Universal Record Statistics of Random Walks

GGI Workshop in Advances in Non-Equilibrium Statistical Mechanics

Grégory Schehr, LPTMS (Orsay)

Universal Record Statistics of Random Walks

GGI Workshop in Advances in Non-Equilibrium Statistical Mechanics

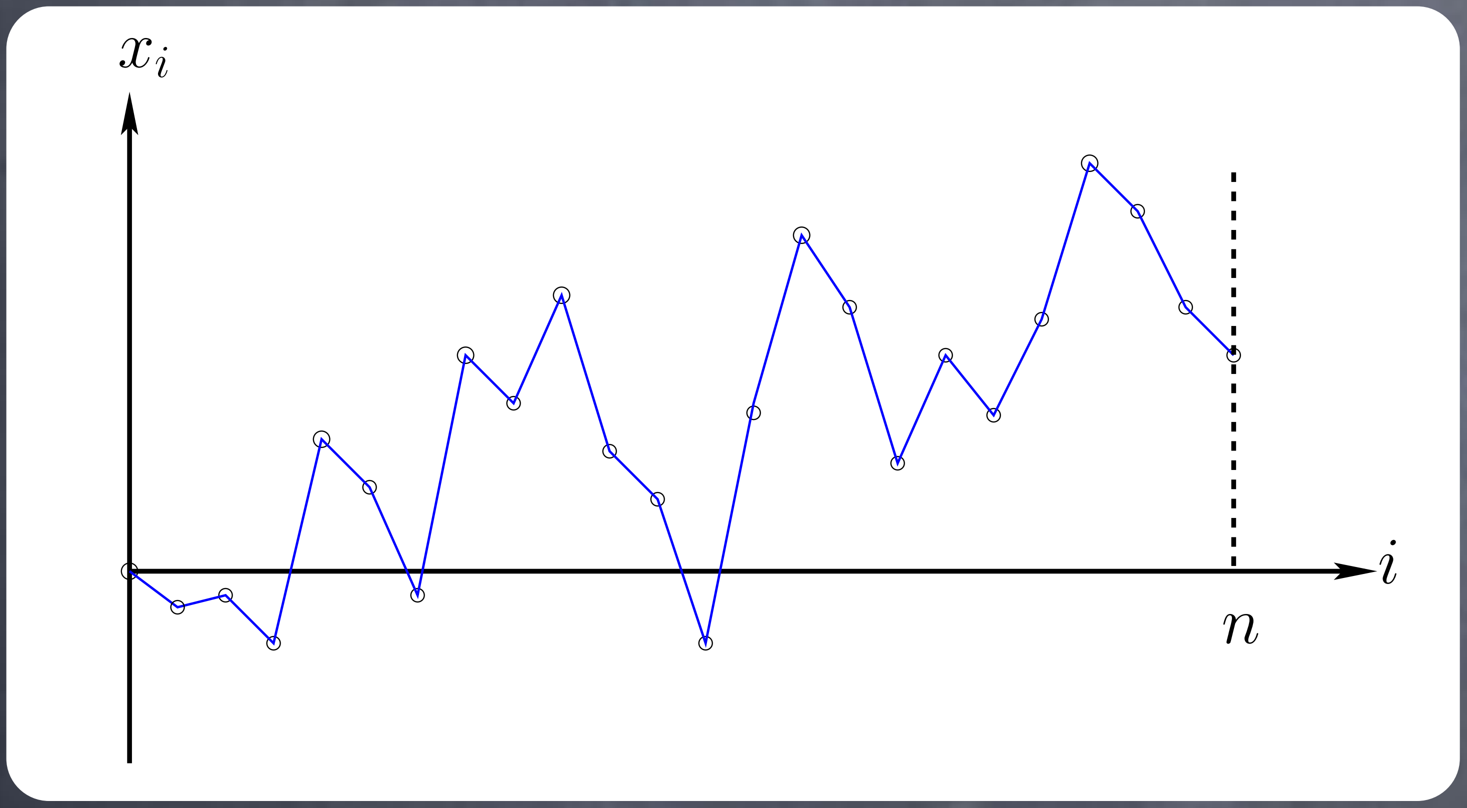
Grégory Schehr, LPTMS (Orsay)

Collaborators:

- C. Godrèche (IPhT, Saclay)
- S. N. Majumdar (LPTMS, Orsay)
- G. Wergen (Uni. of Cologne)

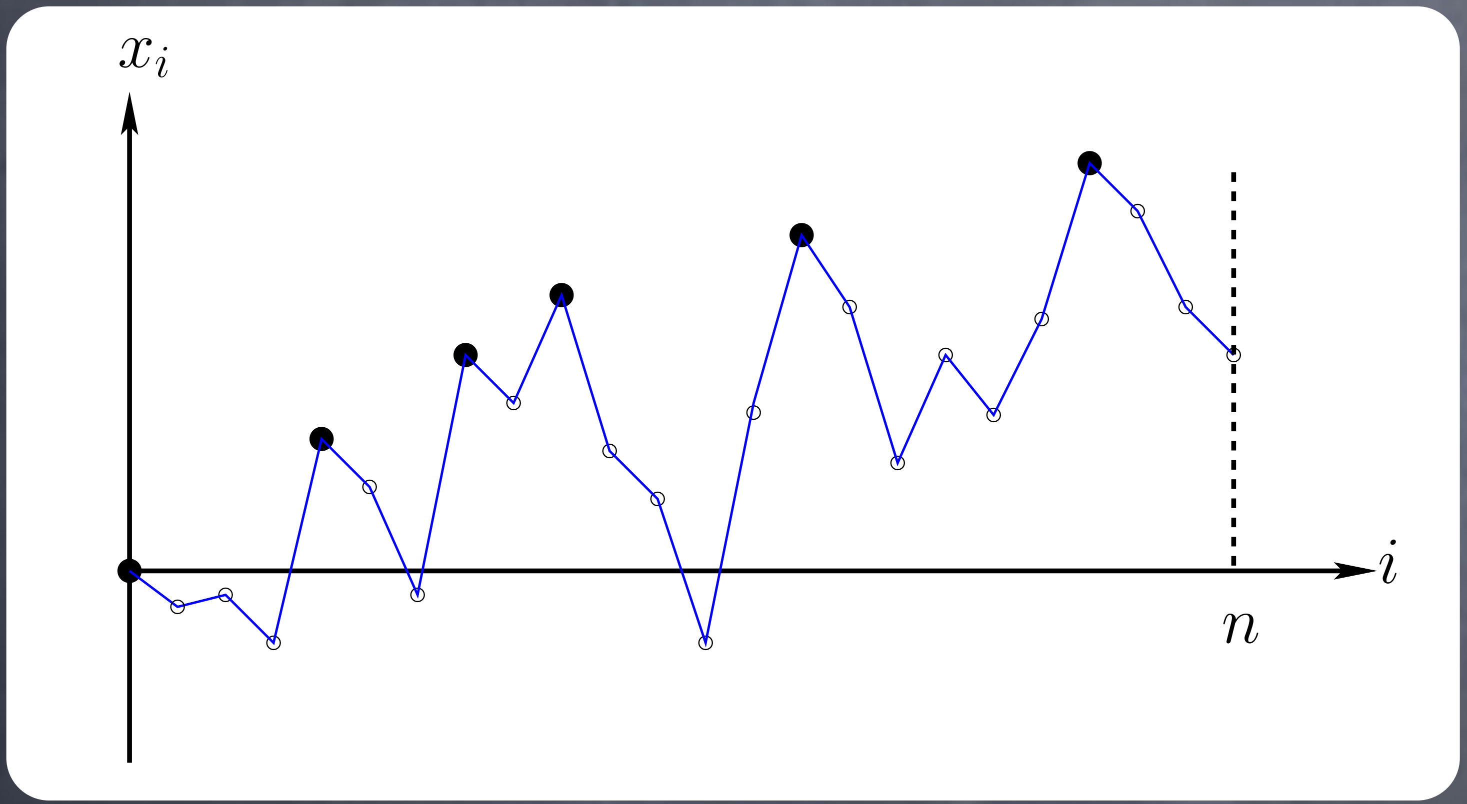
Statement of the problem

$x_1, x_2, \cdots, x_n : n$ random variables (e.g. time series)



Statement of the problem

$x_1, x_2, \cdots, x_n : n$ random variables (e.g. time series)

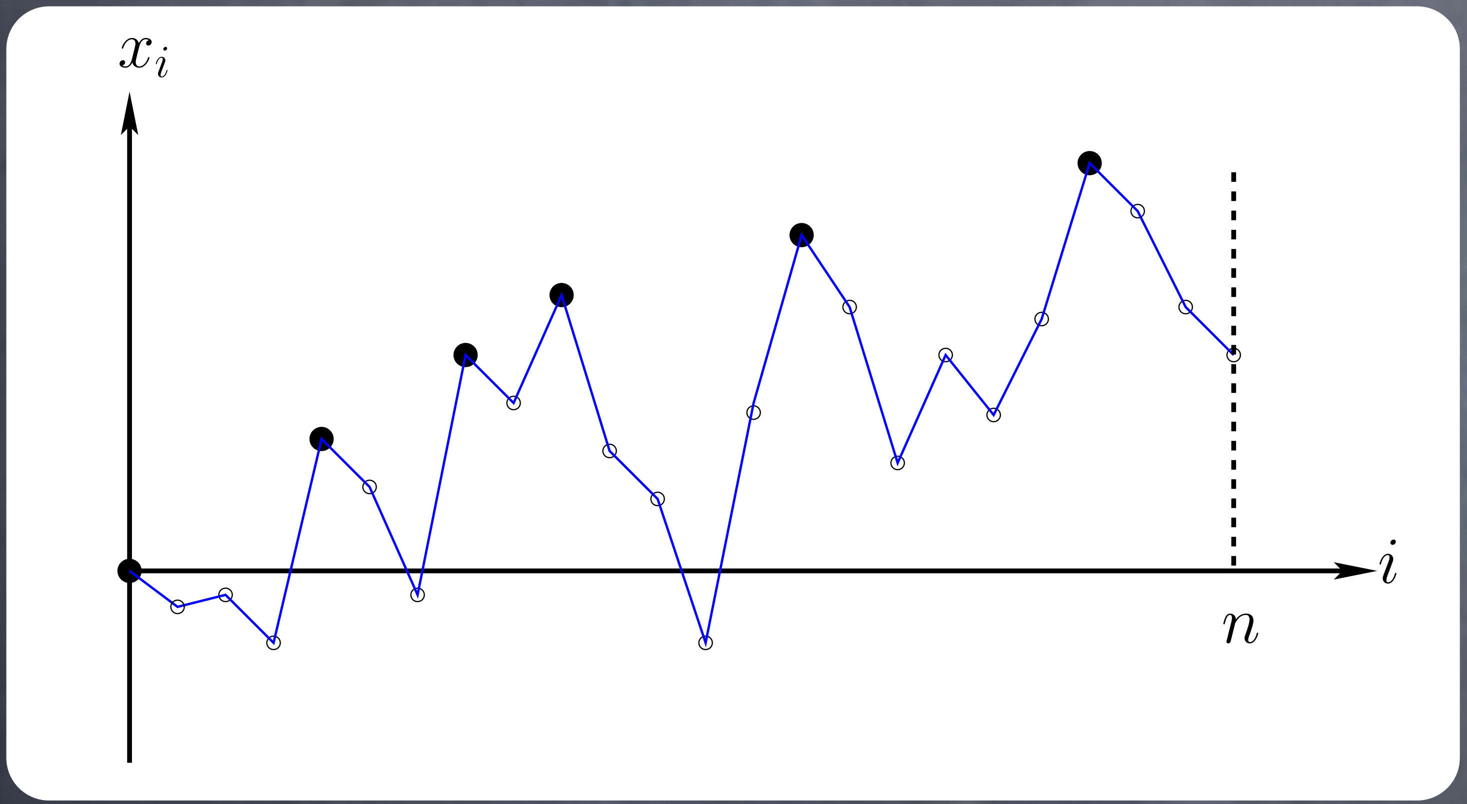


x_k is a record iff

$$x_k \geq \max(x_1, x_2, \cdots, x_{k-1})$$

Statement of the problem

$x_1, x_2, \cdots, x_n : n$ random variables (e.g. time series)



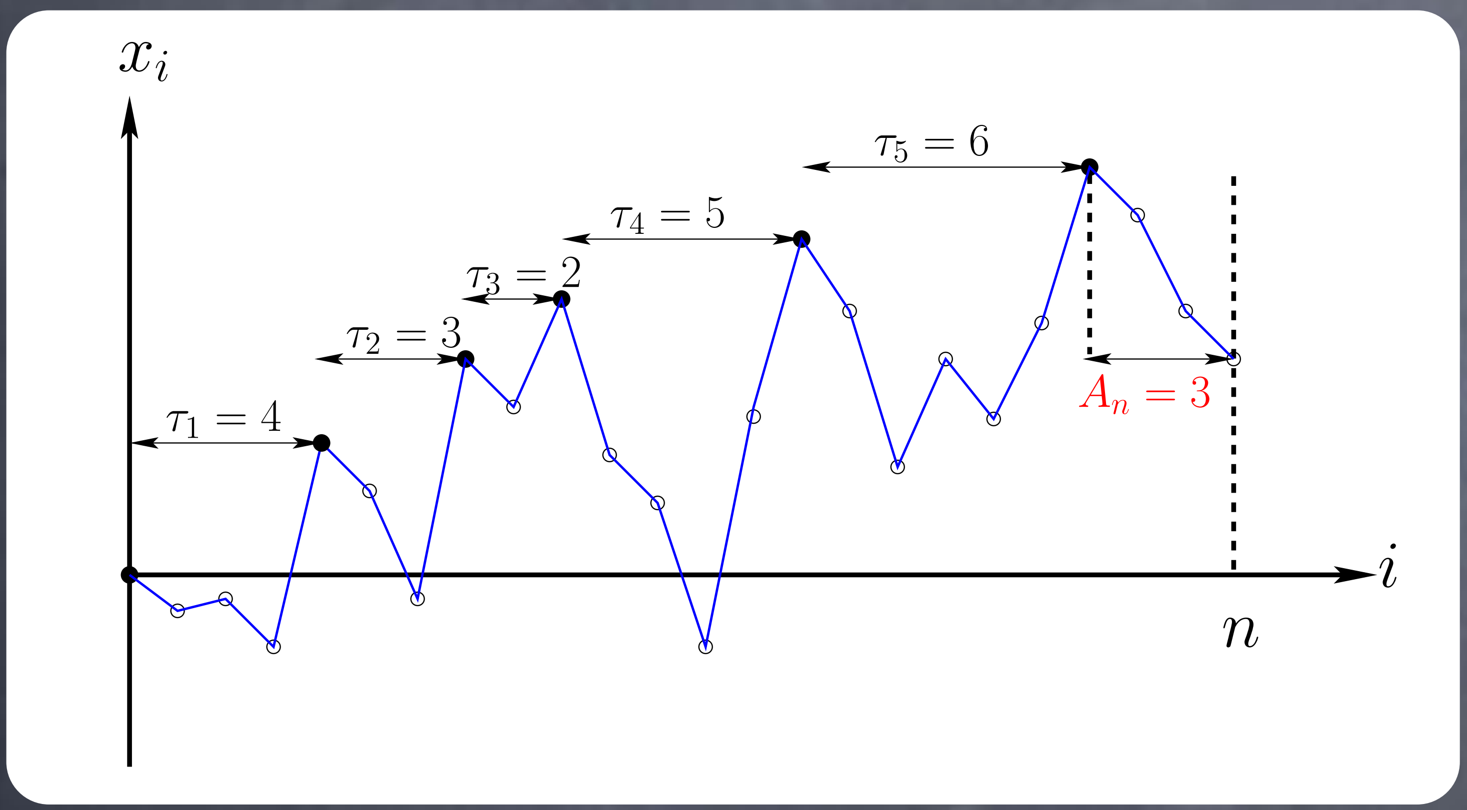
x_k is a record iff

$$x_k \geq \max(x_1, x_2, \cdots, x_{k-1})$$

Questions: Statistics of the number of records R_n ?

Statement of the problem

$x_1, x_2, \cdots, x_n$: n random variables (e.g. time series)



x_k is a record iff

$$x_k \geq \max(x_1, x_2, \cdots, x_{k-1})$$

Questions:
- Statistics of the number of records R_n ?
- Statistics of the ages of records $\tau_1, \tau_2, \cdots, A_n$?

Some recent applications of records in physics

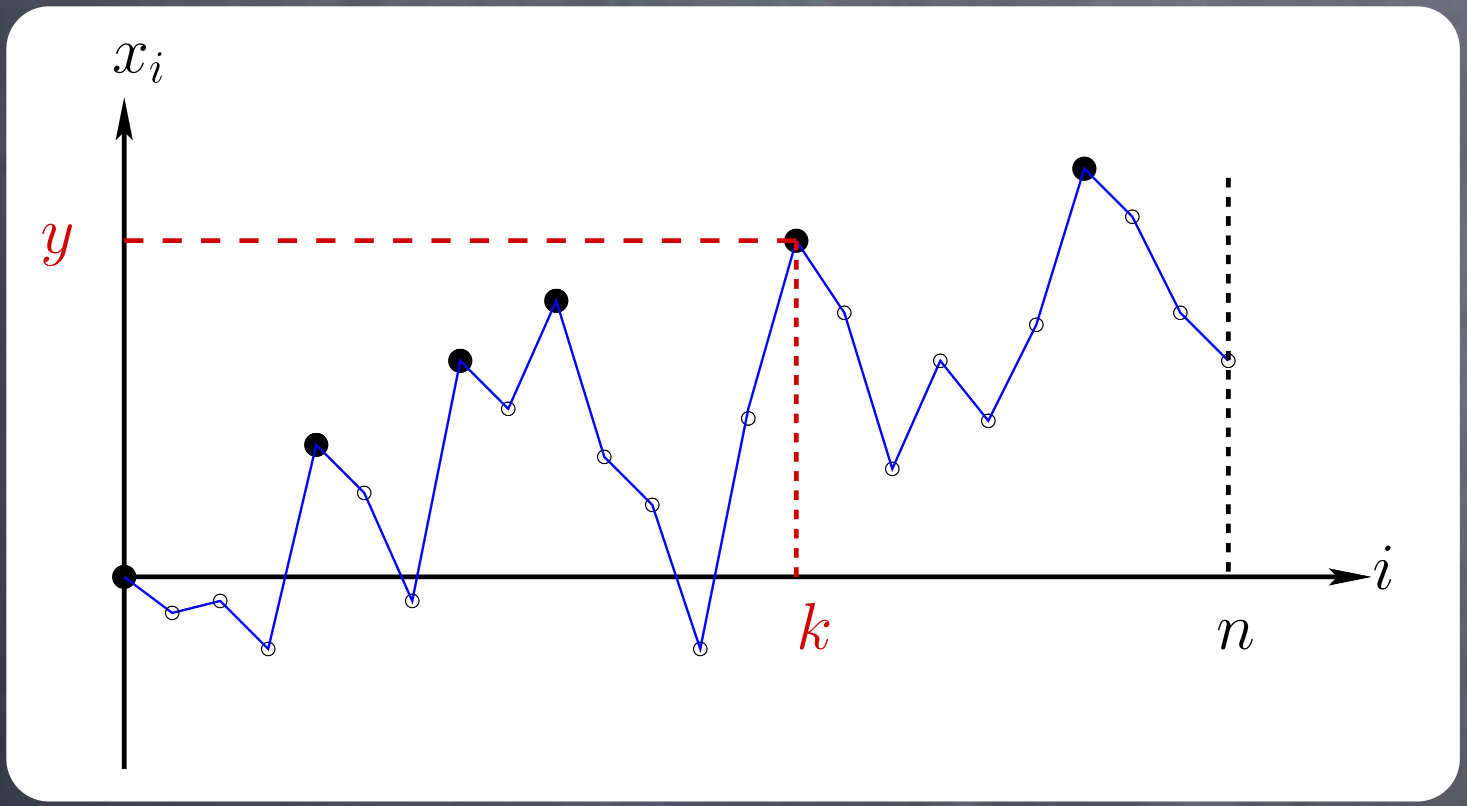
- Domain wall dynamics   Alessandro et al. '90
- Evolutionary biology   Jain & Krug '05
- Global warming   Redner & Petersen '06, Wergen & Krug '10
- Spin-glasses   Sibani '07
- Random walks   Majumdar & Ziff '08, Wergen, Majumdar, G. S. '12
- Growing networks   Godrèche & Luck '08
- Avalanches   Le Doussal & Wiese '09
- Financial data   Wergen, Bogner & Krug '11, Wergen '13

Some recent applications of records in physics

- Domain wall dynamics Alessandro et al. ‘90
- Evolutionary biology Jain & Krug ‘05
- Global warming Redner & Petersen ‘06, Wergen & Krug ‘10
- Spin-glasses Sibani ‘07
- Random walks Majumdar & Ziff ‘08, Wergen, Majumdar, G. S. ‘12
- Growing networks Godrèche & Luck ‘08
- Avalanches Le Doussal & Wiese ‘09
- Financial data Wergen, Bogner & Krug ‘11, Wergen ‘13

Record statistics of i.i.d. random variables

$x_1, x_2, \cdots, x_n : n$ i.i.d. random variables with PDF $p(x)$



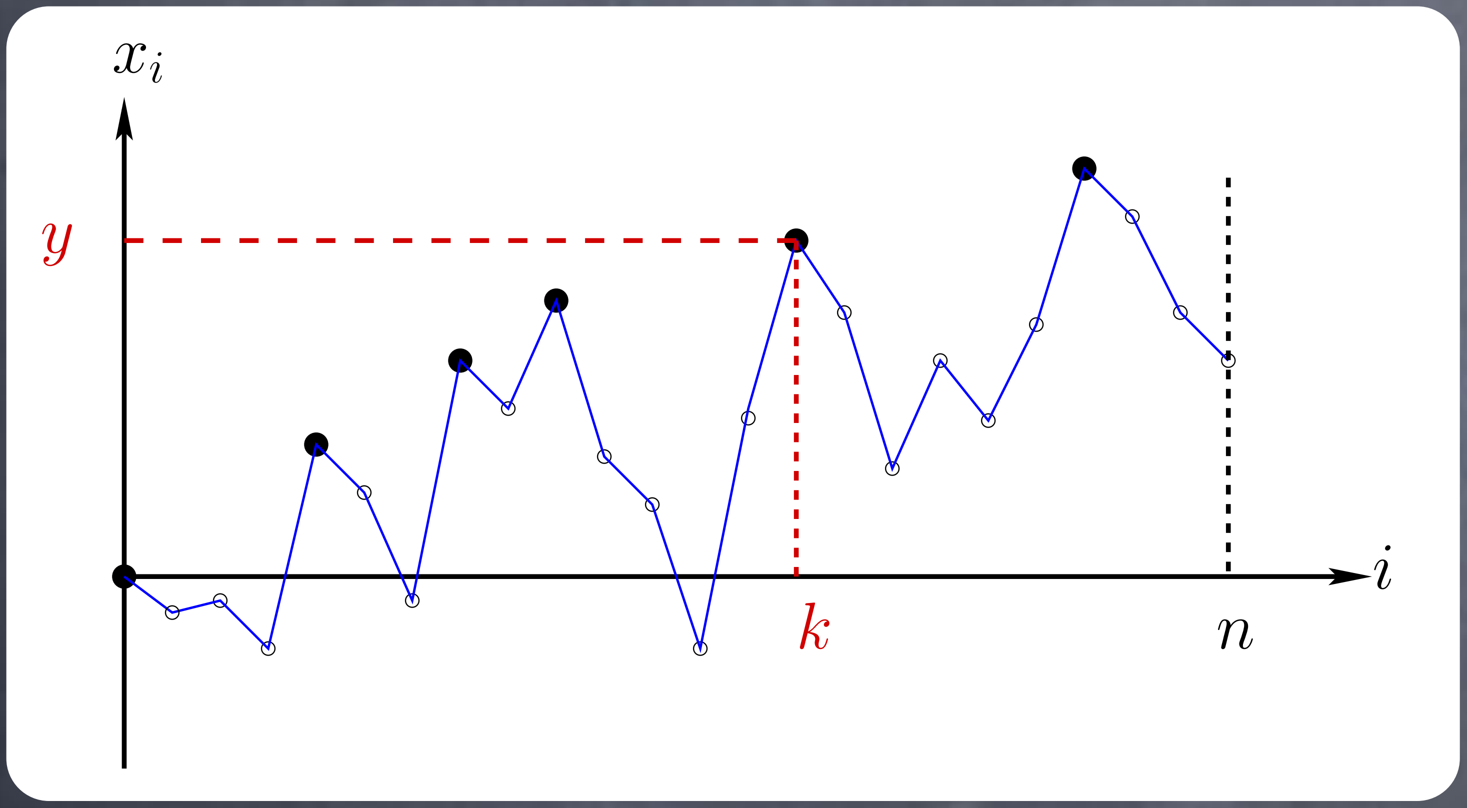
Nber of records R_n

$$R_n = \sum_{k=1}^{n} \sigma_k$$

$$\sigma_k = \begin{cases} 1\,, & \text{if } x_k \text{ is a record} \\ 0\,, & \text{if } x_k \text{ is NOT a record} \end{cases}$$

