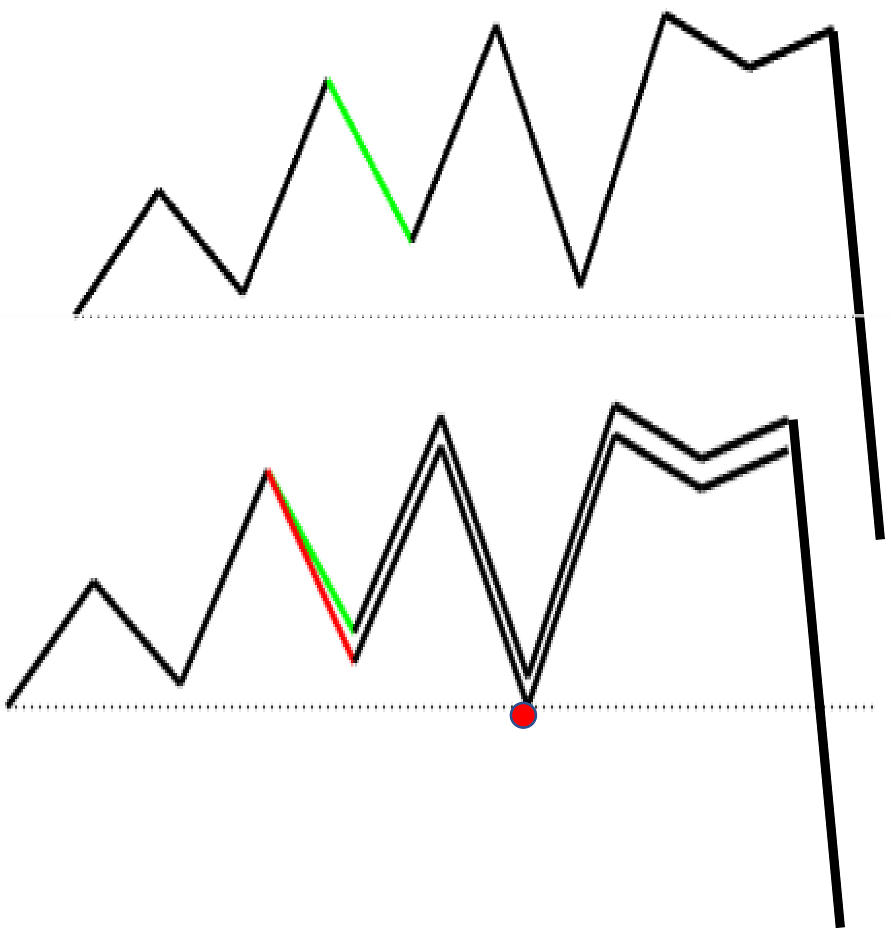
Constructing walks from permutations

- Given a fixed set A_N of N steps with real lengths
- Construct the set W_N of all alternating (zigzag) walks from permutations of these steps
  - There are $N!$ such walks
  - There are $2^N N!$ For symmetrically distributed step length
- Constrain the lengths so that no sub-walk with an even number of steps starts and ends at 0
  - This is a constraint of measure 0
- Claim: the fraction of walks that have first passage to a negative value on step $k \leq N$ is independent of A_N





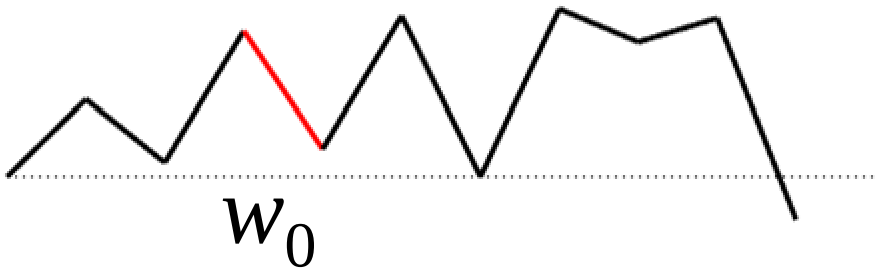
Changing one set of lengths into another



- We change step lengths one at a time
- Small length changes do not affect first passage statistics
- At a critical length some walk returns exactly to 0
- Past that length, the first passage for this walk changes
- These are the only walks that could change first passage statistics
- However …