Record statistics of i.i.d. random variables

$x_1, x_2, \cdots, x_n : n$ i.i.d. random variables with PDF $p(x)$



Nber of records R_n

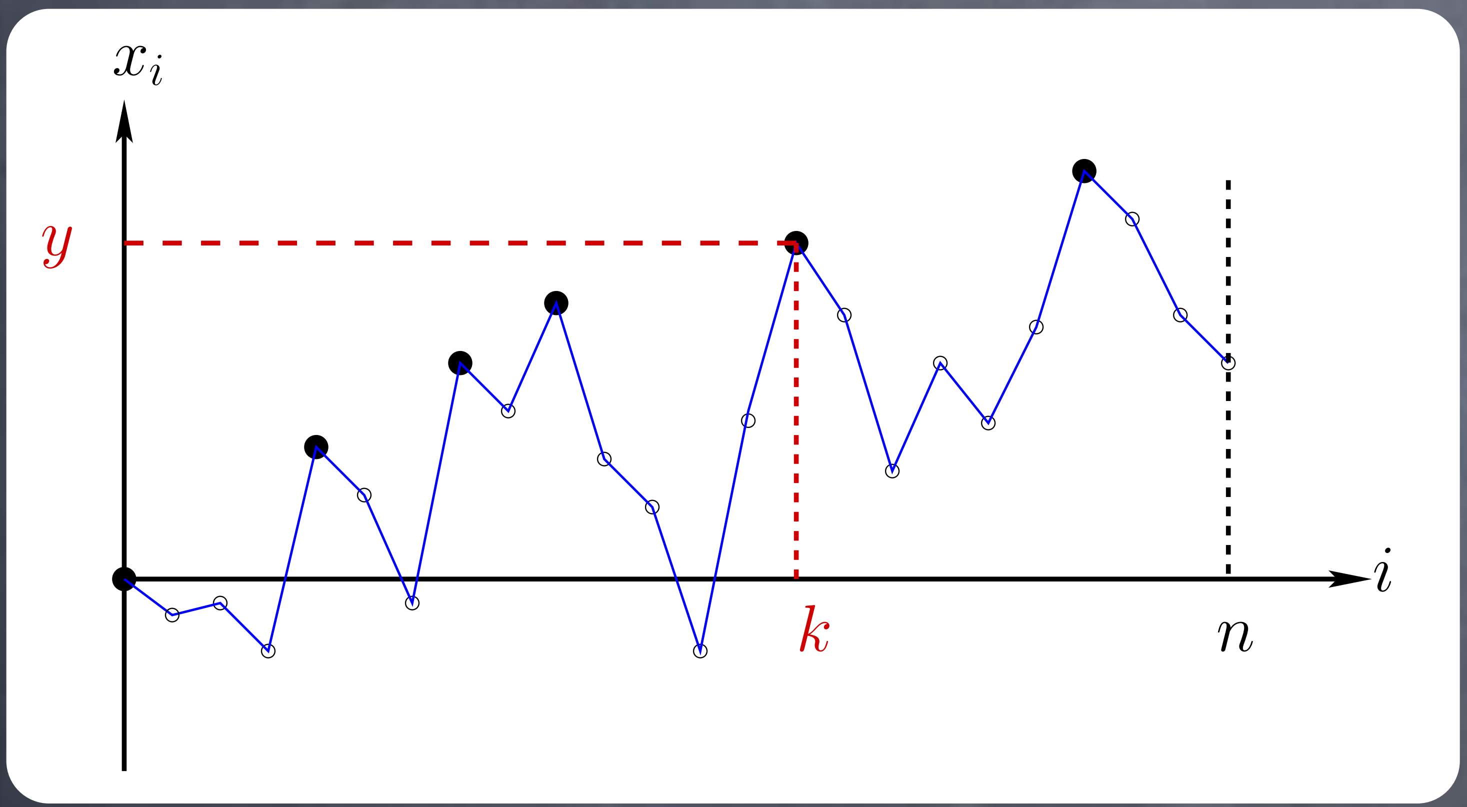
$$R_n = \sum_{k=1}^{n} \sigma_k$$

$$\sigma_k = \begin{cases} 1\ , & \text{if } x_k \text{ is a record} \\ 0\ , & \text{if } x_k \text{ is NOT a record} \end{cases}$$

$\langle R_n \rangle = \sum_{k=1}^{n} r_k \ , \quad r_k = \langle \sigma_k \rangle$ is the proba. that a record is broken at step k

Record statistics of i.i.d. random variables

$x_1, x_2, \cdots, x_n : n$ i.i.d. random variables with PDF $p(x)$



Nber of records R_n

$$R_n = \sum_{k=1}^{n} \sigma_k$$

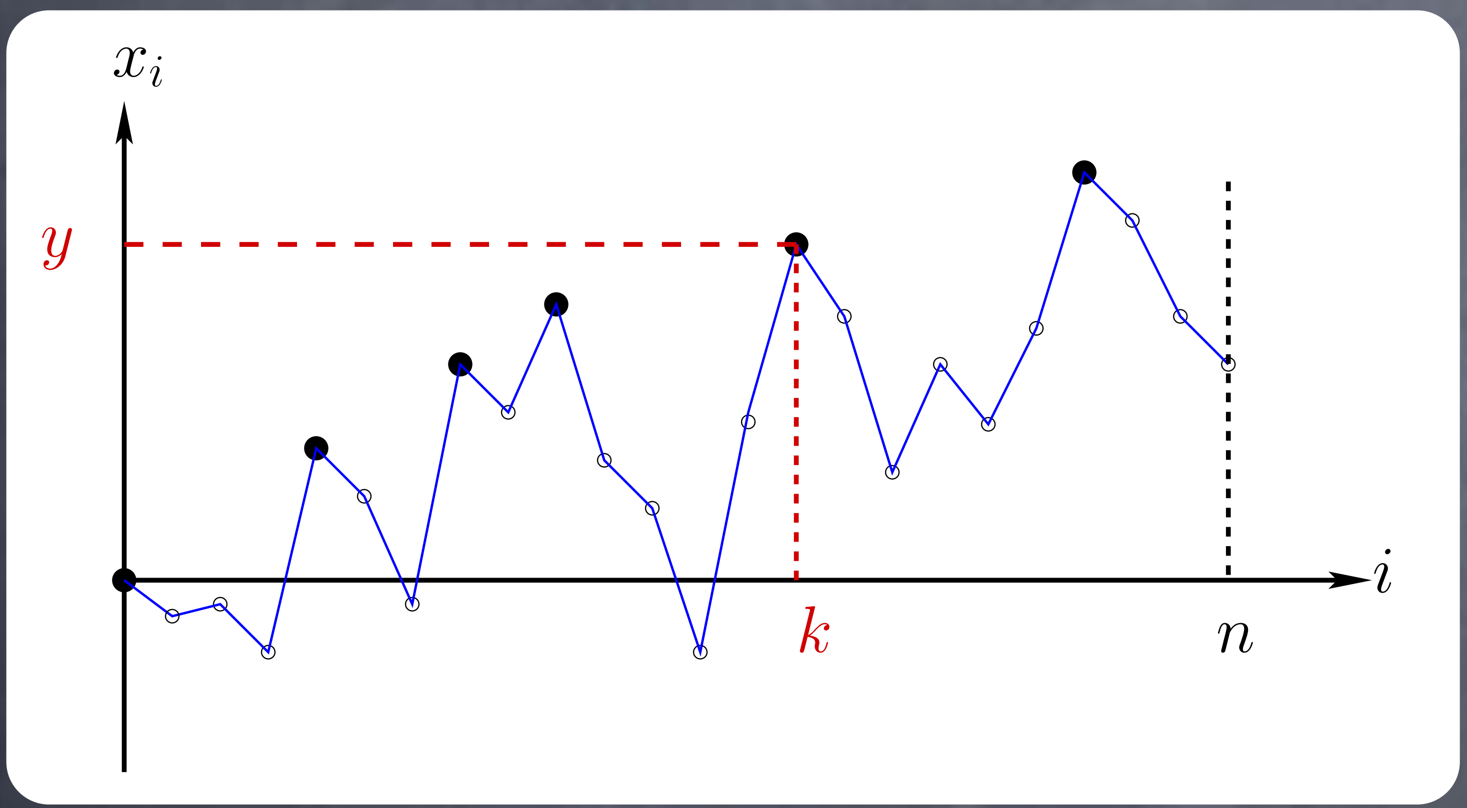
$$\sigma_k = \begin{cases} 1 \, , & \text{if } x_k \text{ is a record} \\ 0 \, , & \text{if } x_k \text{ is NOT a record} \end{cases}$$

$\langle R_n \rangle = \sum_{k=1}^{n} r_k \; , \quad r_k = \langle \sigma_k \rangle$ is the proba. that a record is broken at step k

$$r_k = \int_{-\infty}^{\infty} p(y) \left[\int_{-\infty}^{y} p(x) dx \right]^{k-1} dy$$

Record statistics of i.i.d. random variables

$x_1, x_2, \cdots, x_n : n$ i.i.d. random variables with PDF $p(x)$



Nber of records R_n

$$R_n = \sum_{k=1}^{n} \sigma_k$$

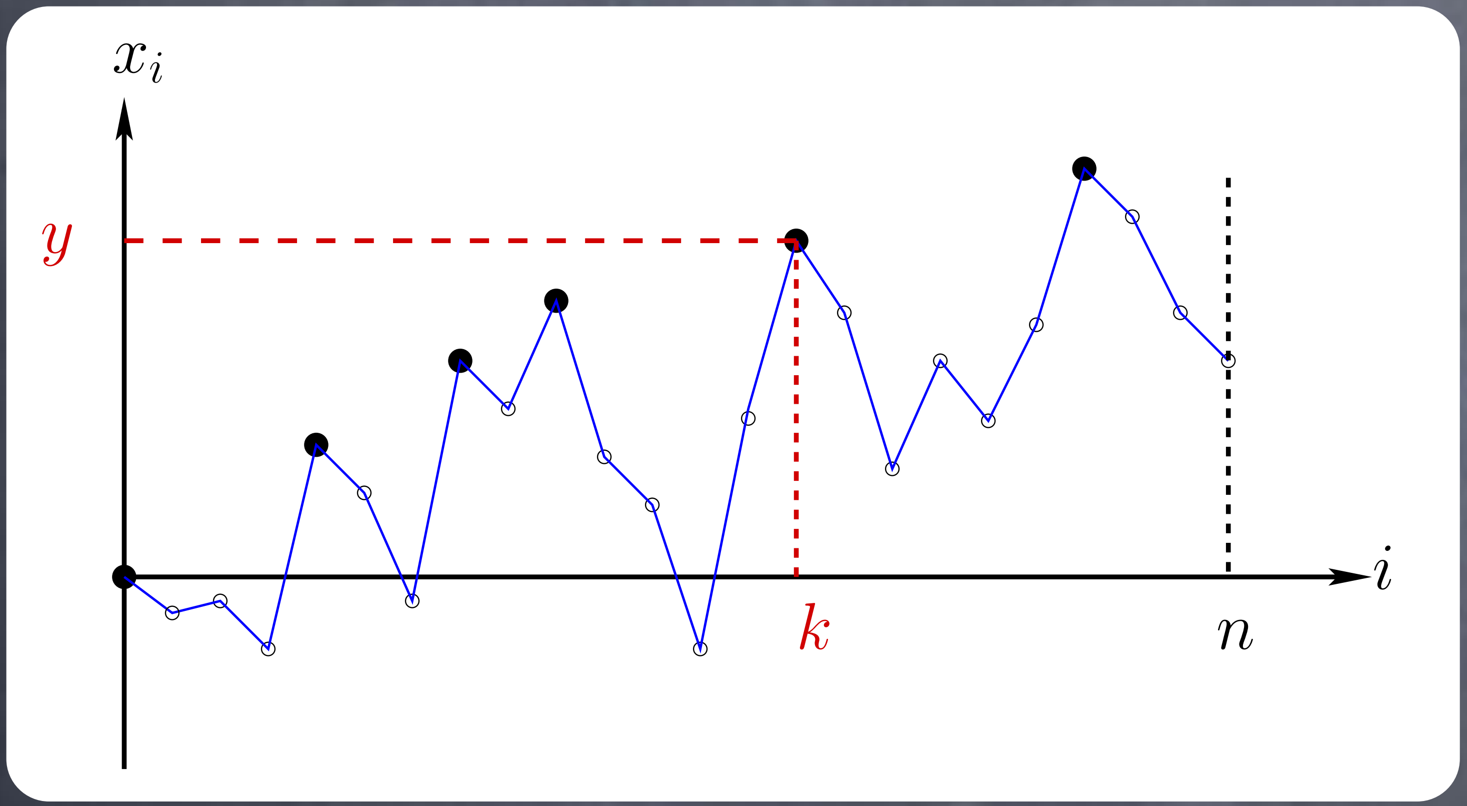
$$\sigma_k = \begin{cases} 1\,, & \text{if } x_k \text{ is a record} \\ 0\,, & \text{if } x_k \text{ is NOT a record} \end{cases}$$

$\langle R_n \rangle = \sum_{k=1}^{n} r_k \,, \quad r_k = \langle \sigma_k \rangle$ is the proba. that a record is broken at step k

$$r_k = \int_{-\infty}^{\infty} p(y) \left[\int_{-\infty}^{y} p(x) dx \right]^{k-1} dy = \frac{1}{k}$$

Record statistics of i.i.d. random variables

$x_1, x_2, \cdots, x_n : n$ i.i.d. random variables with PDF $p(x)$



Nber of records R_n

$$R_n = \sum_{k=1}^{n} \sigma_k$$

$$\sigma_k = \begin{cases} 1\,, & \text{if } x_k \text{ is a record} \\ 0\,, & \text{if } x_k \text{ is NOT a record} \end{cases}$$

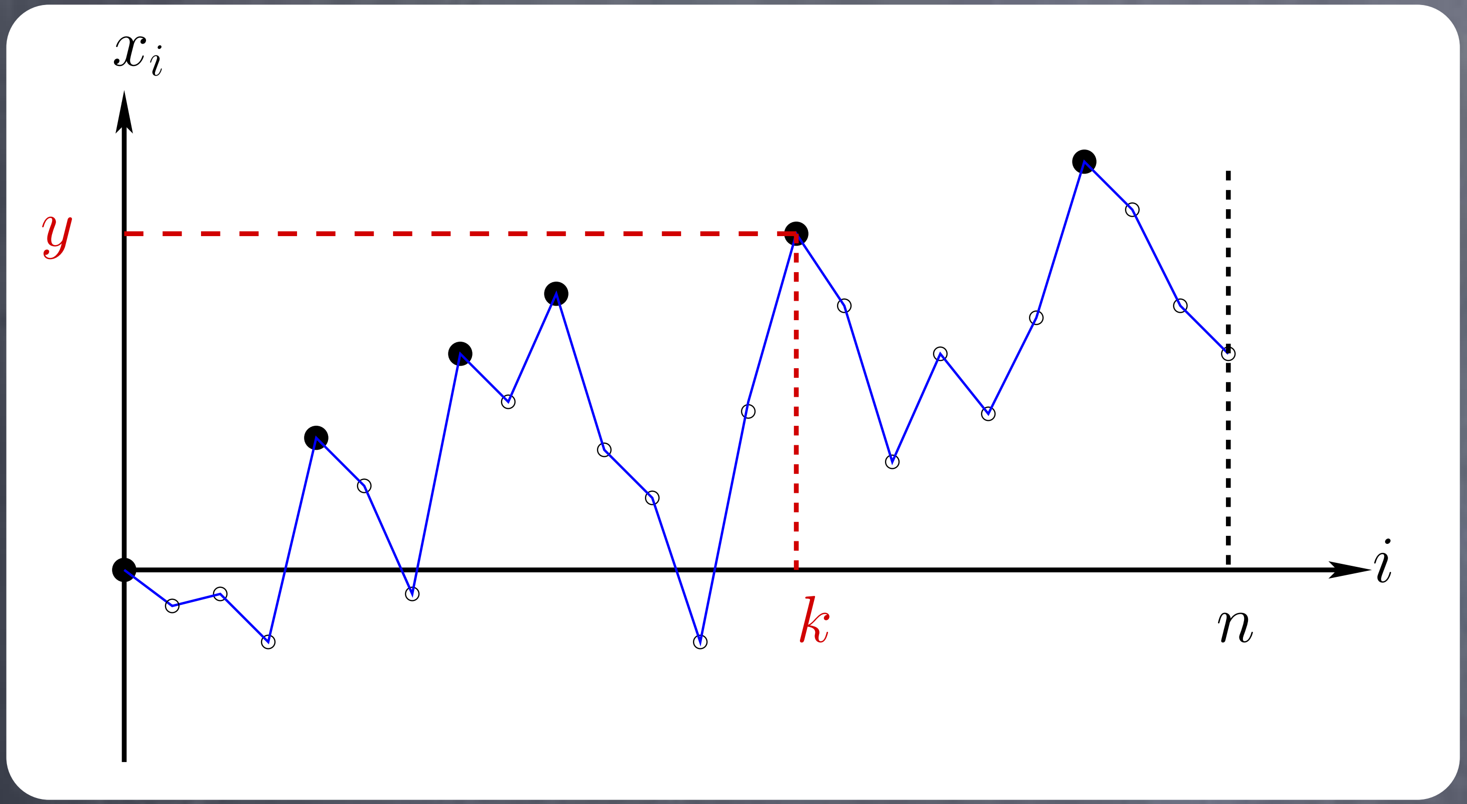
$\langle R_n \rangle = \sum_{k=1}^{n} r_k \;, \quad r_k = \langle \sigma_k \rangle$ is the proba. that a record is broken at step k

Universal !

$$r_k = \int_{-\infty}^{\infty} p(y) \left[\int_{-\infty}^{y} p(x)dx \right]^{k-1} dy = \frac{1}{k}$$

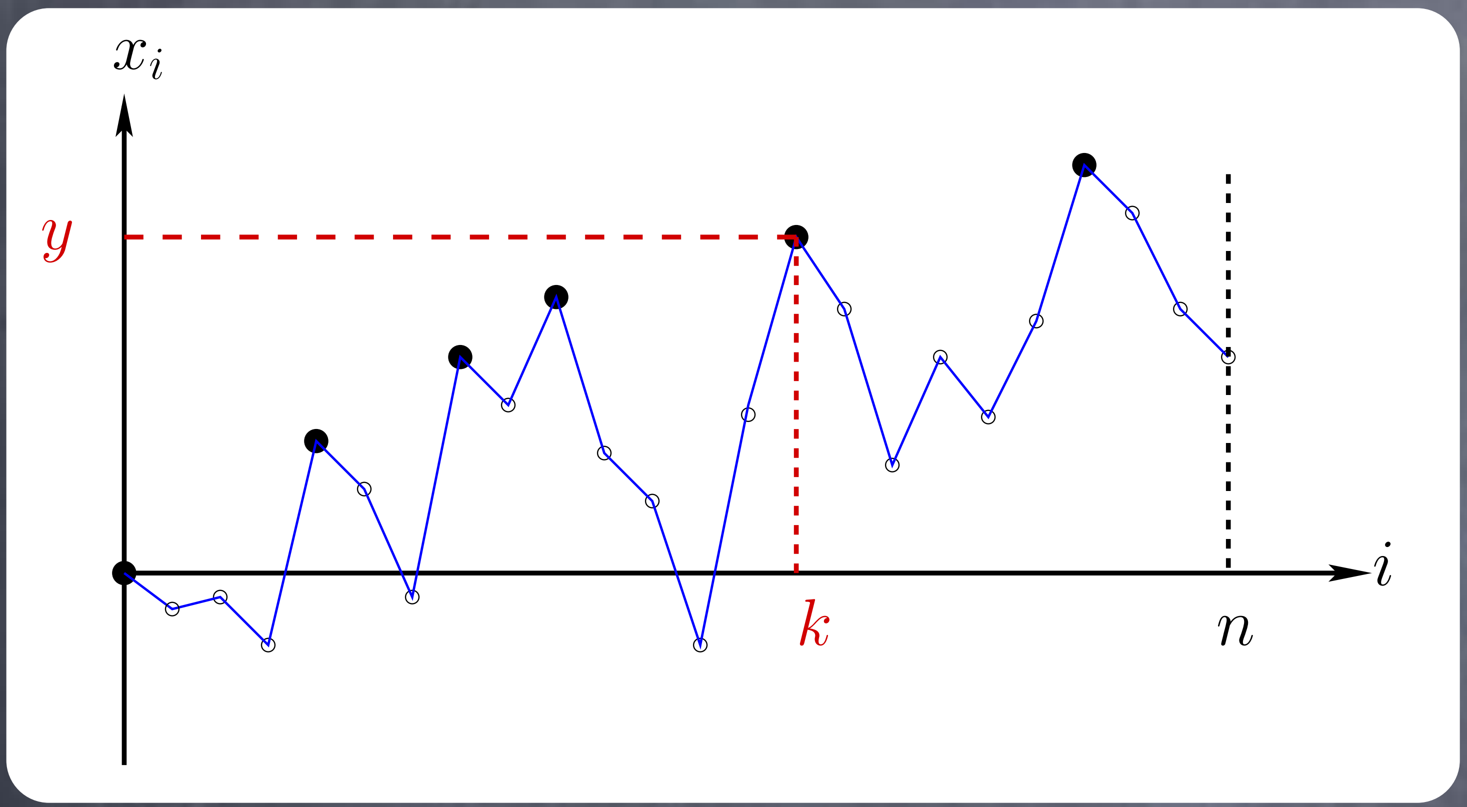
Record statistics of i.i.d. random variables

$x_1, x_2, \cdots, x_N$: N i.i.d. random variables with PDF



Record statistics of i.i.d. random variables

$x_1, x_2, \cdots, x_N : N$ i.i.d. random variables with PDF

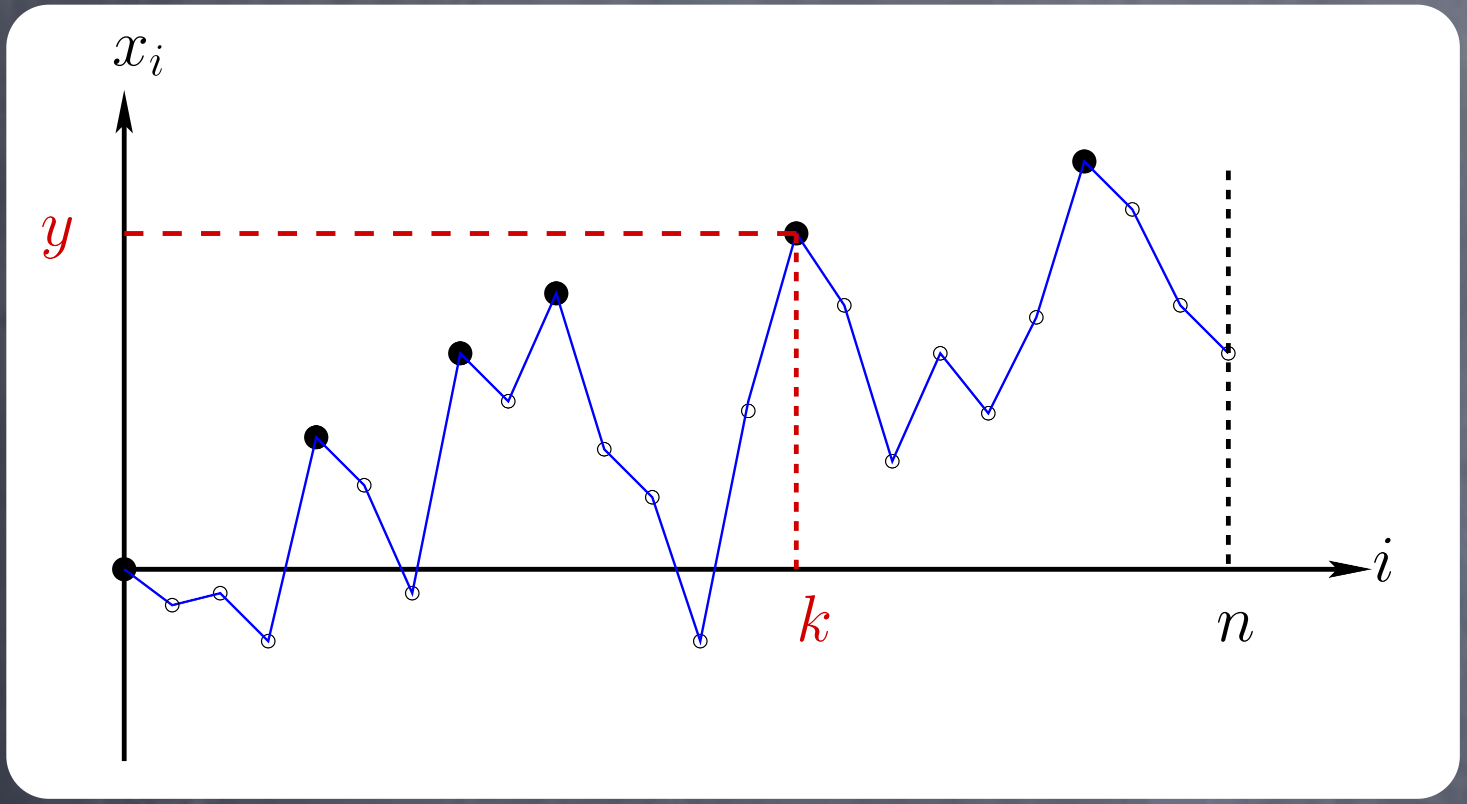


Average nber of records

$$\langle R_n \rangle = \sum_{k=1}^{n} \frac{1}{k} \sim \log n$$

Record statistics of i.i.d. random variables

$x_1, x_2, \cdots, x_N : N$ i.i.d. random variables with PDF



Average nber of records

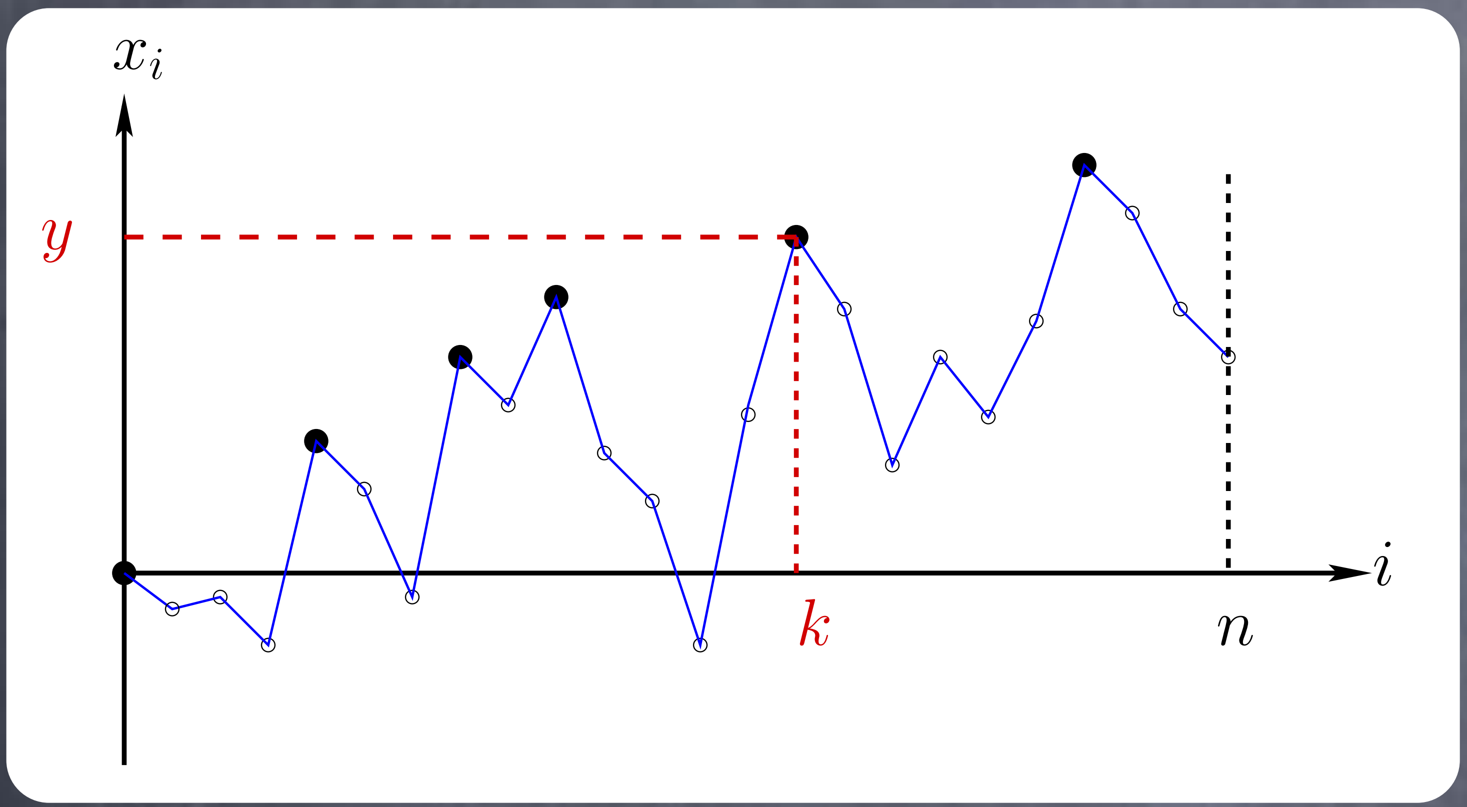
$$\langle R_n \rangle = \sum_{k=1}^{n} \frac{1}{k} \sim \log n$$

Variance

$$\langle R_n^2 \rangle - \langle R_n \rangle^2 \sim \log n$$

Record statistics of i.i.d. random variables

$x_1, x_2, \cdots, x_N$: N i.i.d. random variables with PDF



Average nber of records

$$\langle R_n \rangle = \sum_{k=1}^{n} \frac{1}{k} \sim \log n$$

Variance

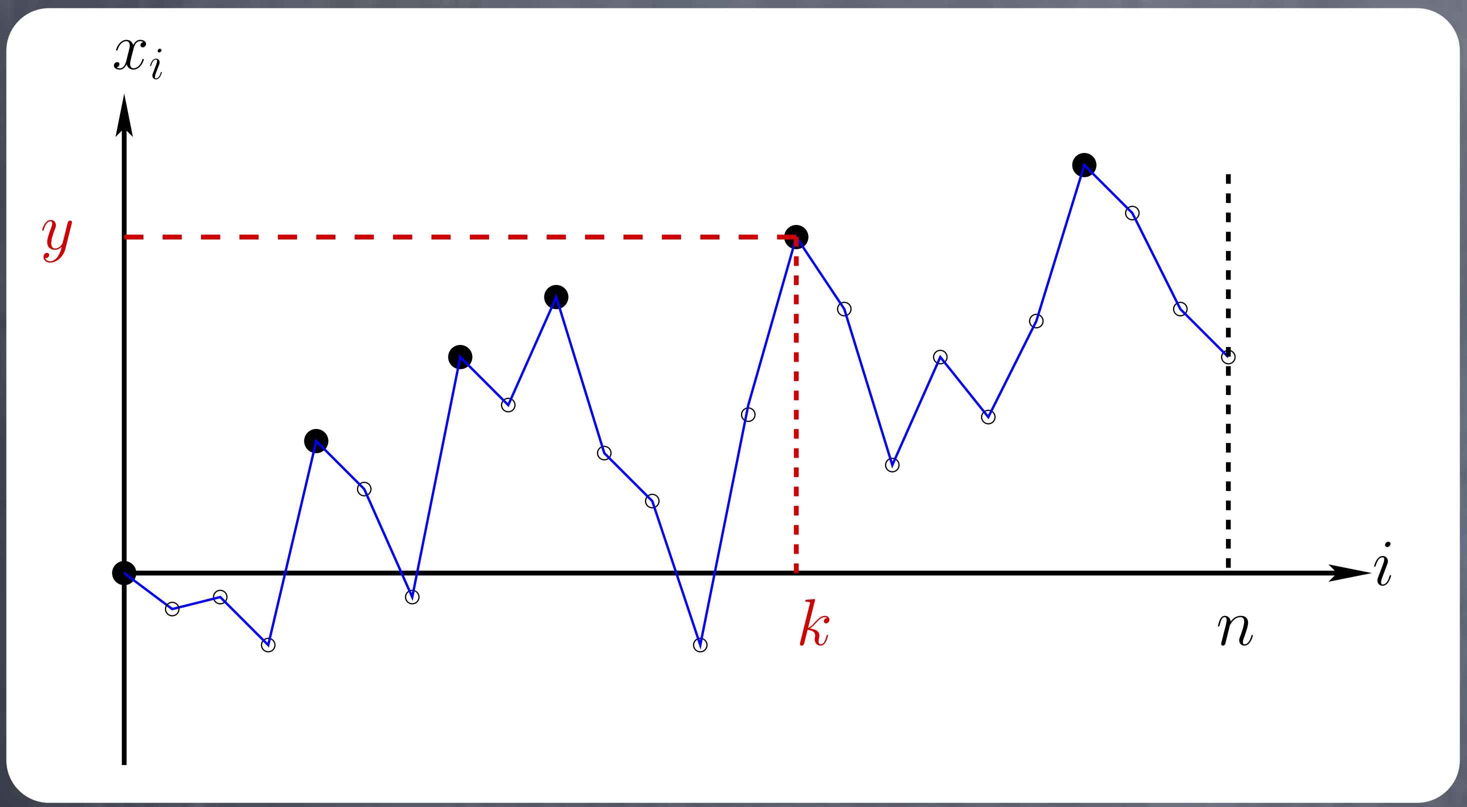
$$\langle R_n^2 \rangle - \langle R_n \rangle^2 \sim \log n$$

Universal probability distribution

$$P(R_n = M) = \frac{\left[n \atop M \right]}{n!}$$

Record statistics of i.i.d. random variables

$x_1, x_2, \cdots, x_N$: N i.i.d. random variables with PDF



Average nber of records

$$\langle R_n \rangle = \sum_{k=1}^{n} \frac{1}{k} \sim \log n$$

Variance

$$\langle R_n^2 \rangle - \langle R_n \rangle^2 \sim \log n$$

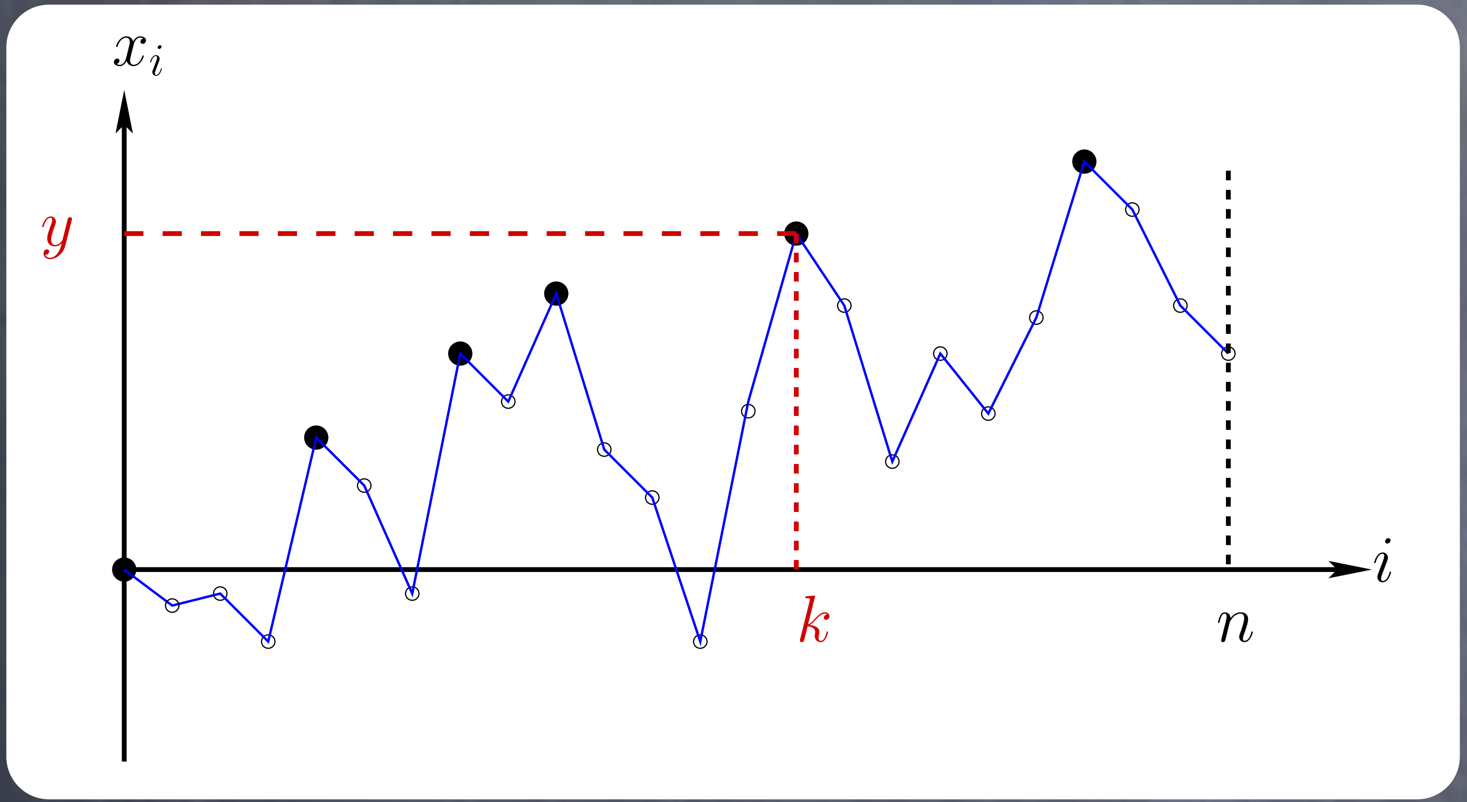
Universal probability distribution

$$P(R_n = M) = \frac{\left[{n \atop M} \right]}{n!}$$

Stirling numbers: number of permutations of n elements with M disjoint cycles

Record statistics of i.i.d. random variables

$x_1, x_2, \cdots, x_N : N$ i.i.d. random variables with PDF



Average nber of records

$$\langle R_n \rangle = \sum_{k=1}^{n} \frac{1}{k} \sim \log n$$

Variance

$$\langle R_n^2 \rangle - \langle R_n \rangle^2 \sim \log n$$

Universal probability distribution

$$P(R_n = M) = \frac{\left[{n \atop M} \right]}{n!}$$

Stirling numbers: number of permutations of n elements with M disjoint cycles

Gaussian for large n

$$\sim \frac{1}{\sqrt{2\pi \log n}} \exp\left(-\frac{(M - \log n)^2}{2 \log n} \right)$$

Record statistics of random walks

$x_0 = 0$

$x_i = x_{i-1} + \eta_i$ where the jumps η_is are i.i.d. with PDF $p(\eta)$

continuous & symmetric

Record statistics of random walks

$x_0 = 0$

$x_i = x_{i-1} + \eta_i$ where the jumps η_is are i.i.d. with PDF $p(\eta)$

continuous & symmetric

Including

- Ordinary random walks $\sigma^2 = \int_{-\infty}^{\infty} \eta^2\, p(\eta)\, d\eta < \infty$

$$x_n \sim \sigma \sqrt{n}$$

- Lévy flights

$$p(\eta) \propto a^{\mu}\, |\eta|^{-1-\mu}\ , \ |\eta| \to \infty$$

$$0 < \mu < 2$$

$x_n \sim a\, n^{1/\mu}$ μ is the Lévy index

Record statistics of random walks

$x_0 = 0$

$x_i = x_{i-1} + \eta_i$ where the jumps η_is are i.i.d. with PDF $p(\eta)$

continuous & symmetric

Including

- Ordinary random walks $\sigma^2 = \int_{-\infty}^{\infty} \eta^2 \, p(\eta) \, d\eta < \infty$

$$x_n \sim \sigma \sqrt{n}$$

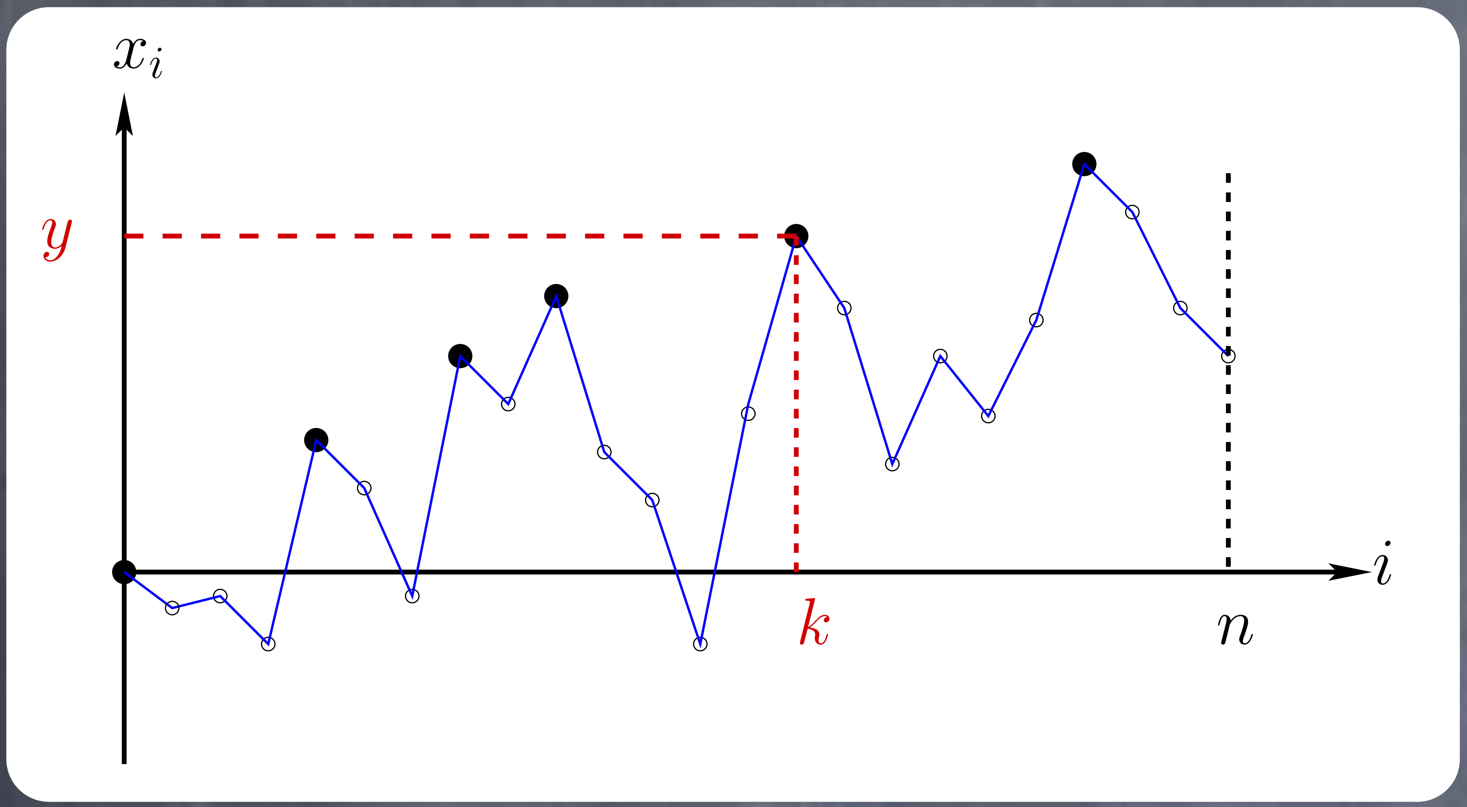
- Lévy flights $p(\eta) \propto a^\mu \, |\eta|^{-1-\mu} \, , \; |\eta| \to \infty$

$$0 < \mu < 2$$

$x_n \sim a \, n^{1/\mu}$ μ is the Lévy index

Q: Dependence of records on the jump distribution ?