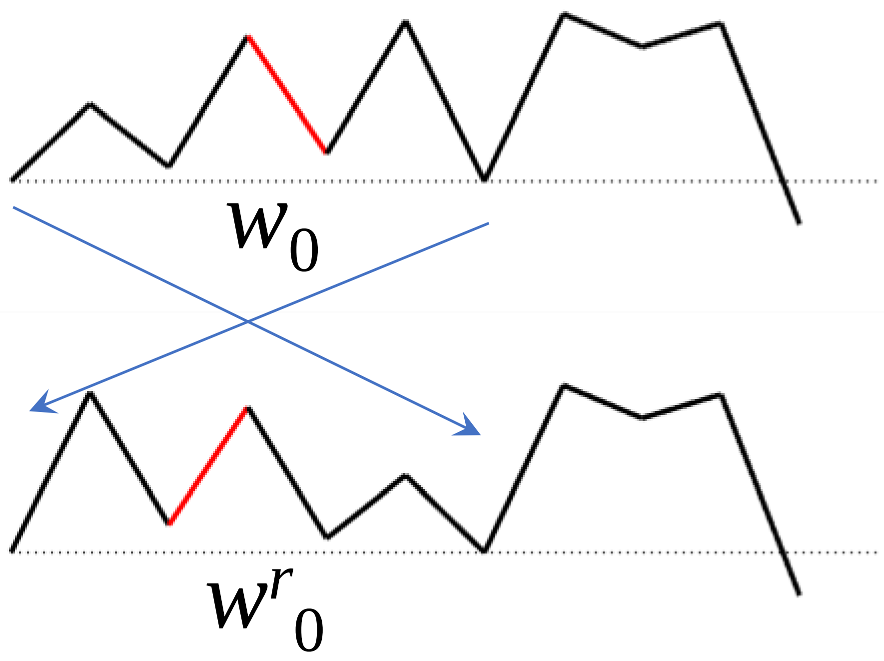
Pairing walks around a critical value

- When a length change causes one even sub-walk w_0 to return to 0, many others also return to 0:
  - Permutations of the upward steps
  - Permutations of the downward steps
  - Exchanging the upward and downward steps



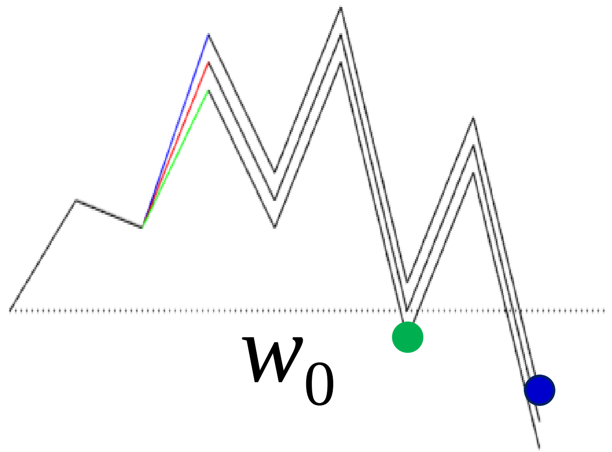
Pairing walks around a critical value

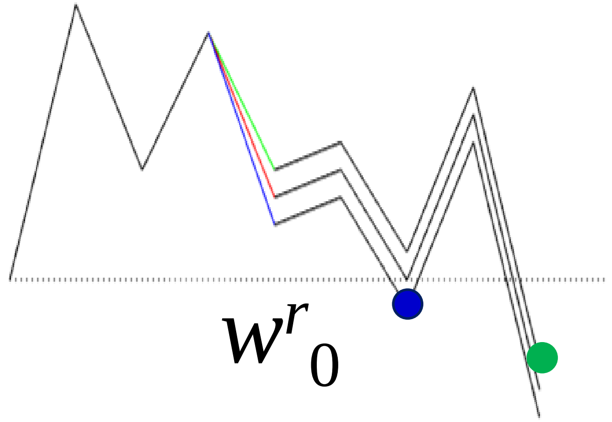
- When a length change causes one even sub-walk w_0 to return to 0, many others do also:
  - Permutation of the upward steps
  - Permutation of the downward steps
  - Exchanging the upward and downward steps
- In particular, the time reversed sub-walk $w^r{}_0$ goes from 0 to 0
- Pair each critical walk that begins with w_0 with a walk with the same tail but begins with $w^r{}_0$



Pairing walks around a critical value

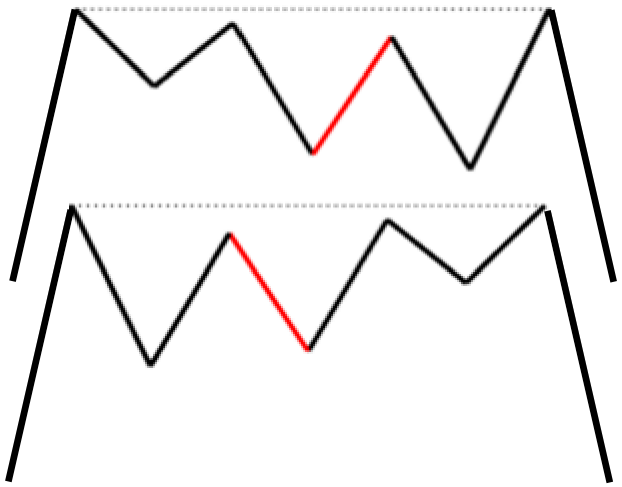
- As a step is changed through a critical value
  - If w_0 ends higher than it's start, then the time reversed walk $w^r{}_0$ ends lower than it's start
- First passages of the paired walks beginning with w_0 and $w^r{}_0$ exchange step number
- The total number of first passages at each step is unchanged





Pairing walks works for other statistics

- Order statistics (location of k^{th} highest peak or valley) are distribution independent
- Location of k^{th} passage
- For symmetrically distributed step size average over $2^N N!$ permutations and step sign choice



Distribution independent statistics are combinatoric

- We can use any set of lengths satisfying the constraint
- A good choice is $\{2^k | k = 1 \cdots N\}$
- The steps are positive as long as every negative step is preceded by a longer positive step
- First passage at the k^{th} step requires the k^{th} step be longer that all previous steps and negative
- The fraction with first passage at step k is $\frac{C_{k-1}}{2^{2k-1}}$, where $C_m = \frac{1}{m+1}\binom{2m}{m}$ is the m^{th} Catalan number