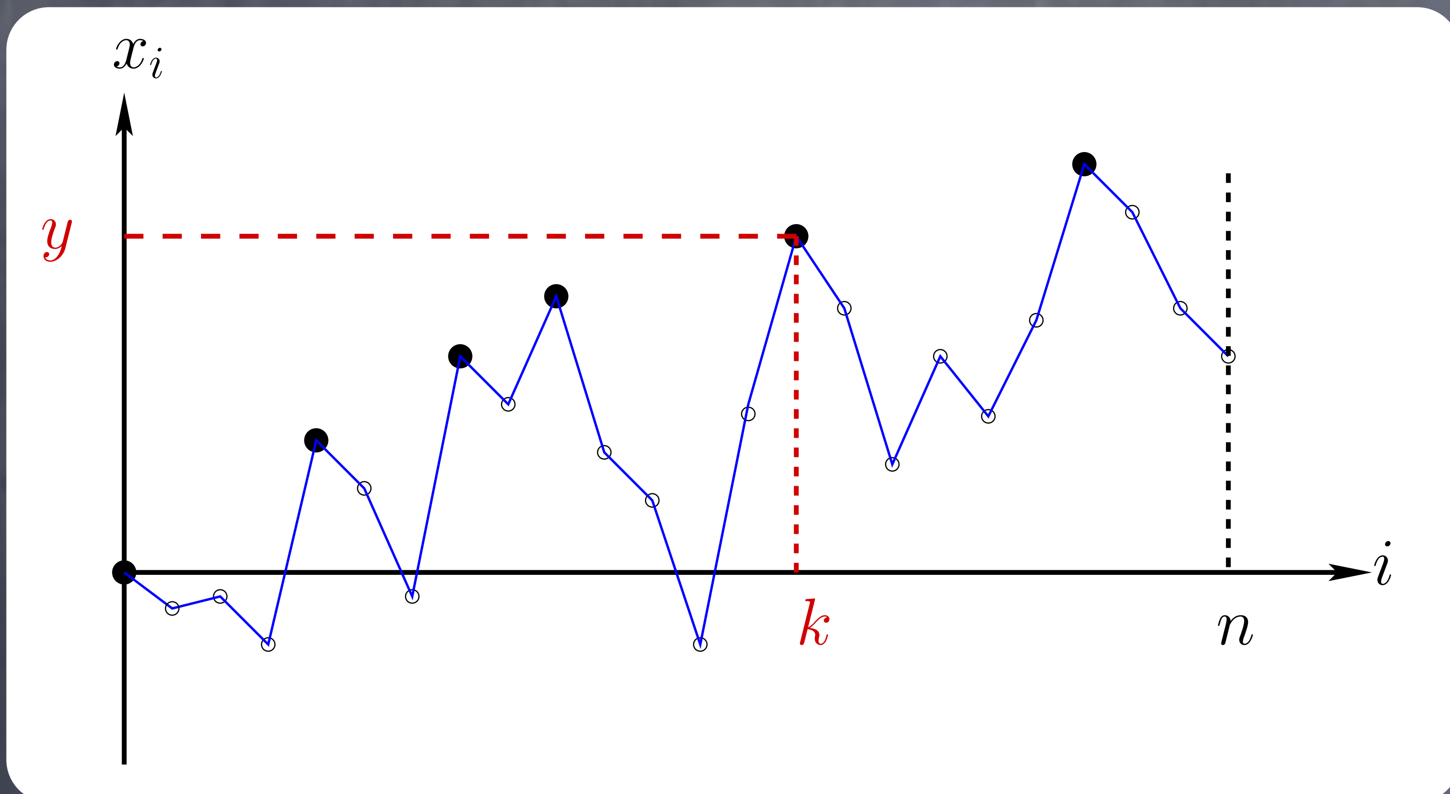
Mean record number of random walks



$$\langle R_n \rangle = \sum_{k=0}^{n} r_k \ , \ \ r_k = \langle \sigma_k \rangle$$

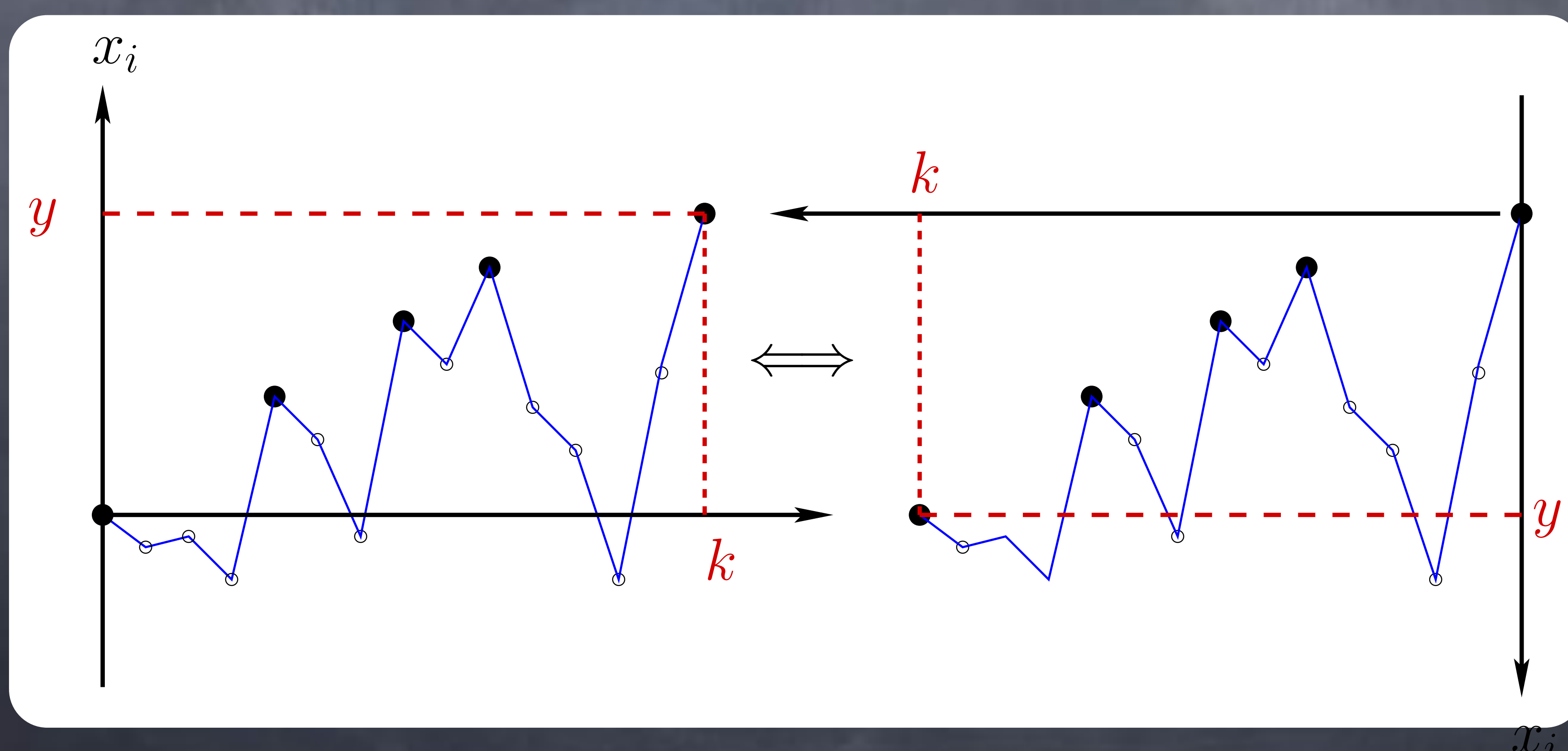
r_k is the proba. that a record is broken at step k

Mean record number of random walks

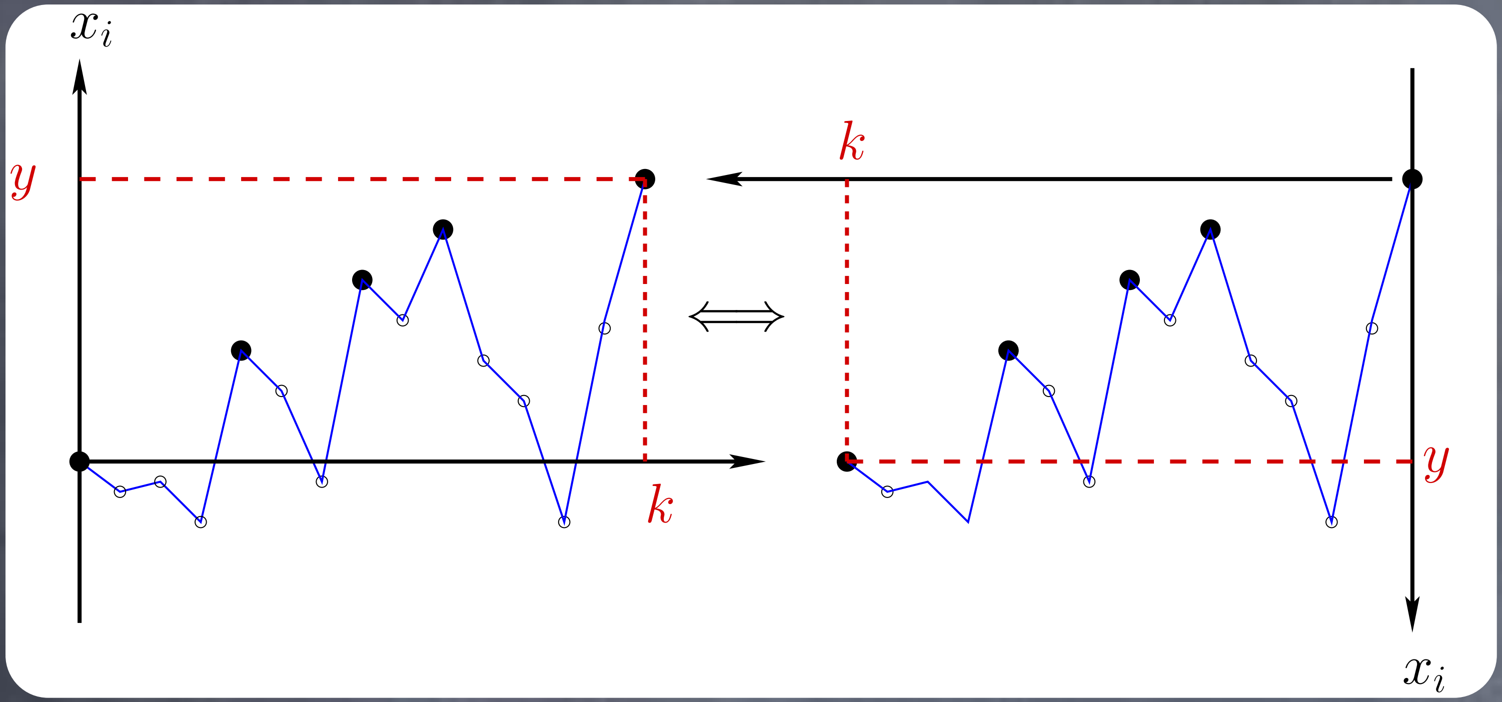


$$\langle R_n \rangle = \sum_{k=0}^{n} r_k \ , \ \ r_k = \langle \sigma_k \rangle$$

r_k is the proba. that a record is broken at step k



Mean record number of random walks



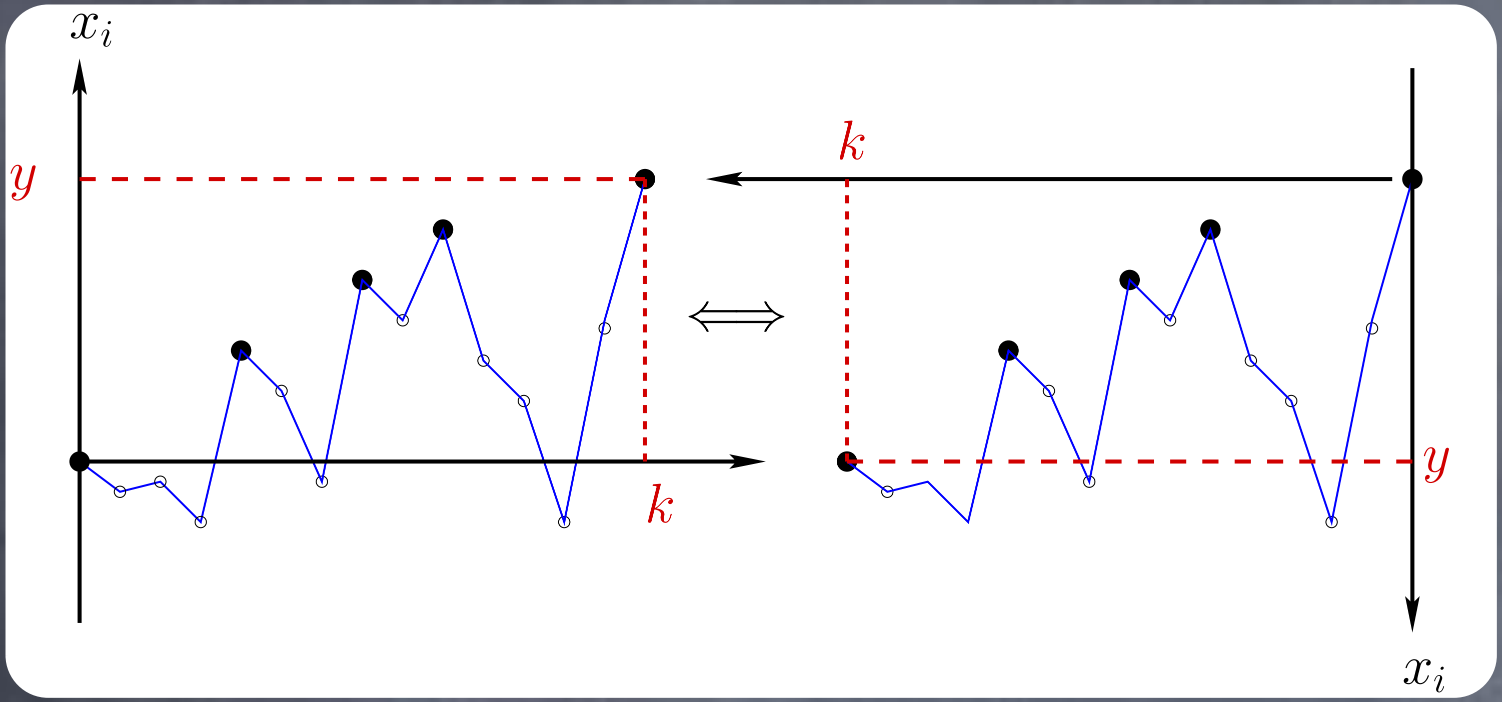
$$\langle R_n \rangle = \sum_{k=0}^{n} r_k \; , \; r_k = \langle \sigma_k \rangle$$

r_k is the proba. that a record is broken at step k

$$\Longrightarrow r_k = q_-(k)$$

$=$ Proba. that the walker stays negative up to step k starting from the origin

Mean record number of random walks



$$\langle R_n \rangle = \sum_{k=0}^{n} r_k \; , \; r_k = \langle \sigma_k \rangle$$

r_k is the proba. that a record is broken at step k

$$\Longrightarrow r_k = q_-(k)$$

$=$ Proba. that the walker stays negative up to step k starting from the origin

$q_-(k)$ is given by the Sparre Andersen Theorem

Mean record number of random walks

$$\langle R_n \rangle = \sum_{k=0}^{n} r_k = \sum_{k=0}^{n} q_-(k)$$

$q_-(k)$ is given by the Sparre Andersen Theorem

Mean record number of random walks

$$\langle R_n \rangle = \sum_{k=0}^{n} r_k = \sum_{k=0}^{n} q_-(k)$$

$q_-(k)$ is given by the Sparre Andersen Theorem

For symmetric RW

$$\sum_{k=0}^{\infty} q_-(k) z^k = \frac{1}{\sqrt{1-z}} \Longrightarrow q_-(k) = \frac{1}{2^{2k}} \binom{2k}{k} \sim \frac{1}{\sqrt{\pi k}}$$

Mean record number of random walks

$$\langle R_n \rangle = \sum_{k=0}^{n} r_k = \sum_{k=0}^{n} q_-(k)$$

$q_-(k)$ is given by the Sparre Andersen Theorem

For symmetric RW

$$\sum_{k=0}^{\infty} q_-(k) z^k = \frac{1}{\sqrt{1-z}} \Longrightarrow q_-(k) = \frac{1}{2^{2k}} \binom{2k}{k} \sim \frac{1}{\sqrt{\pi k}}$$

Universal, i.e. independent of the jump distribution !

Mean record number of random walks

$$\langle R_n \rangle = \sum_{k=0}^{n} r_k = \sum_{k=0}^{n} q_-(k)$$

$q_-(k)$ is given by the Sparre Andersen Theorem

For symmetric RW

$$\sum_{k=0}^{\infty} q_-(k) z^k = \frac{1}{\sqrt{1-z}} \Longrightarrow q_-(k) = \frac{1}{2^{2k}} \binom{2k}{k} \sim \frac{1}{\sqrt{\pi k}}$$

Universal, i.e. independent of the jump distribution !

$$\langle R_n \rangle = \frac{2\,\Gamma(3/2+n)}{\sqrt{\pi}\, n!} \sim \frac{2}{\sqrt{\pi}} \sqrt{n}$$

Majumdar, Ziff `08

Record statistics of random walks with a drift

$x_0 = 0$

$x_i = x_{i-1} + \eta_i$ where the jumps η_is are i.i.d. with PDF $p(\eta)$

continuous & symmetric

RW with a drift $y_n = x_n + c\,n$

Record statistics of random walks with a drift

$x_0 = 0$

$x_i = x_{i-1} + \eta_i$ where the jumps η_is are i.i.d. with PDF $p(\eta)$

continuous & symmetric

RW with a drift $y_n = x_n + c\,n$

Mean number of records of y_n: $\langle R_n \rangle = \sum_{k=0}^{n} r_k = \sum_{k=0}^{n} q_-(k)$

$$q_-(k) = \Pr(y_1 < 0, y_2 < 0, \cdots, y_k < 0)$$

Record statistics of random walks with a drift

$x_0 = 0$

$x_i = x_{i-1} + \eta_i$ where the jumps η_is are i.i.d. with PDF $p(\eta)$

continuous & symmetric

RW with a drift $y_n = x_n + c\,n$

Mean number of records of y_n: $\langle R_n \rangle = \sum_{k=0}^{n} r_k = \sum_{k=0}^{n} q_-(k)$

$$q_-(k) = \Pr(y_1 < 0, y_2 < 0, \cdots, y_k < 0)$$

(Generalized) Sparre Andersen theorem

$$\boxed{\sum_{k=0}^{\infty} q_-(k) z^k = \exp\left(\sum_{k=1}^{\infty} \frac{z^k}{k} \Pr(y_k < 0) \right)}$$

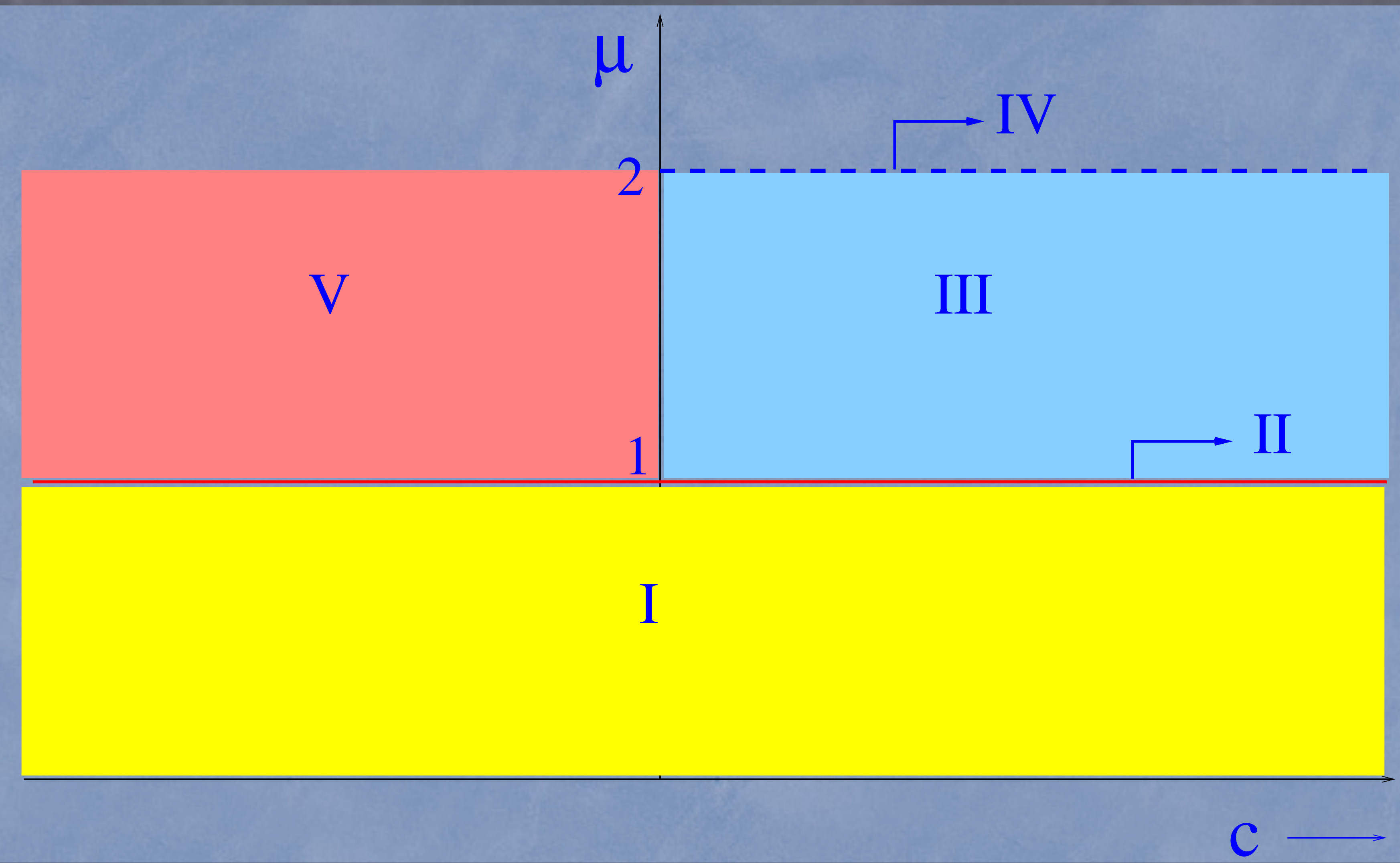
Record statistics of random walks with a drift

$$x_0 = 0$$
$$x_i = x_{i-1} + \eta_i, \qquad \hat{p}(k) = \int_{-\infty}^{+\infty} p(\eta) e^{ik\eta}\, d\eta = 1 - |ak|^\mu + \cdots$$

$$y_n = x_n + c\, n$$

RW with a drift

Majumdar, G. S., Wergen `12



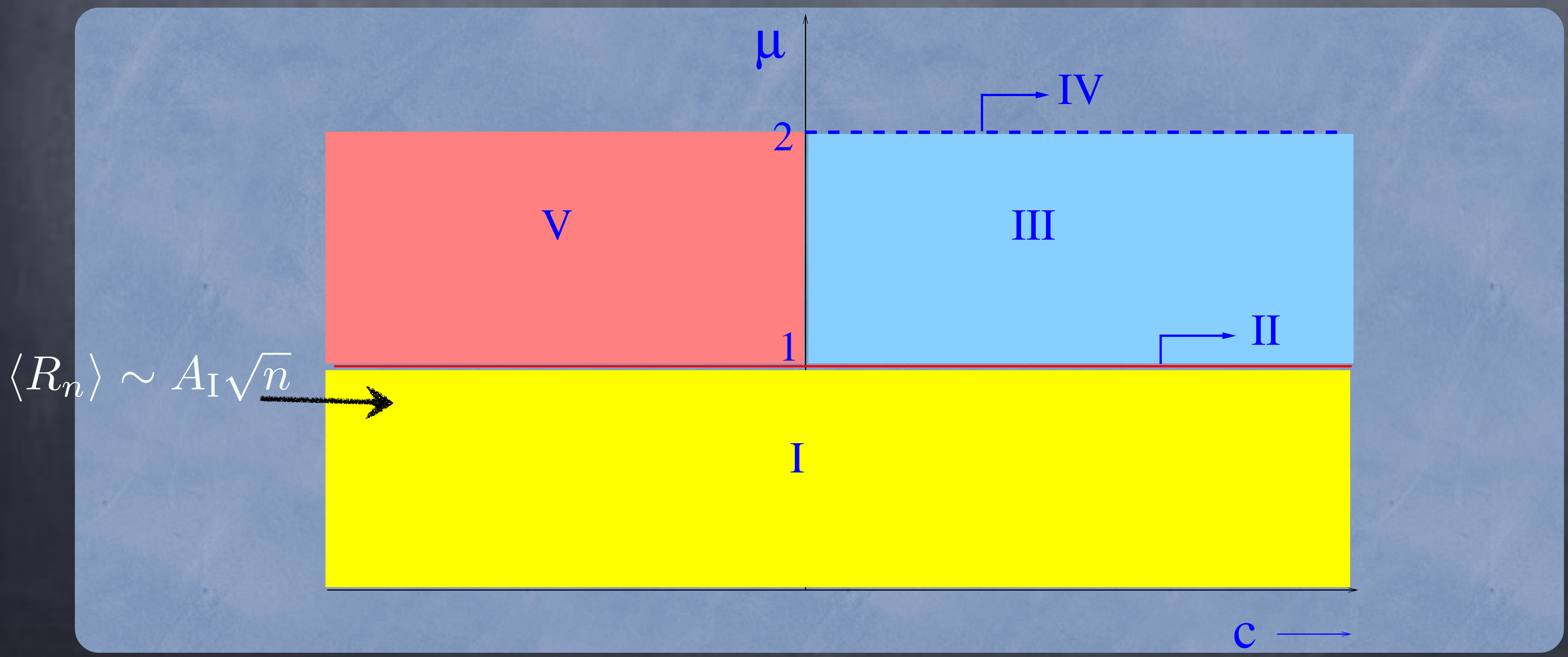
Record statistics of random walks with a drift

$$x_0 = 0$$
$$x_i = x_{i-1} + \eta_i, \qquad \hat{p}(k) = \int_{-\infty}^{+\infty} p(\eta) e^{ik\eta}\, d\eta = 1 - |ak|^\mu + \cdots$$

$$y_n = x_n + c\, n$$

RW with a drift

Majumdar, G. S., Wergen `12



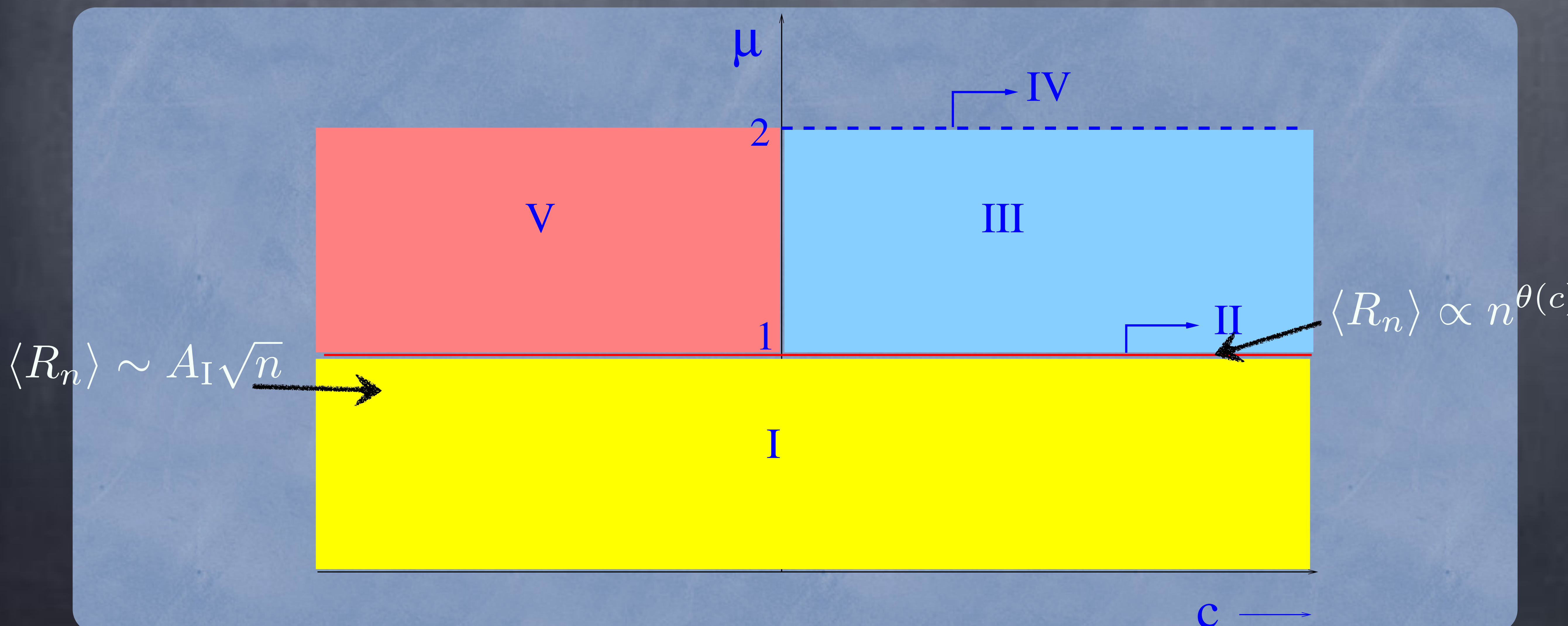
Record statistics of random walks with a drift

$$x_0 = 0$$
$$x_i = x_{i-1} + \eta_i, \qquad \hat{p}(k) = \int_{-\infty}^{+\infty} p(\eta) e^{ik\eta}\, d\eta = 1 - |ak|^\mu + \cdots$$

$$y_n = x_n + c\, n$$

RW with a drift

Majumdar, G. S., Wergen `12



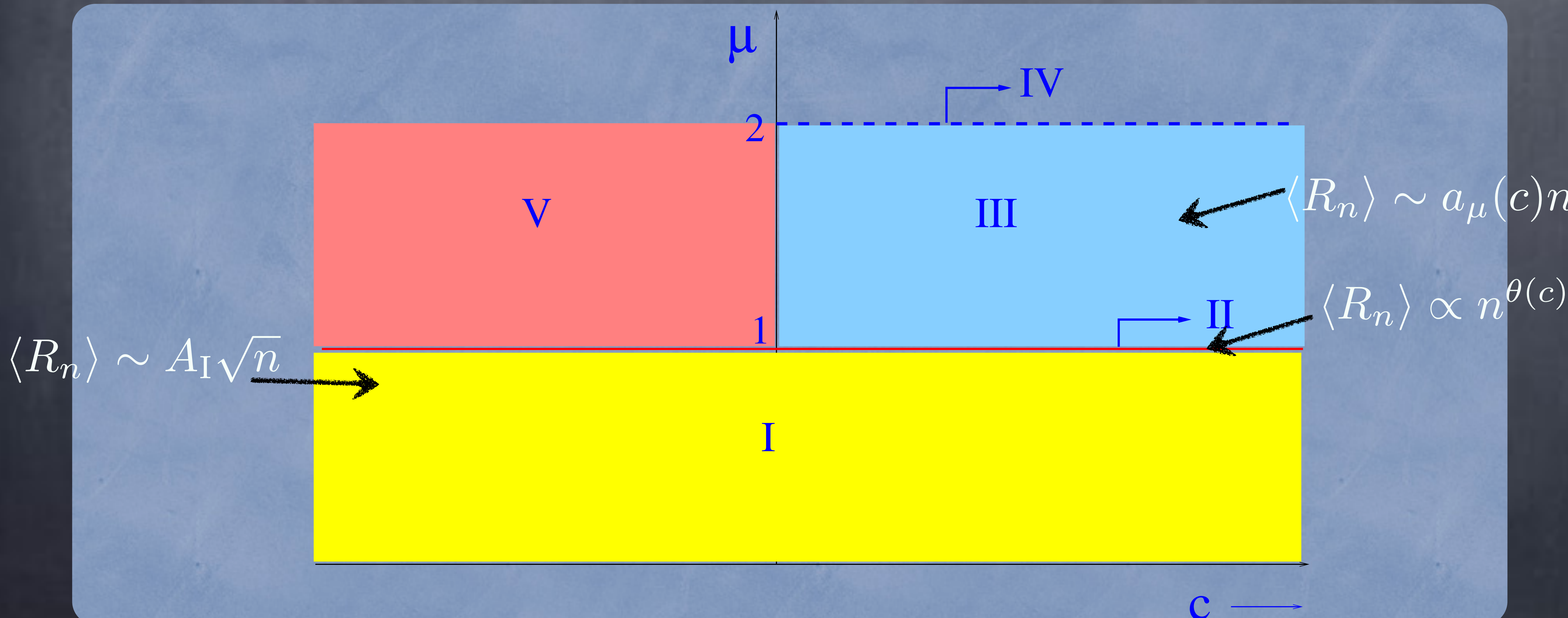
Record statistics of random walks with a drift

$$x_0 = 0$$
$$x_i = x_{i-1} + \eta_i, \qquad \hat{p}(k) = \int_{-\infty}^{+\infty} p(\eta) e^{ik\eta}\, d\eta = 1 - |ak|^\mu + \cdots$$

$$y_n = x_n + c\, n$$

RW with a drift

Majumdar, G. S., Wergen `12



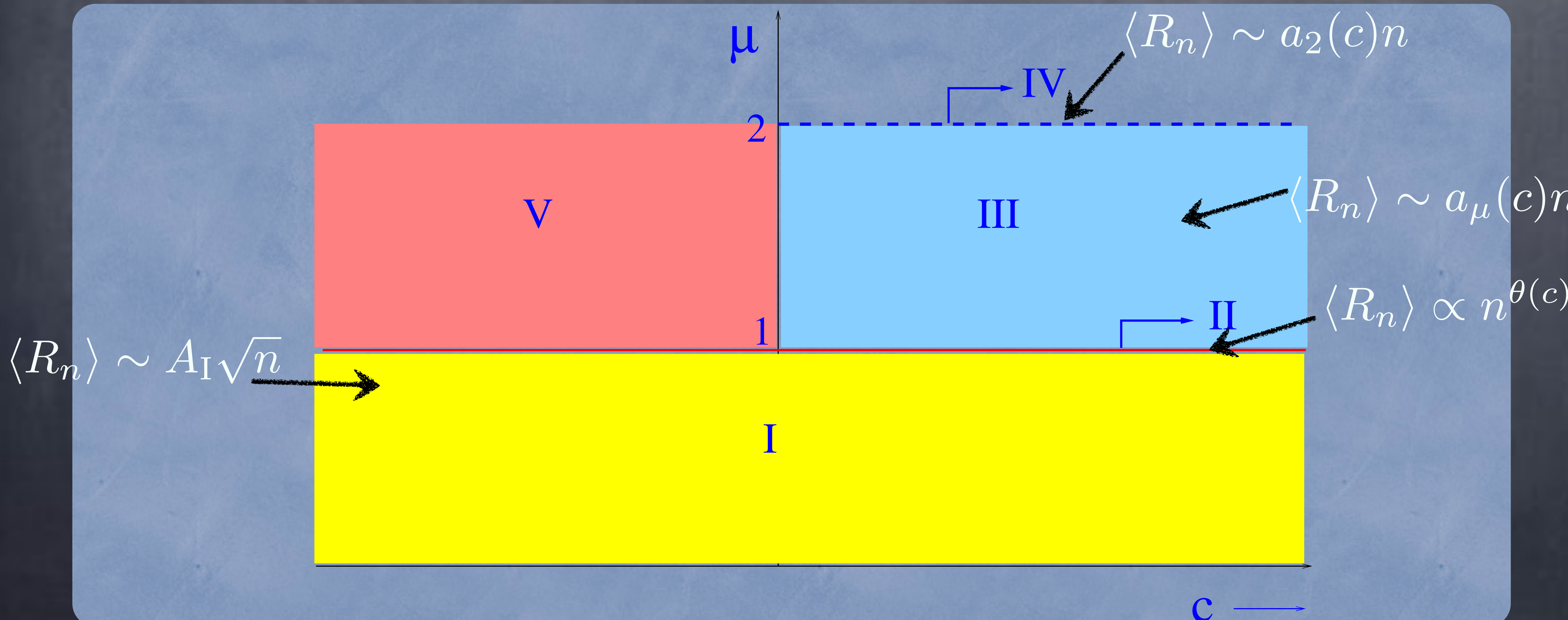
Record statistics of random walks with a drift

$$x_0 = 0$$
$$x_i = x_{i-1} + \eta_i, \qquad \hat{p}(k) = \int_{-\infty}^{+\infty} p(\eta) e^{ik\eta}\, d\eta = 1 - |ak|^\mu + \cdots$$

$$y_n = x_n + c\, n$$

RW with a drift

Majumdar, G. S., Wergen `12



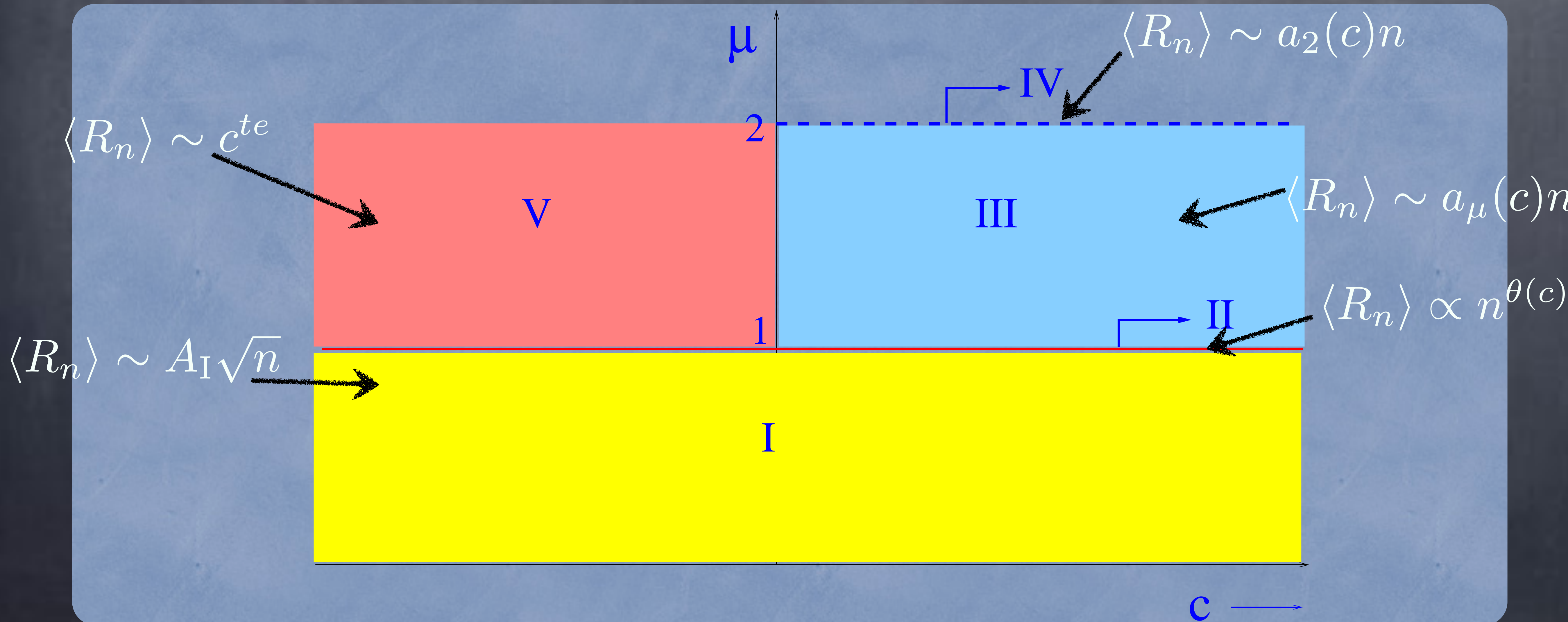
Record statistics of random walks with a drift

$$x_0 = 0$$
$$x_i = x_{i-1} + \eta_i, \qquad \hat{p}(k) = \int_{-\infty}^{+\infty} p(\eta) e^{ik\eta}\, d\eta = 1 - |ak|^\mu + \cdots$$

$$y_n = x_n + c\, n$$

RW with a drift

Majumdar, G. S., Wergen `12



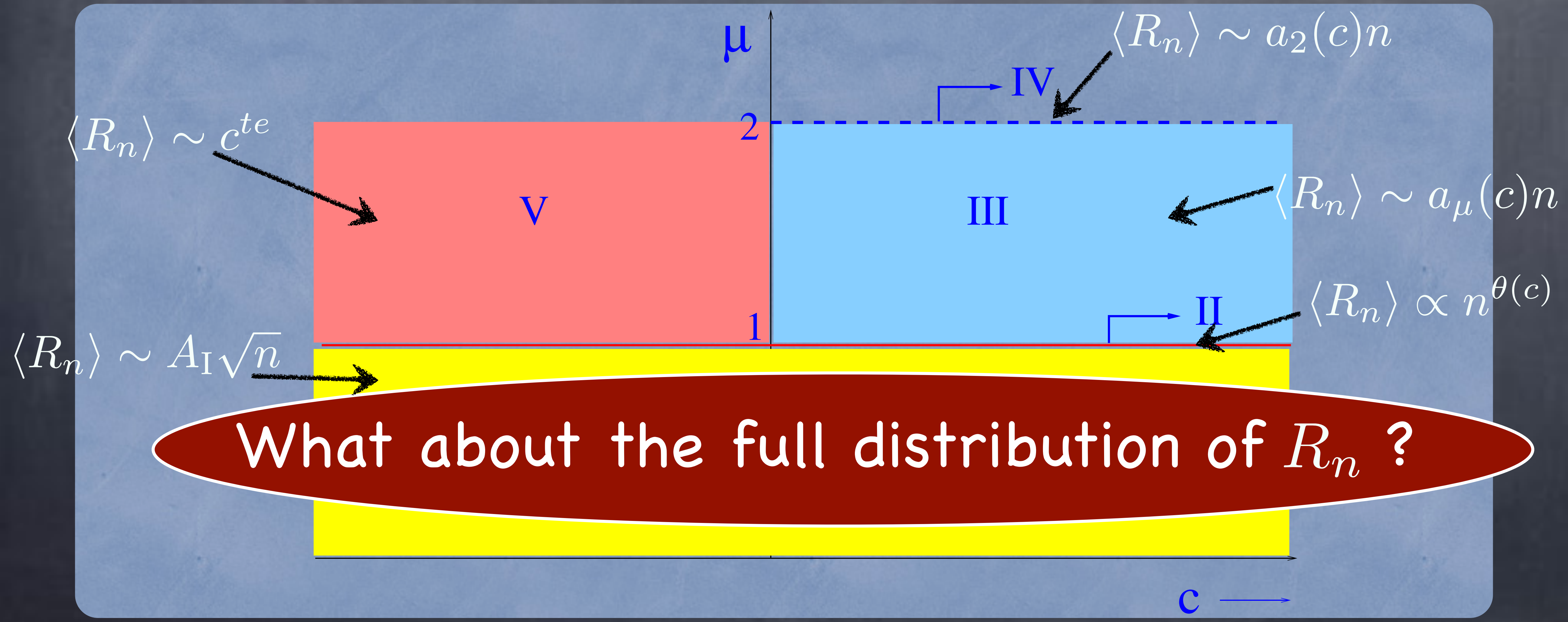
Record statistics of random walks with a drift

$$x_0 = 0$$
$$x_i = x_{i-1} + \eta_i, \qquad \hat{p}(k) = \int_{-\infty}^{+\infty} p(\eta) e^{ik\eta}\, d\eta = 1 - |ak|^\mu + \cdots$$

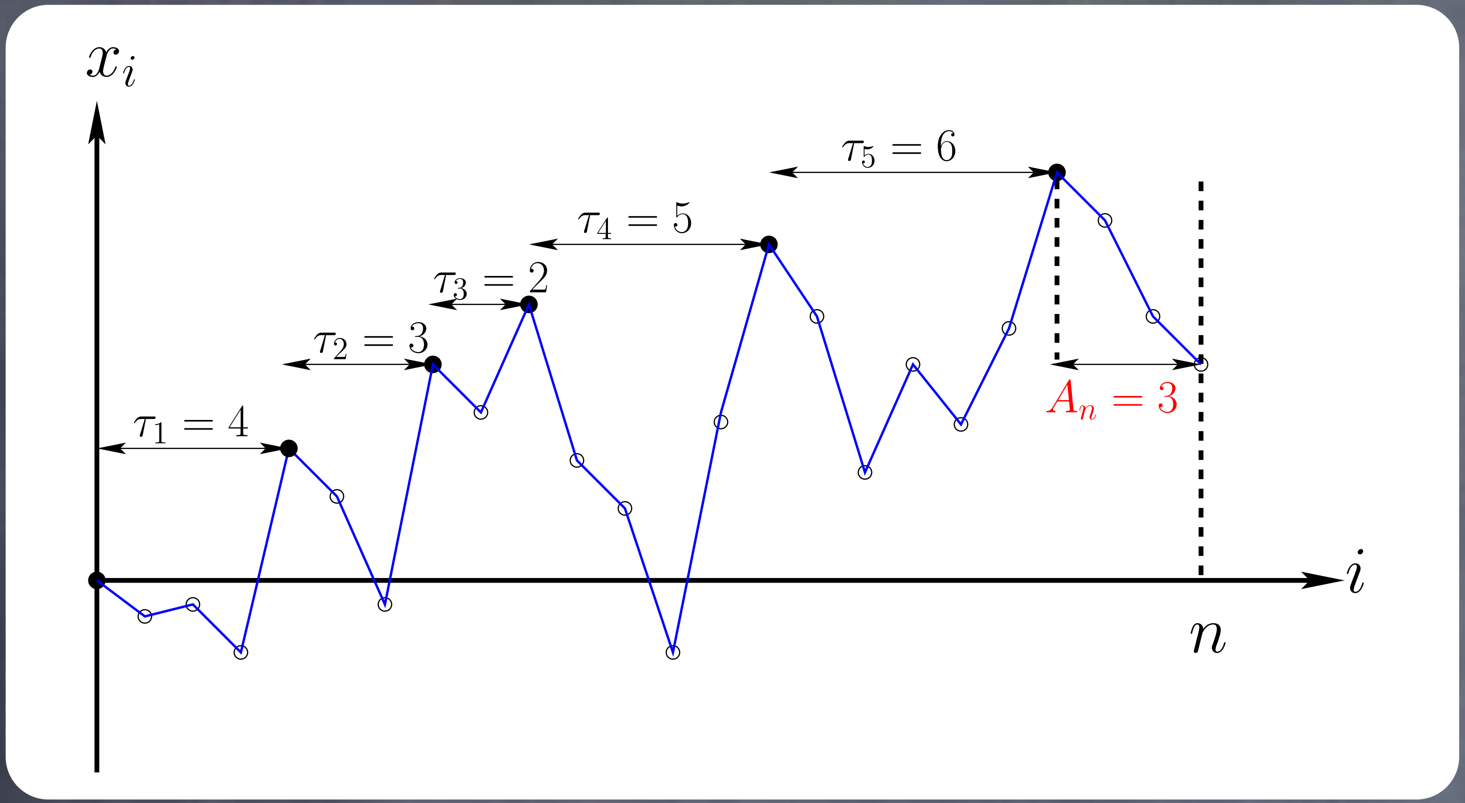
$$y_n = x_n + c\, n$$

RW with a drift

Majumdar, G. S., Wergen `12



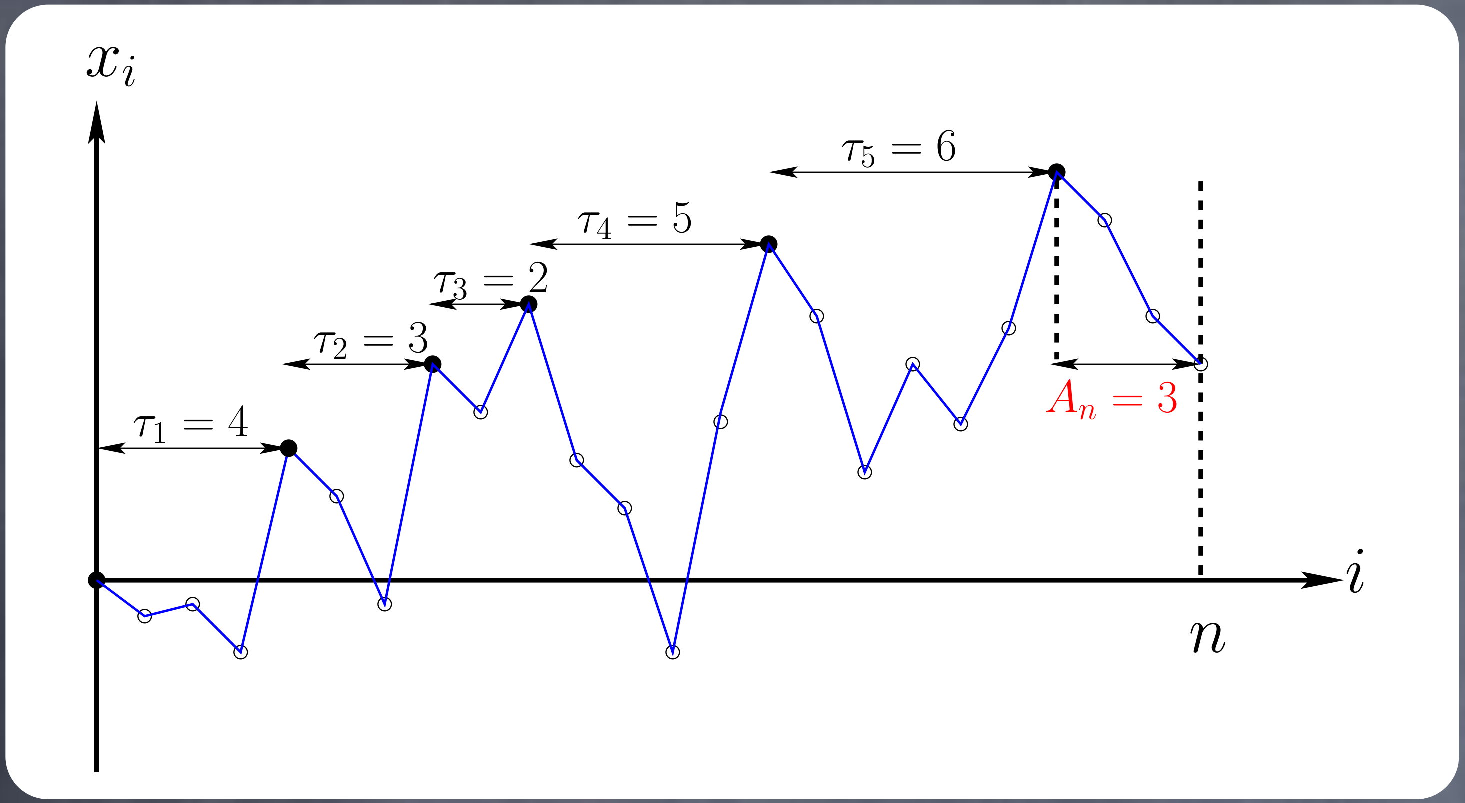
Renewal approach to records of RW



Joint distribution of

$$R_n, \tau_1, \tau_2, \cdots, \tau_{R_n - 1}, A_n \ ?$$

Renewal approach to records of RW



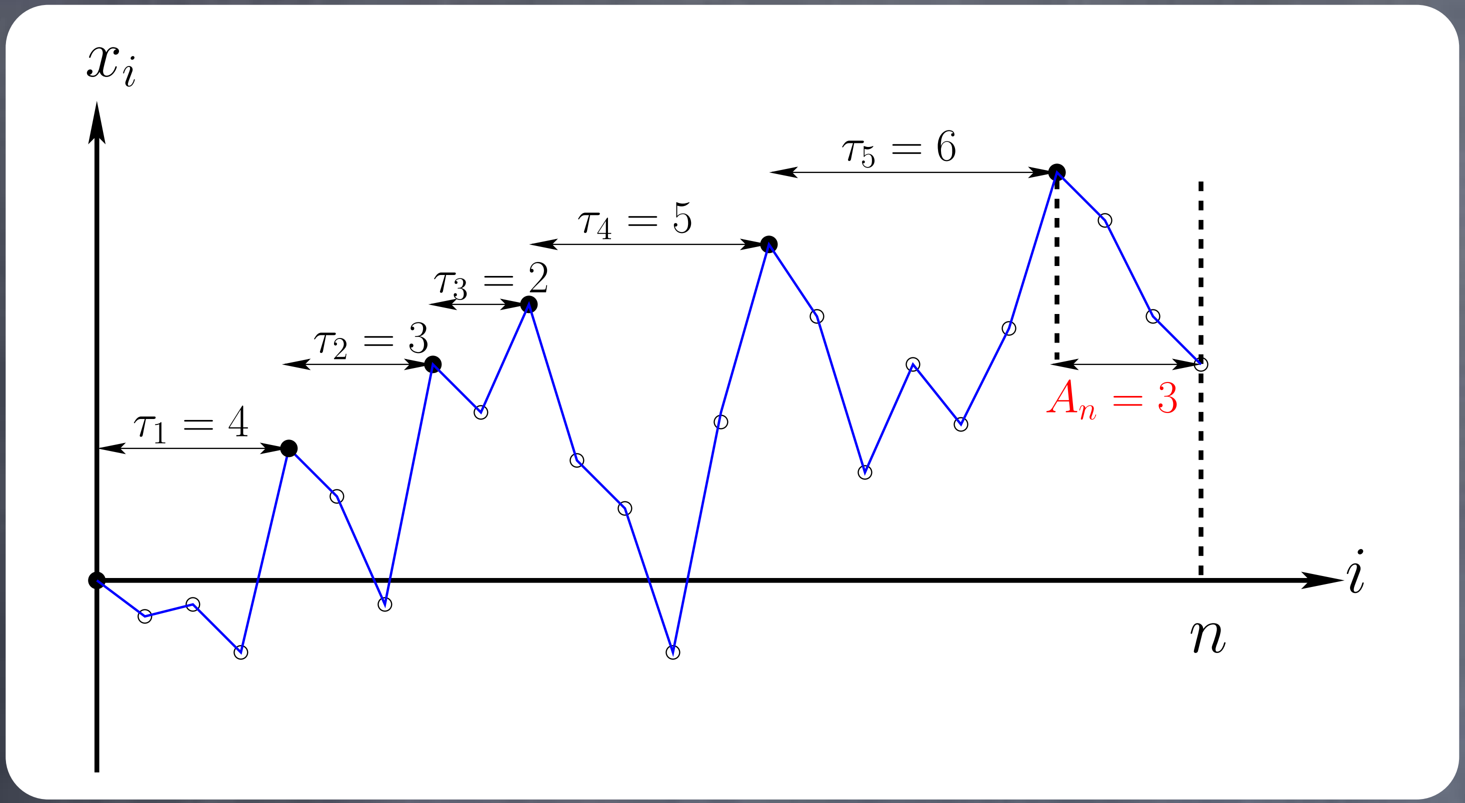
Joint distribution of

$R_n, \tau_1, \tau_2, \cdots, \tau_{R_n-1}, A_n$?

- RW is a Markov process $\Longrightarrow$ $\tau_1, \tau_2, \cdots, \tau_{R_n-1}, A_n$ are independent except for the global constraint

$$\sum_{i=1}^{R_n-1} \tau_i + A_n = n$$

Renewal approach to records of RW



Joint distribution of

$$R_n, \tau_1, \tau_2, \cdots, \tau_{R_n - 1}, A_n \ ?$$

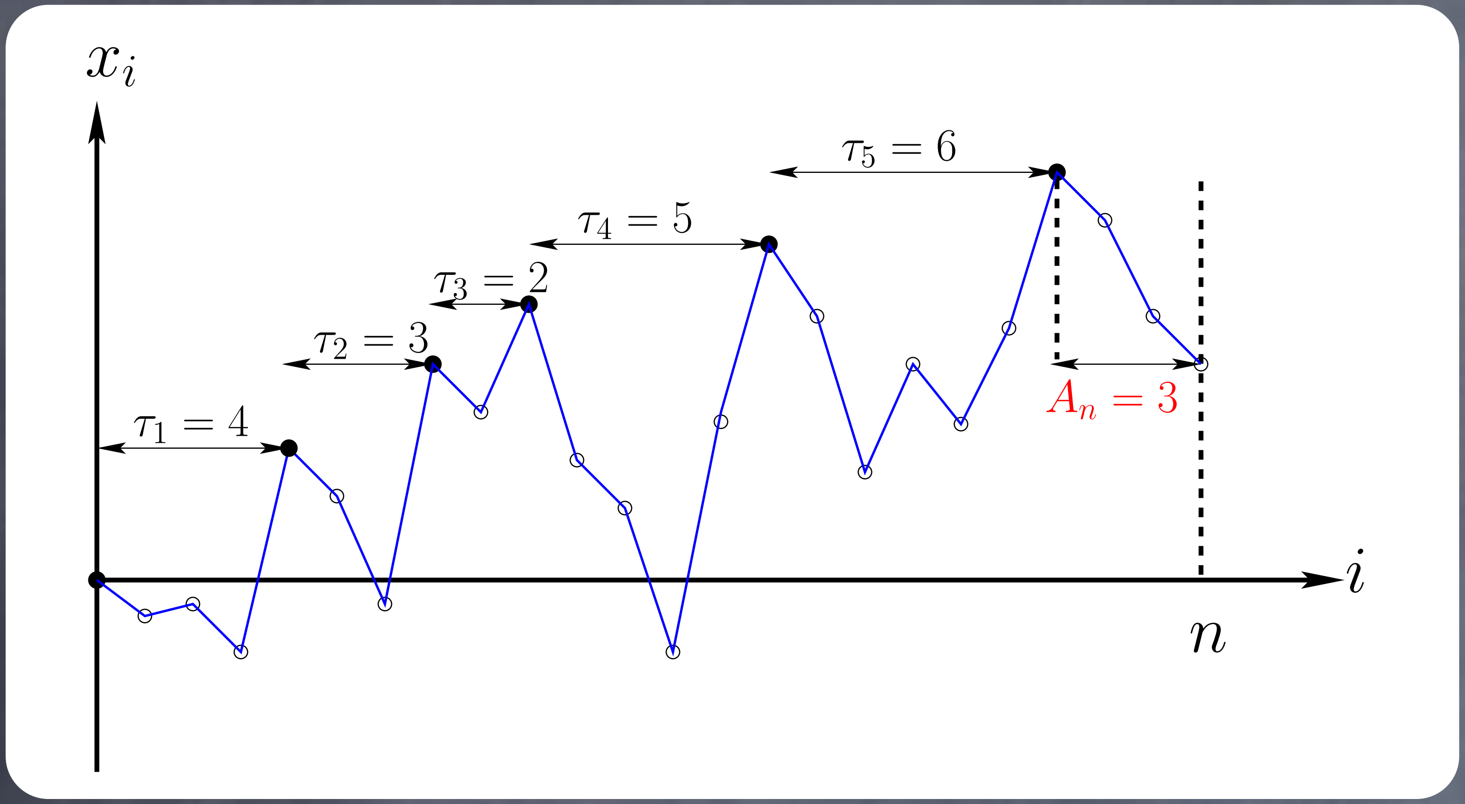
- RW is a Markov process $\Longrightarrow$ $\tau_1, \tau_2, \cdots, \tau_{R_n-1}, A_n$ are independent except for the global constraint

$$\sum_{i=1}^{R_n - 1} \tau_i + A_n = n$$

- RW is translationally invariant $\Longrightarrow$ τ_is are identical

while A_n has different statistics

Renewal approach to records of RW



Joint distribution of

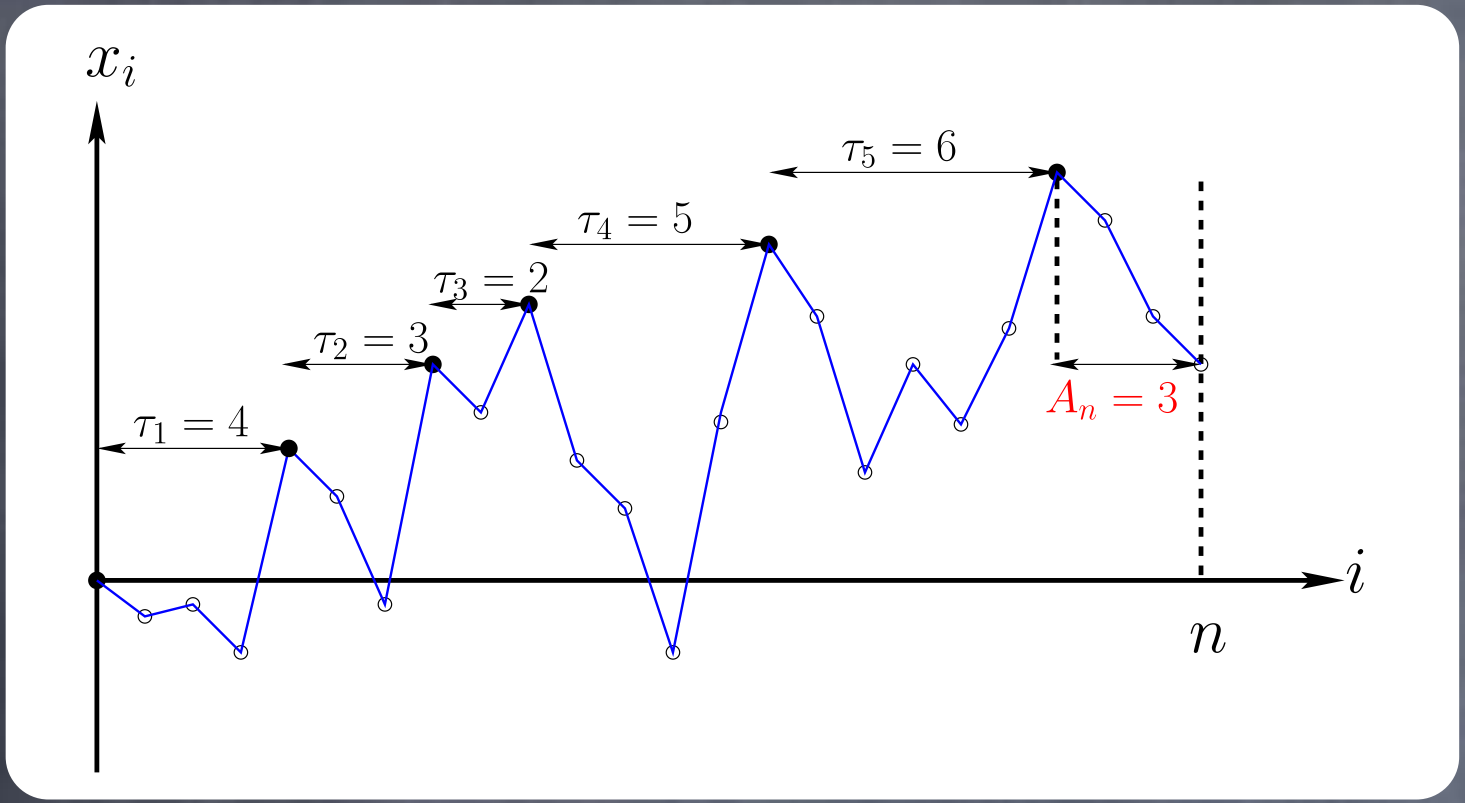
$$R_n, \tau_1, \tau_2, \cdots, \tau_{R_n - 1}, A_n \ ?$$

Two main objects:

- Persistence (or survival) probability
  $q_-(k) = \Pr(y_1 < y_0, y_2 < y_0, \cdots, y_k < y_0)$ indep. of y_0

- Distribution of first-passage time (from below)
  $f_-(k) = \Pr(y_1 < y_0, y_2 < y_0, \cdots, y_{k-1} < y_0, y_k > y_0)$
  $= q_-(k) - q_-(k-1)$ indep. of y_0

Renewal approach to records of RW



Joint distribution of

$$R_n, \tau_1, \tau_2, \cdots, \tau_{R_n - 1}, A_n$$

$$\Pr(R_n = m, \tau_1 = \ell_1, \cdots, \tau_{m-1} = \ell_{m-1}, A_n = a) = P(\vec{\ell}, m, n)$$

$$\boxed{P(\vec{l}, m, n) = f_-(\ell_1) f_-(\ell_2) \cdots f_-(\ell_{m-1}) q_-(a) \delta\left(\sum_{k=1}^{m-1} \ell_k + a, n\right)}$$

first passage proba.

survival proba.

Proba. distribution of the number of records

$$P(m,n) = \Pr.(R_n = m) = \sum_{\ell_1=1}^{\infty} \sum_{\ell_2=1}^{\infty} \cdots \sum_{\ell_{m-1}=1}^{\infty} \sum_{a=0}^{\infty} P(\vec{l}, m, n)$$

with

$$P(\vec{l}, m, n) = f_-(\ell_1) f_-(\ell_2) \cdots f_-(\ell_{m-1}) q_-(a) \delta\left(\sum_{k=1}^{m-1} \ell_k + a, n\right)$$

Proba. distribution of the number of records

$$P(m,n) = \Pr.(R_n = m) = \sum_{\ell_1=1}^{\infty}\sum_{\ell_2=1}^{\infty}\cdots\sum_{\ell_{m-1}=1}^{\infty}\sum_{a=0}^{\infty} P(\vec{l},m,n)$$

with

$$P(\vec{l},m,n) = f_-(\ell_1)f_-(\ell_2)\cdots f_-(\ell_{m-1})q_-(a)\delta\left(\sum_{k=1}^{m-1}\ell_k + a, n\right)$$

Generating function w.r.t. the number of steps

$$\sum_{n=0}^{\infty} P(m,n)z^n = \left(\sum_{\ell\geq 1} z^\ell f_-(\ell)\right)^{m-1} \sum_{a\geq 0} z^a q_-(a)$$

$$= \left[\tilde{f}_-(z)\right]^{m-1} \tilde{q}_-(z)$$

Proba. distribution of the number of records

$$P(m,n) = \Pr.(R_n = m) = \sum_{\ell_1=1}^{\infty}\sum_{\ell_2=1}^{\infty}\cdots\sum_{\ell_{m-1}=1}^{\infty}\sum_{a=0}^{\infty} P(\vec{l},m,n)$$

with

$$P(\vec{l},m,n) = f_-(\ell_1)f_-(\ell_2)\cdots f_-(\ell_{m-1})q_-(a)\delta\left(\sum_{k=1}^{m-1}\ell_k + a, n\right)$$

Generating function w.r.t. the number of steps

$$\sum_{n=0}^{\infty} P(m,n)z^n = \left(\sum_{\ell\geq 1} z^{\ell} f_-(\ell)\right)^{m-1} \sum_{a\geq 0} z^a q_-(a)$$

$$= \left[\tilde{f}_-(z)\right]^{m-1} \tilde{q}_-(z)$$

(for symmetric jumps) $= \left[1-\sqrt{1-z}\right]^{m-1} \dfrac{1}{\sqrt{1-z}}$

Proba. distribution of the number of records

- By ``inverting'' the GF (for symmetric jumps):

Majumdar, Ziff `08

$$P(m,n) = \binom{2n-m+1}{n} 2^{-2n+m-1} \ , \ m \leq n+1$$

Proba. distribution of the number of records

- By ``inverting'' the GF (for symmetric jumps):

Majumdar, Ziff `08

$$P(m,n) = \binom{2n-m+1}{n} 2^{-2n+m-1} \; , \; m \leq n+1$$

For $n \gg 1:$ $\quad P(m,n) \sim \frac{1}{\sqrt{n}} g_0\left(\frac{m}{\sqrt{n}}\right) \; , \; g_0(x) = \frac{1}{\sqrt{\pi}} e^{-\frac{x^2}{4}} \; , \; x > 0$

Proba. distribution of the number of records

- By ``inverting'' the GF (for symmetric jumps):

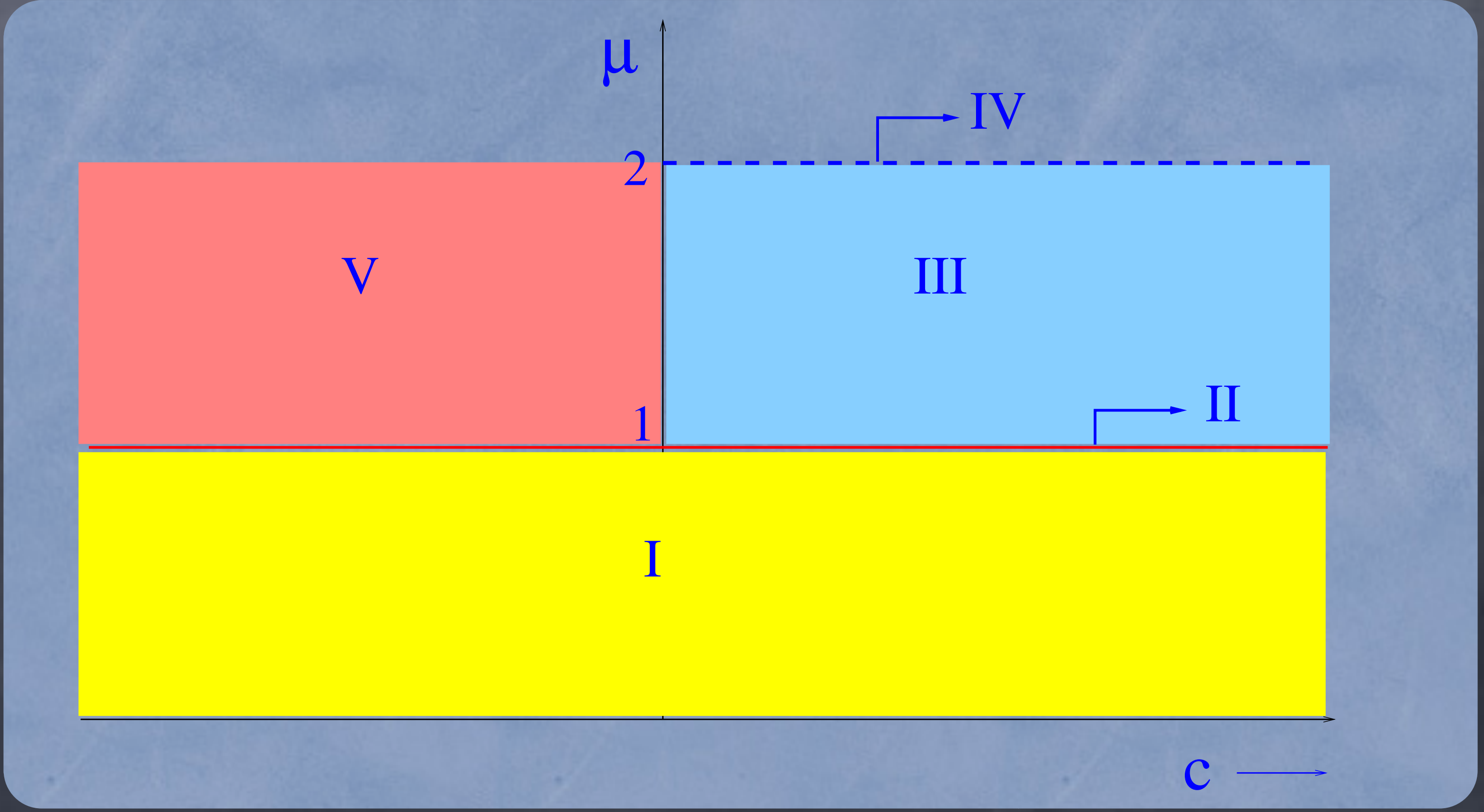
Majumdar, Ziff `08

$$P(m,n) = \binom{2n-m+1}{n} 2^{-2n+m-1} \; , \; m \leq n+1$$

For $n \gg 1:$ $\quad P(m,n) \sim \frac{1}{\sqrt{n}} g_0\left(\frac{m}{\sqrt{n}}\right) \; , \; g_0(x) = \frac{1}{\sqrt{\pi}} e^{-\frac{x^2}{4}} \; , \; x > 0$

- RW with a drift

Majumdar, G. S., Wergen `12



Proba. distribution of the number of records

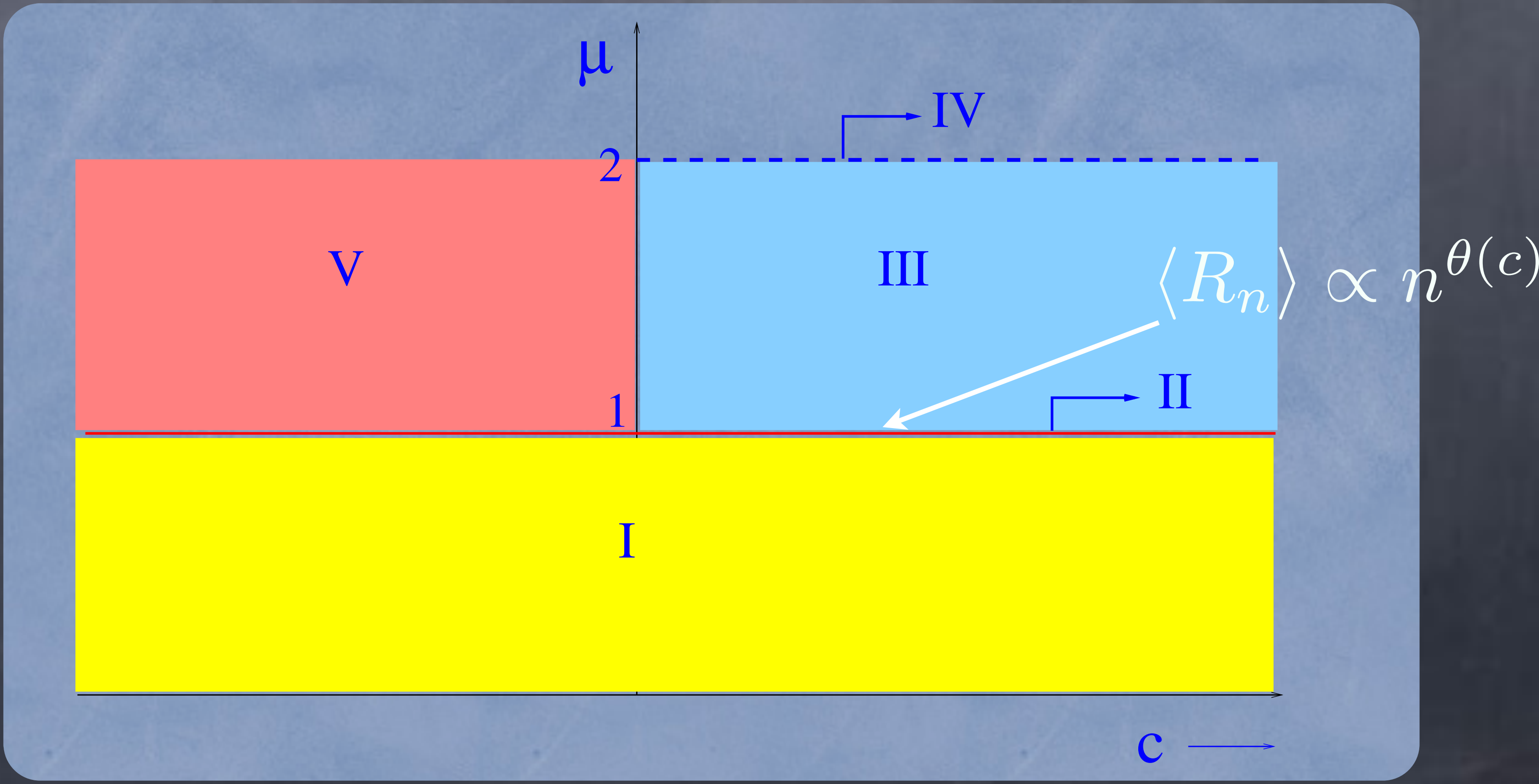
- By ``inverting'' the GF (for symmetric jumps):

Majumdar, Ziff `08

$$P(m,n) = \binom{2n-m+1}{n} 2^{-2n+m-1} \; , \; m \leq n+1$$

For $n \gg 1:\quad P(m,n) \sim \frac{1}{\sqrt{n}} g_0\left(\frac{m}{\sqrt{n}}\right) \; , \; g_0(x) = \frac{1}{\sqrt{\pi}} e^{-\frac{x^2}{4}} \; , \; x > 0$

- RW with a drift Majumdar, G. S., Wergen `12



Proba. distribution of the number of records

- By "inverting" the GF (for symmetric jumps):

Majumdar, Ziff '08

$$P(m,n) = \binom{2n-m+1}{n} 2^{-2n+m-1} \; , \; m \leq n+1$$

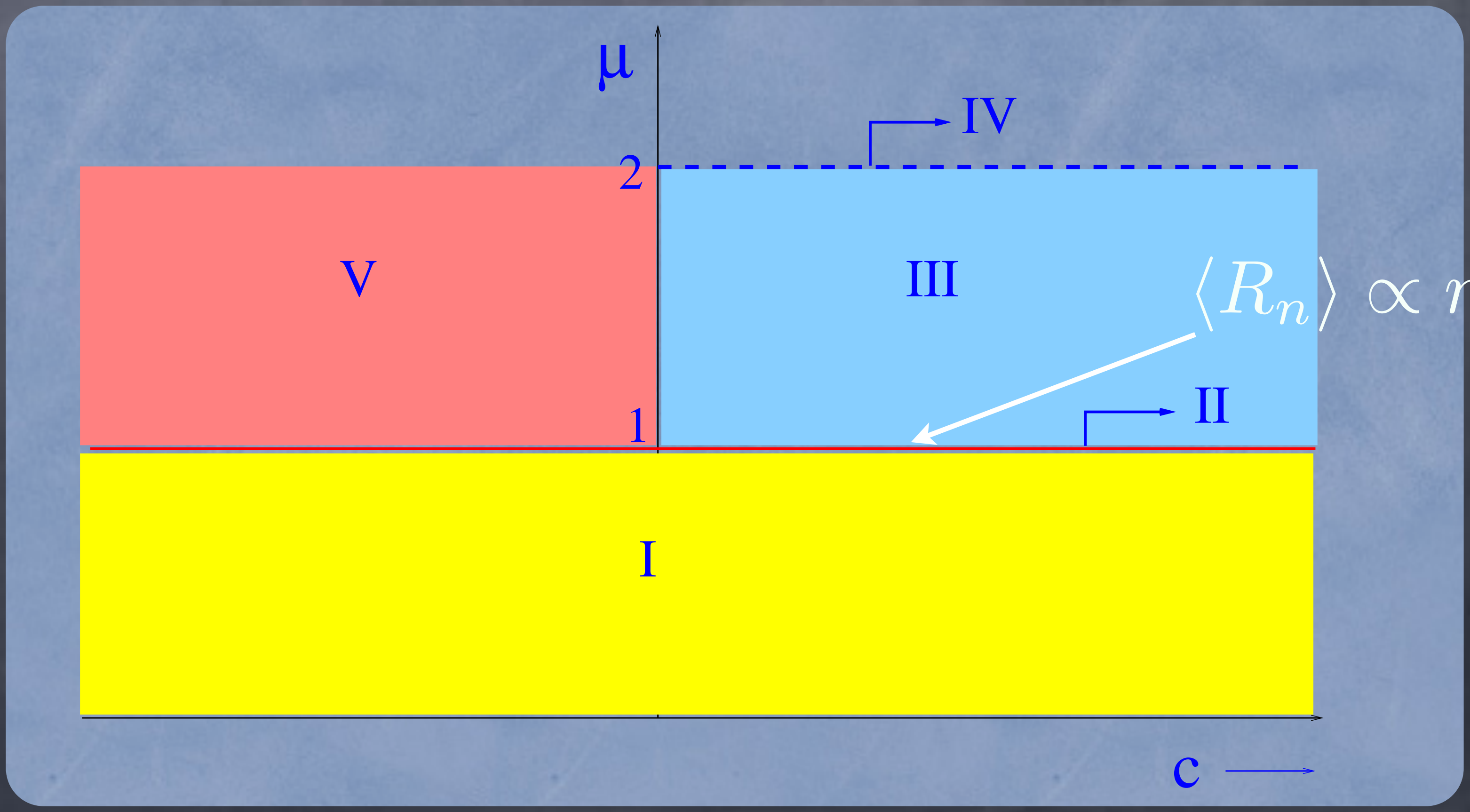
For $n \gg 1$: $\quad P(m,n) \sim \frac{1}{\sqrt{n}} g_0\left(\frac{m}{\sqrt{n}}\right) \; , \; g_0(x) = \frac{1}{\sqrt{\pi}} e^{-\frac{x^2}{4}} \; , \; x > 0$

- RW with a drift   Majumdar, G. S., Wergen '12

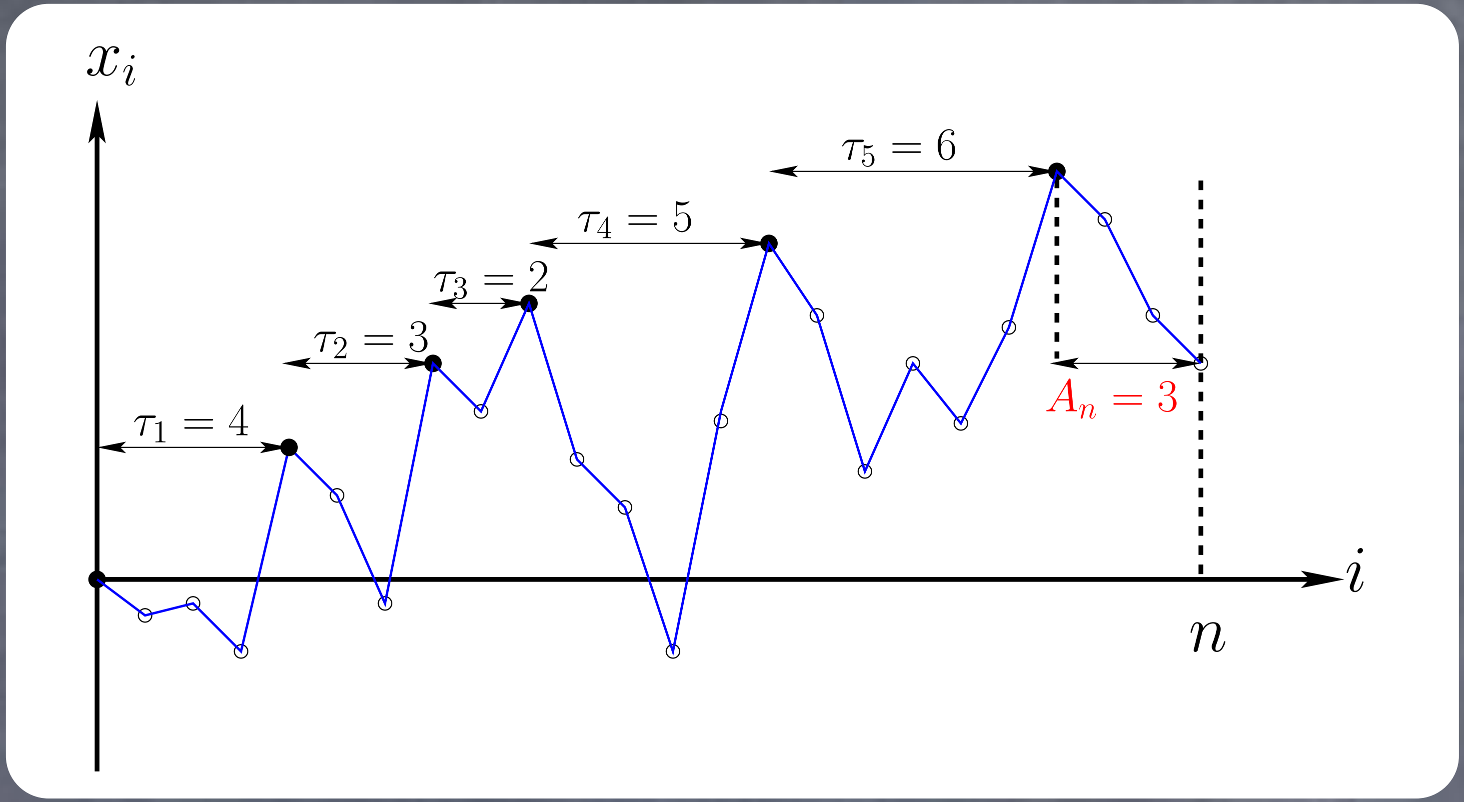
$$P(m,n) = \frac{1}{n^{\theta(c)}} g_c\left(\frac{m}{n^{\theta(c)}}\right)$$

e.g. for $\mu = 1 \; , \; \theta(c) = 1/3$

$$g_c(x) = 3^{2/3} \mathrm{Ai}\left(\frac{x}{3^{1/3}}\right)$$

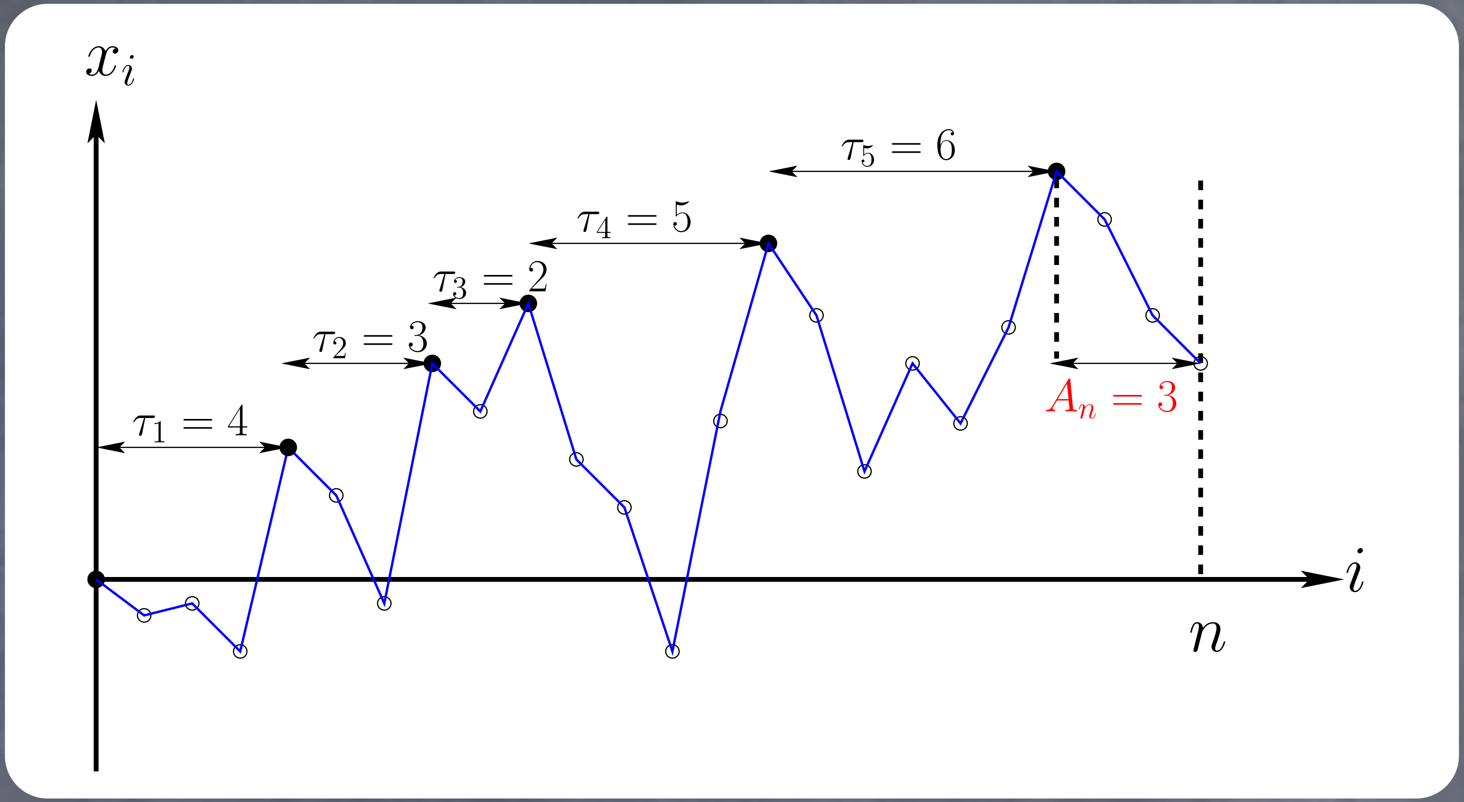


Statistics of the ages of records



sym. RW

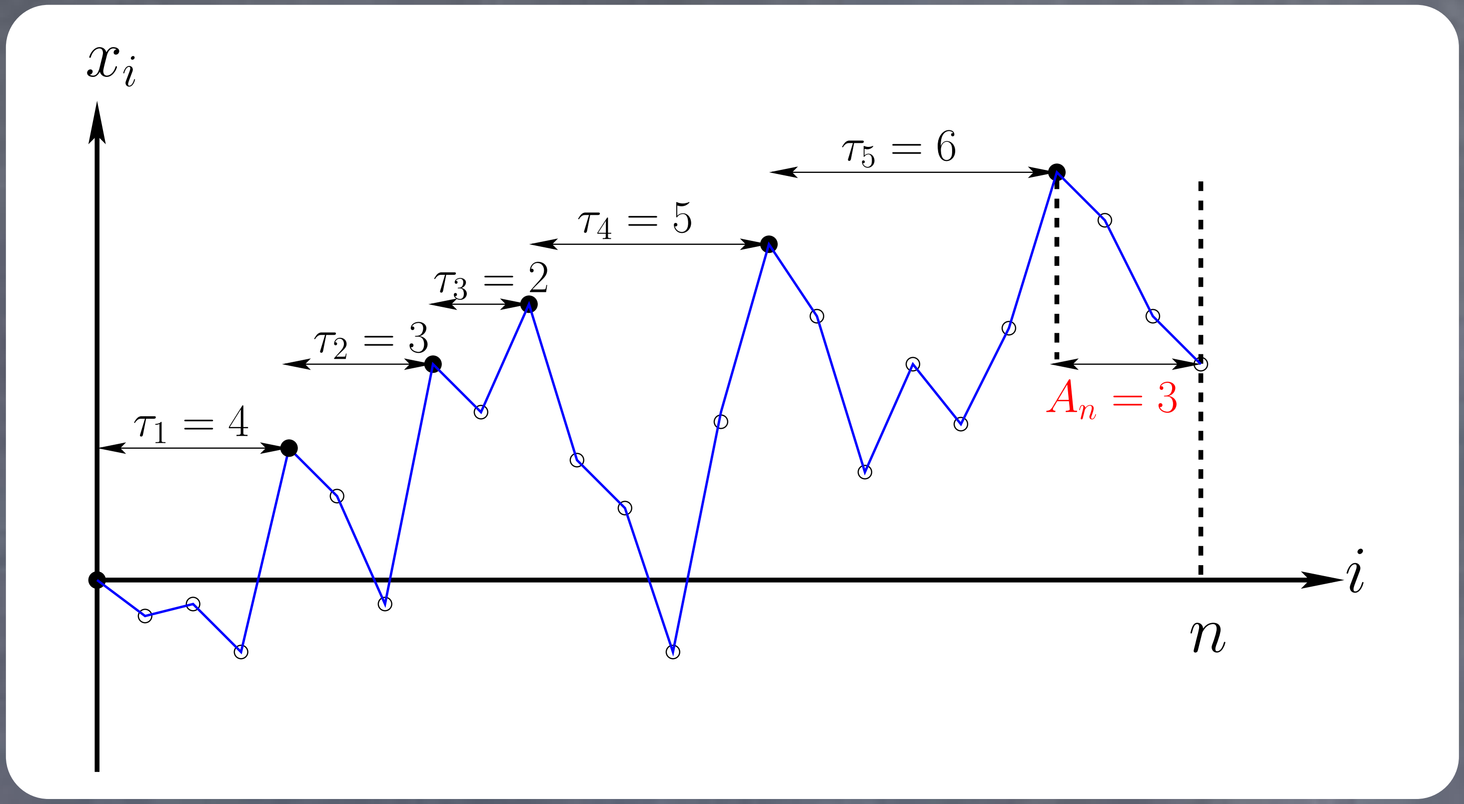
Statistics of the ages of records



sym. RW

- Typical age of a record: $$\ell_{\text{typ}} \sim \frac{n}{\langle R_n \rangle} \sim \sqrt{n}$$

Statistics of the ages of records

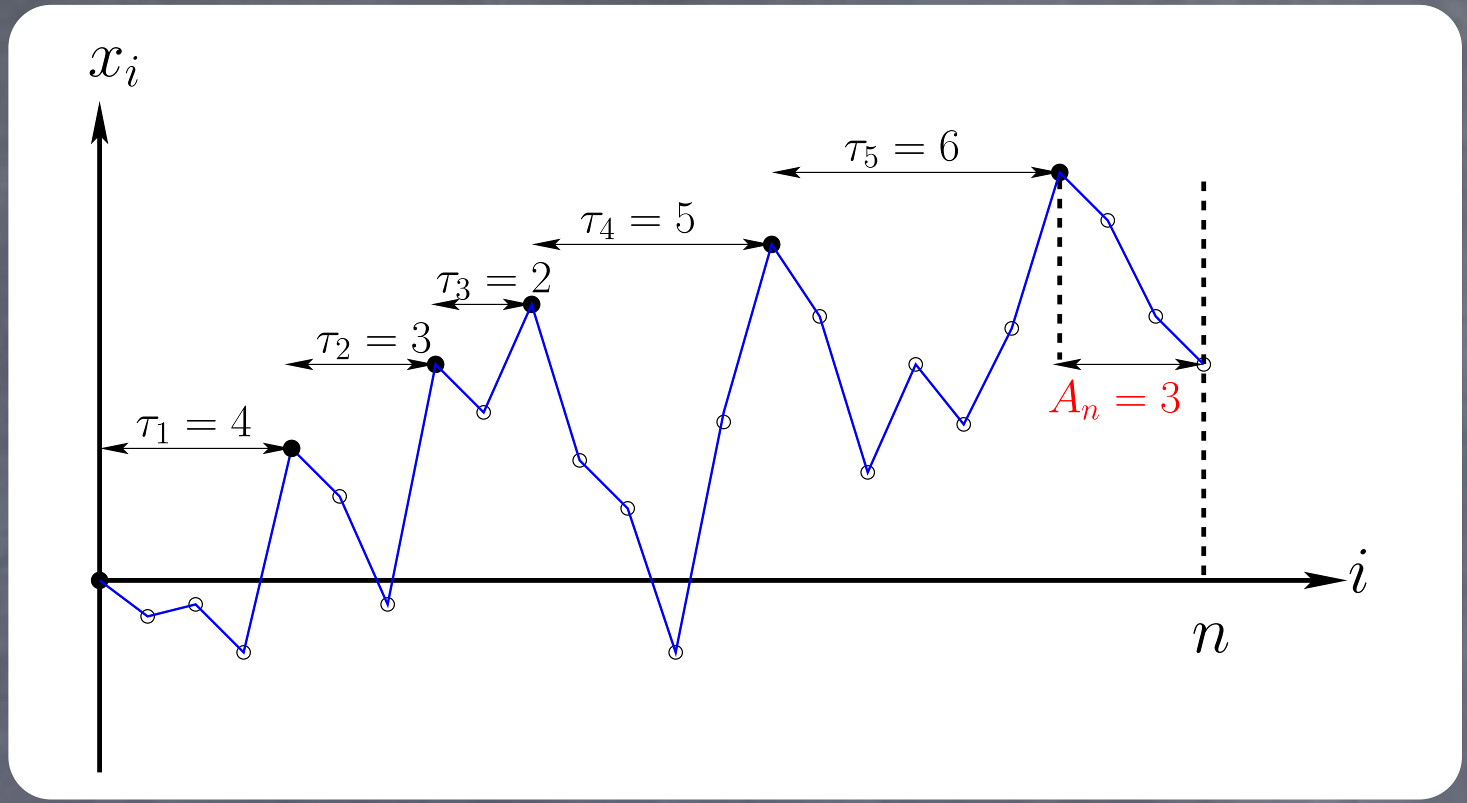


sym. RW

- Typical age of a record: $\ell_{\text{typ}} \sim \dfrac{n}{\langle R_n \rangle} \sim \sqrt{n}$

- What about the longest or shortest age of a record ?

Statistics of the ages of records



sym. RW

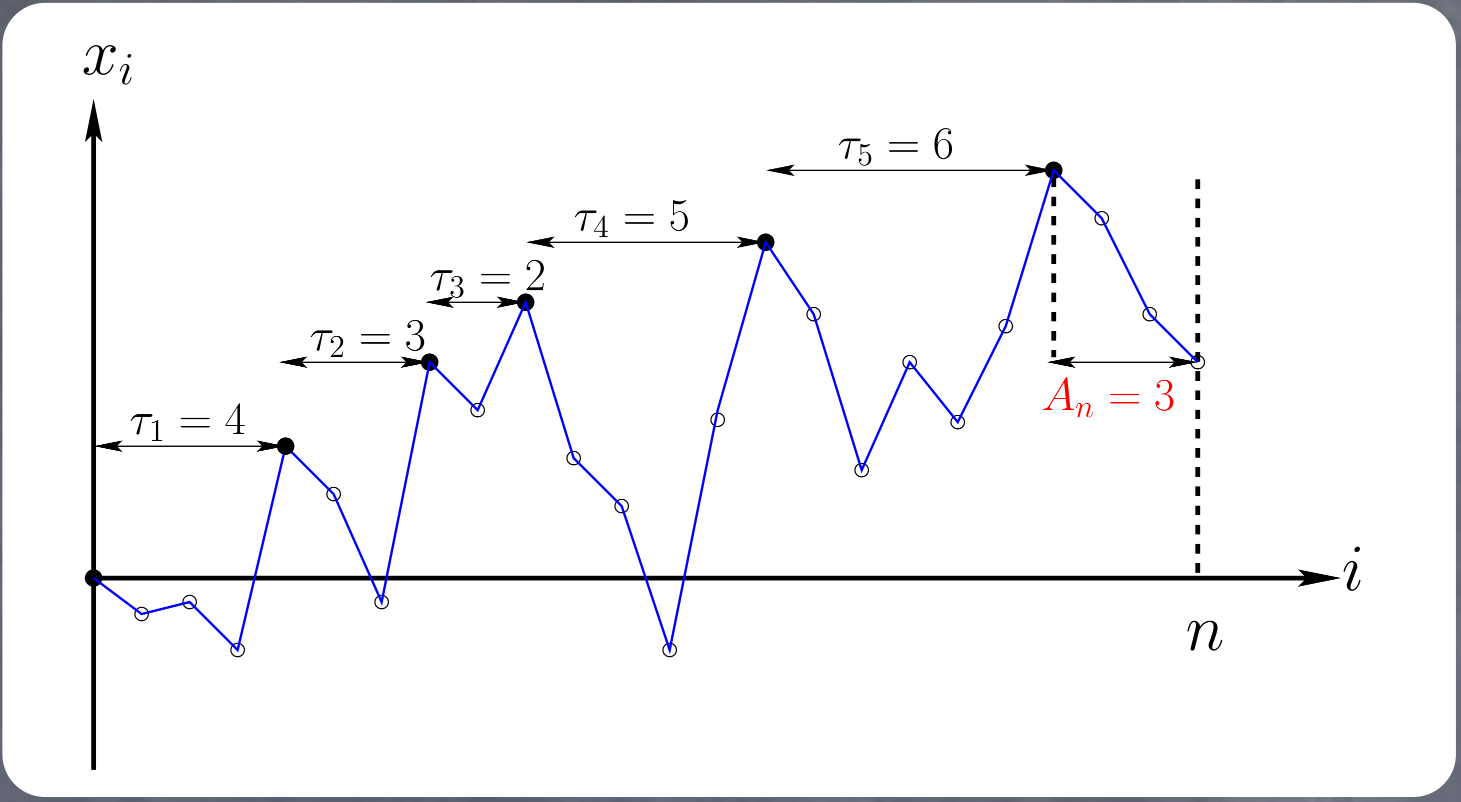
- Typical age of a record: $\ell_{\mathrm{typ}} \sim \dfrac{n}{\langle R_n \rangle} \sim \sqrt{n}$

- What about the longest or shortest age of a record ?

- What is the proba. that the current record is the oldest one ?

Godrèche, Majumdar, G. S., `14

Statistics of the ages of records



sym. RW

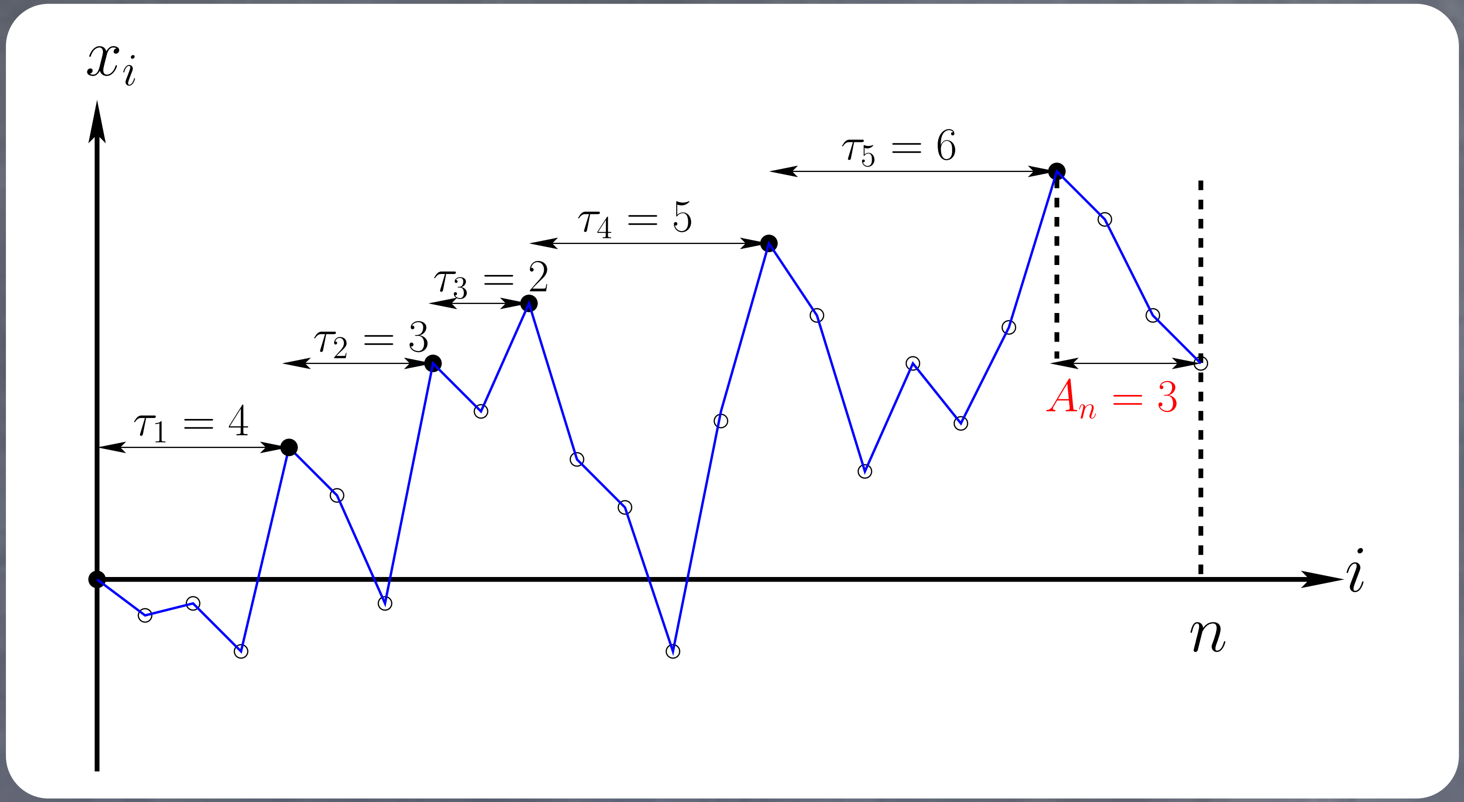
- Typical age of a record: $\ell_{\rm typ} \sim \dfrac{n}{\langle R_n \rangle} \sim \sqrt{n}$

- What about the longest or shortest age of a record ?

- What is the proba. that the current record is the oldest one ?

Godrèche, Majumdar, G. S., `14

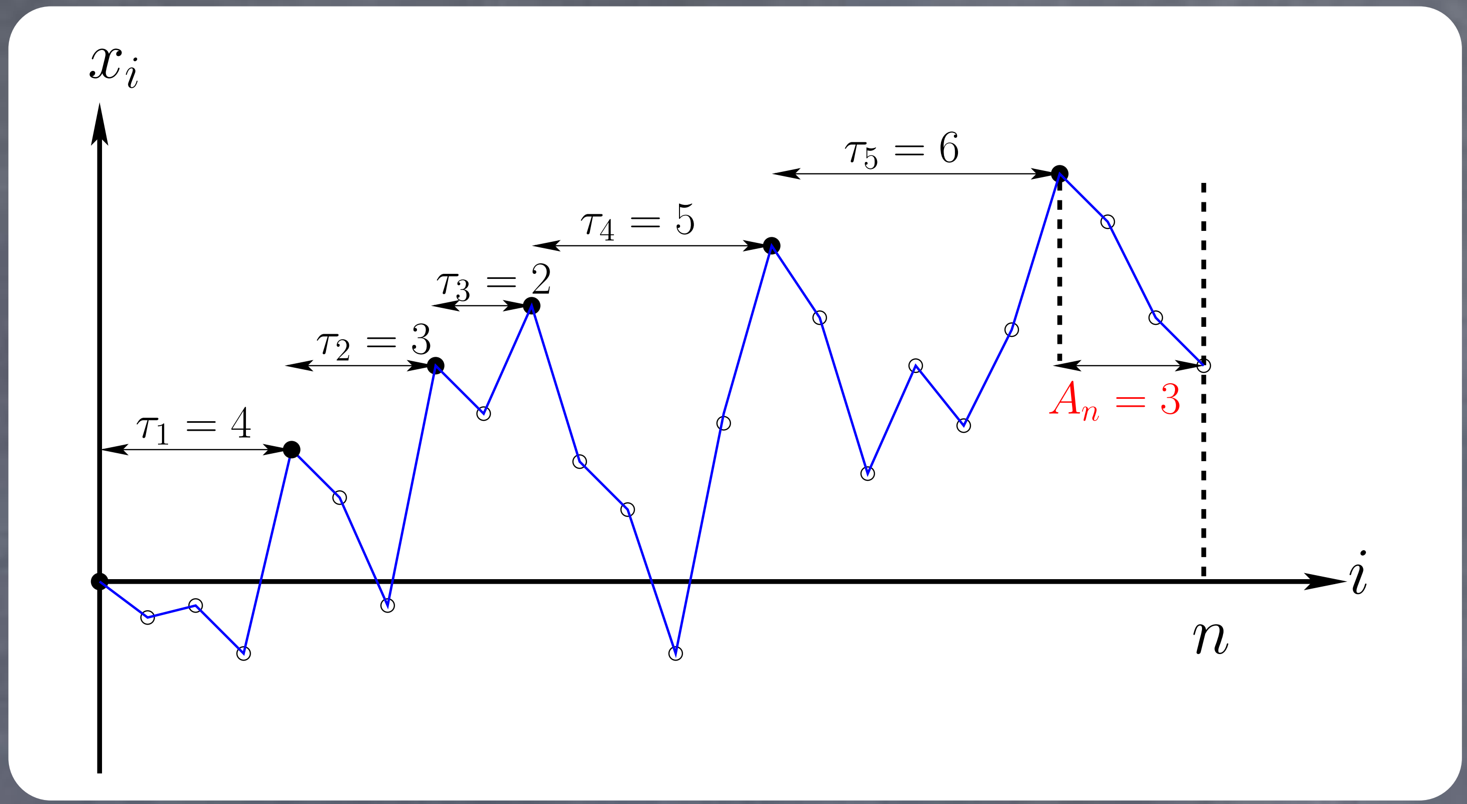
Statistics of the ages of records



sym. RW

$$Q(n) = \Pr\left[A_n \geq \max(\tau_1, \tau_2, \cdots, \tau_{m-1})\right]$$ = ?

Statistics of the ages of records



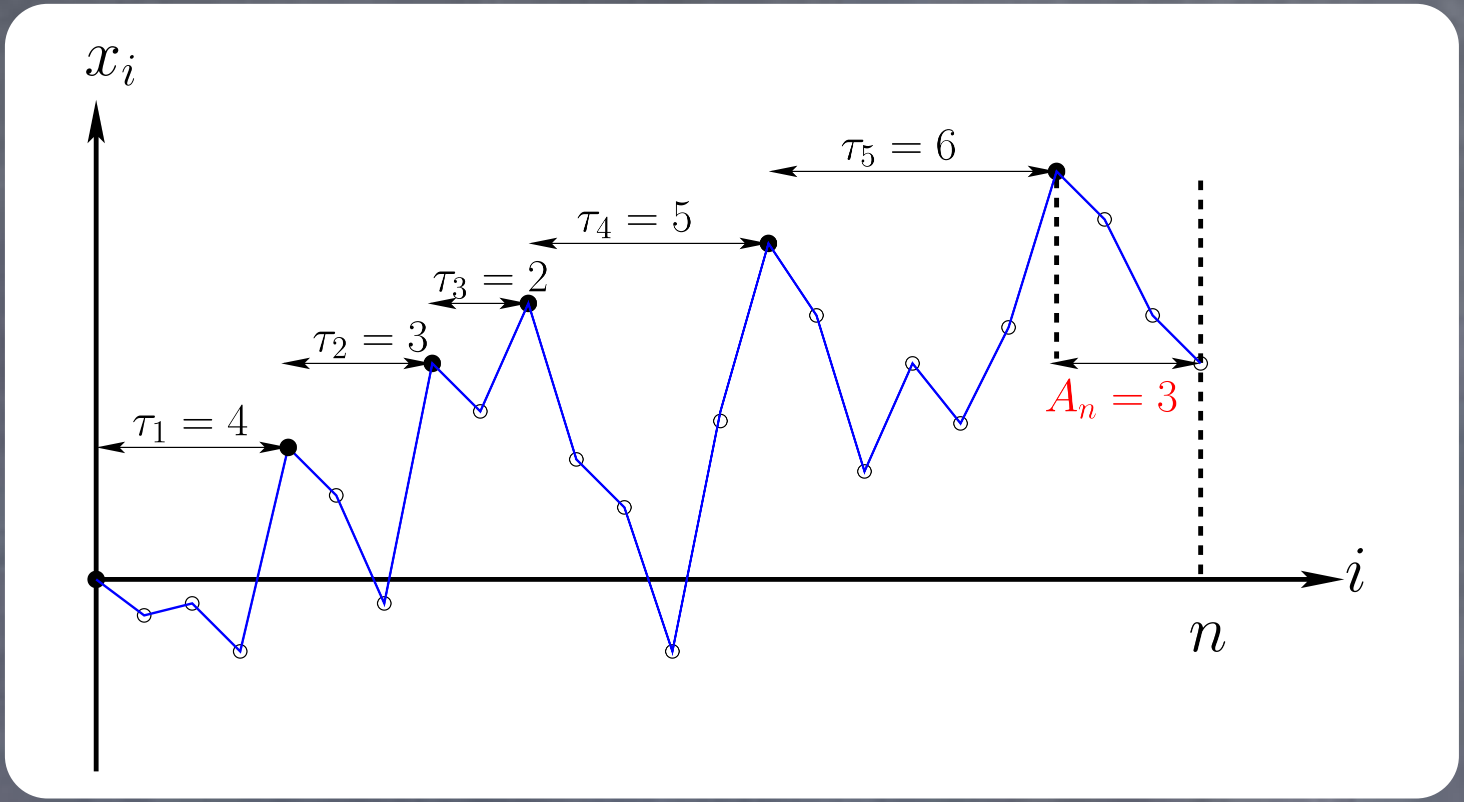
sym. RW

$$Q(n) = \Pr\left[A_n \geq \max(\tau_1, \tau_2, \cdots, \tau_{m-1})\right] = ?$$

$$Q(n) = \sum_{m \geq 1} \Pr\left[A_n \geq \max(\tau_1, \tau_2, \cdots, \tau_{m-1}, R_n = m)\right]$$

Statistics of the ages of records

sym. RW



$$Q(n) = \Pr\left[A_n \geq \max(\tau_1, \tau_2, \cdots, \tau_{m-1})\right] \;\; = \; ?$$

$$Q(n) = \sum_{m \geq 1} \underbrace{\Pr\left[A_n \geq \max(\tau_1, \tau_2, \cdots, \tau_{m-1}, R_n = m)\right]}_{Q(m,n)}$$

Statistics of the ages of records

$$Q(n) = \sum_{m \geq 1} \underbrace{\Pr\left[A_n \geq \max(\tau_1, \tau_2, \cdots, \tau_{m-1}, R_n = m)\right]}_{Q(m,n)}$$

Statistics of the ages of records

$$Q(n) = \sum_{m \geq 1} \underbrace{\Pr\left[A_n \geq \max(\tau_1, \tau_2, \cdots, \tau_{m-1}, R_n = m)\right]}_{Q(m,n)}$$

$$Q(m,n) = \sum_{a \geq 0} \sum_{\ell_1 = 1}^{a} \cdots \sum_{\ell_{m-1}=1}^{a} \Pr(R_n = m, \tau_1 = \ell_1, \cdots, \tau_{m-1} = \ell_{m-1}, A_n = a)$$

Statistics of the ages of records

$$Q(n) = \sum_{m \geq 1} \underbrace{\Pr\left[A_n \geq \max(\tau_1, \tau_2, \cdots, \tau_{m-1}, R_n = m)\right]}_{Q(m,n)}$$

$$Q(m,n) = \sum_{a \geq 0} \sum_{\ell_1 = 1}^{a} \cdots \sum_{\ell_{m-1}=1}^{a} \underbrace{\Pr(R_n = m, \tau_1 = \ell_1, \cdots, \tau_{m-1} = \ell_{m-1}, A_n = a)}_{P(\vec{\ell}, m, n)}$$

Statistics of the ages of records

$$Q(n) = \sum_{m \geq 1} \underbrace{\Pr\left[A_n \geq \max(\tau_1, \tau_2, \cdots, \tau_{m-1}, R_n = m)\right]}_{Q(m,n)}$$

$$Q(m,n) = \sum_{a \geq 0} \sum_{\ell_1 = 1}^{a} \cdots \sum_{\ell_{m-1}=1}^{a} \underbrace{\Pr(R_n = m, \tau_1 = \ell_1, \cdots, \tau_{m-1} = \ell_{m-1}, A_n = a)}_{P(\vec{\ell}, m, n)}$$

with $P(\vec{l}, m, n) = f_-(\ell_1) f_-(\ell_2) \cdots f_-(\ell_{m-1}) q_-(a) \delta\left(\sum_{k=1}^{m-1} \ell_k + a, n\right)$

Statistics of the ages of records

$$Q(n) = \sum_{m \geq 1} \underbrace{\Pr\left[A_n \geq \max(\tau_1, \tau_2, \cdots, \tau_{m-1}, R_n = m)\right]}_{Q(m,n)}$$

$$Q(m,n) = \sum_{a \geq 0} \sum_{\ell_1 = 1}^{a} \cdots \sum_{\ell_{m-1}=1}^{a} \underbrace{\Pr(R_n = m, \tau_1 = \ell_1, \cdots, \tau_{m-1} = \ell_{m-1}, A_n = a)}_{P(\vec{\ell}, m, n)}$$

with $P(\vec{l}, m, n) = f_-(\ell_1) f_-(\ell_2) \cdots f_-(\ell_{m-1}) q_-(a) \delta\left(\sum_{k=1}^{m-1} \ell_k + a, n\right)$

Generating function

$$\sum_{n \geq 0} z^n Q(m,n) = \sum_{a \geq 0} \left(\sum_{\ell=1}^{a} f_-(\ell) z^\ell\right)^{m-1} q_-(a) z^a$$

Statistics of the ages of records

$$Q(n) = \sum_{m \geq 1} \underbrace{\Pr\left[A_n \geq \max(\tau_1, \tau_2, \cdots, \tau_{m-1}, R_n = m)\right]}_{Q(m,n)}$$

$$Q(m,n) = \sum_{a \geq 0} \sum_{\ell_1 = 1}^{a} \cdots \sum_{\ell_{m-1}=1}^{a} \underbrace{\Pr(R_n = m, \tau_1 = \ell_1, \cdots, \tau_{m-1} = \ell_{m-1}, A_n = a)}_{P(\vec{\ell}, m, n)}$$

with $P(\vec{l}, m, n) = f_-(\ell_1) f_-(\ell_2) \cdots f_-(\ell_{m-1}) q_-(a) \delta\left(\sum_{k=1}^{m-1} \ell_k + a, n\right)$

Generating function

$$\sum_{n \geq 0} z^n Q(m,n) = \sum_{a \geq 0} \left(\sum_{\ell=1}^{a} f_-(\ell) z^\ell\right)^{m-1} q_-(a) z^a$$

$$\sum_{n \geq 0} z^n Q(n) = \sum_{n \geq 1} z^n \sum_{m \geq 1} Q(m,n) = \sum_{a \geq 0} \sum_{m \geq 1} \left(\sum_{\ell=1}^{a} f_-(\ell) z^\ell\right)^{m-1} q_-(a) z^a$$

Statistics of the ages of records

$$Q(n) = \Pr\left[A_n \geq \max(\tau_1, \tau_2, \cdots, \tau_{m-1})\right]$$

$$\sum_{n\geq 0} z^n Q(n) = \sum_{a\geq 0} \frac{z^a q_-(a)}{1 - \sum_{\ell=1}^{a} f_-(\ell) z^\ell}$$

Statistics of the ages of records

$$Q(n) = \Pr\left[A_n \geq \max(\tau_1, \tau_2, \cdots, \tau_{m-1})\right]$$

$$\sum_{n\geq 0} z^n Q(n) = \sum_{a\geq 0} \frac{z^a q_-(a)}{1 - \sum_{\ell=1}^{a} f_-(\ell) z^\ell}$$

For symmetric RW

$$q_-(k) = \frac{1}{2^{2k}}\binom{2k}{k}, \; f_-(k) = q_-(k) - q_-(k-1)$$

Statistics of the ages of records

$$Q(n) = \Pr\left[A_n \geq \max(\tau_1, \tau_2, \cdots, \tau_{m-1})\right]$$

$$\sum_{n\geq 0} z^n Q(n) = \sum_{a\geq 0} \frac{z^a q_-(a)}{1 - \sum_{\ell=1}^{a} f_-(\ell) z^\ell}$$

For symmetric RW

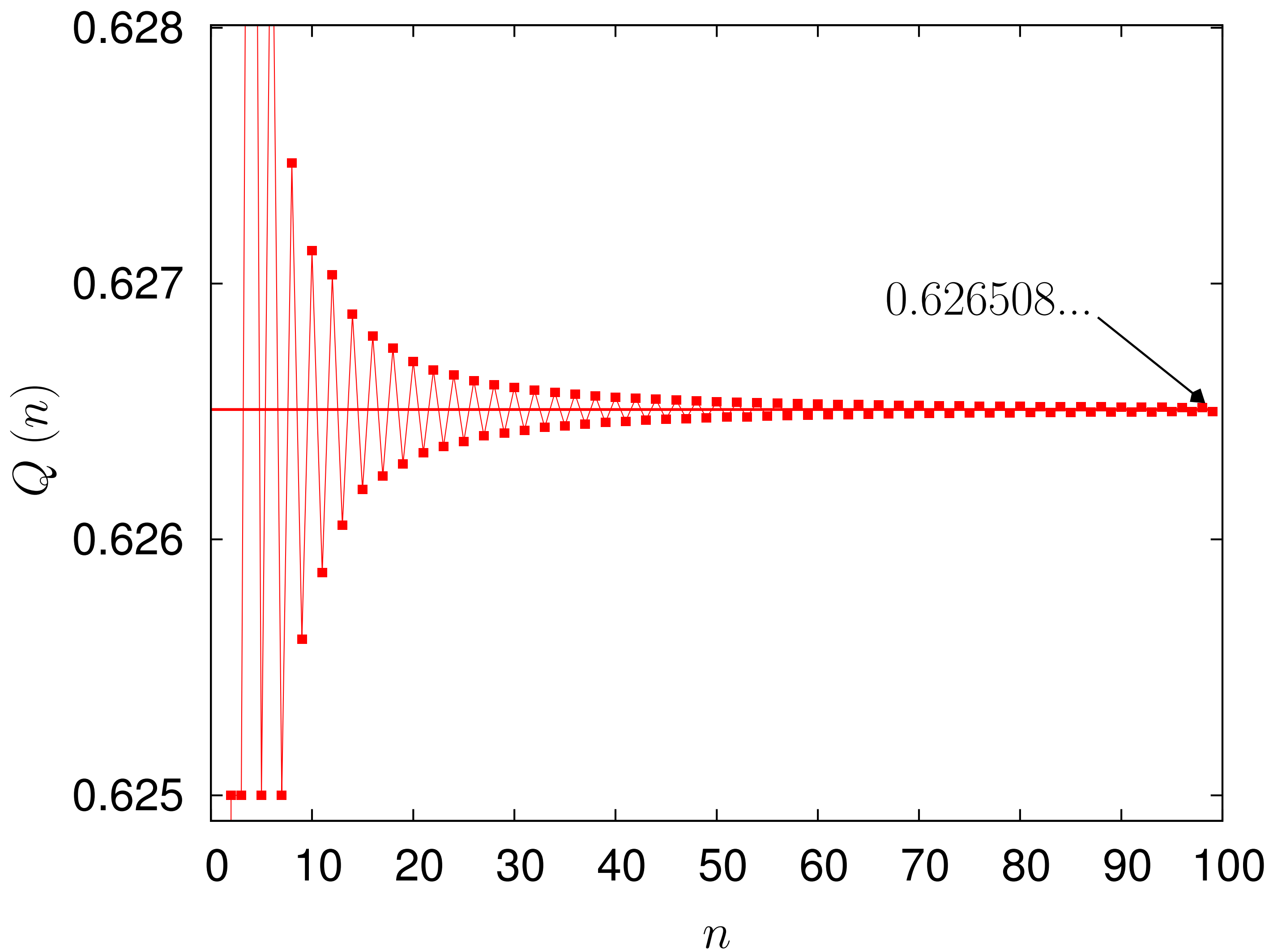
$$q_-(k) = \frac{1}{2^{2k}}\binom{2k}{k}\,,\ f_-(k) = q_-(k) - q_-(k-1)$$

One finds

$$\sum_{n\geq 0} z^n Q(n) = 1 + \frac{1}{2}z + \frac{5}{8}z^2 + \frac{5}{8}z^3 + \frac{81}{128}z^4 + \frac{5}{8}z^5 + \frac{161}{256}z^6 + \cdots$$

Statistics of the ages of records

$$Q(n) = \Pr\left[A_n \geq \max(\tau_1, \tau_2, \cdots, \tau_{m-1})\right]$$

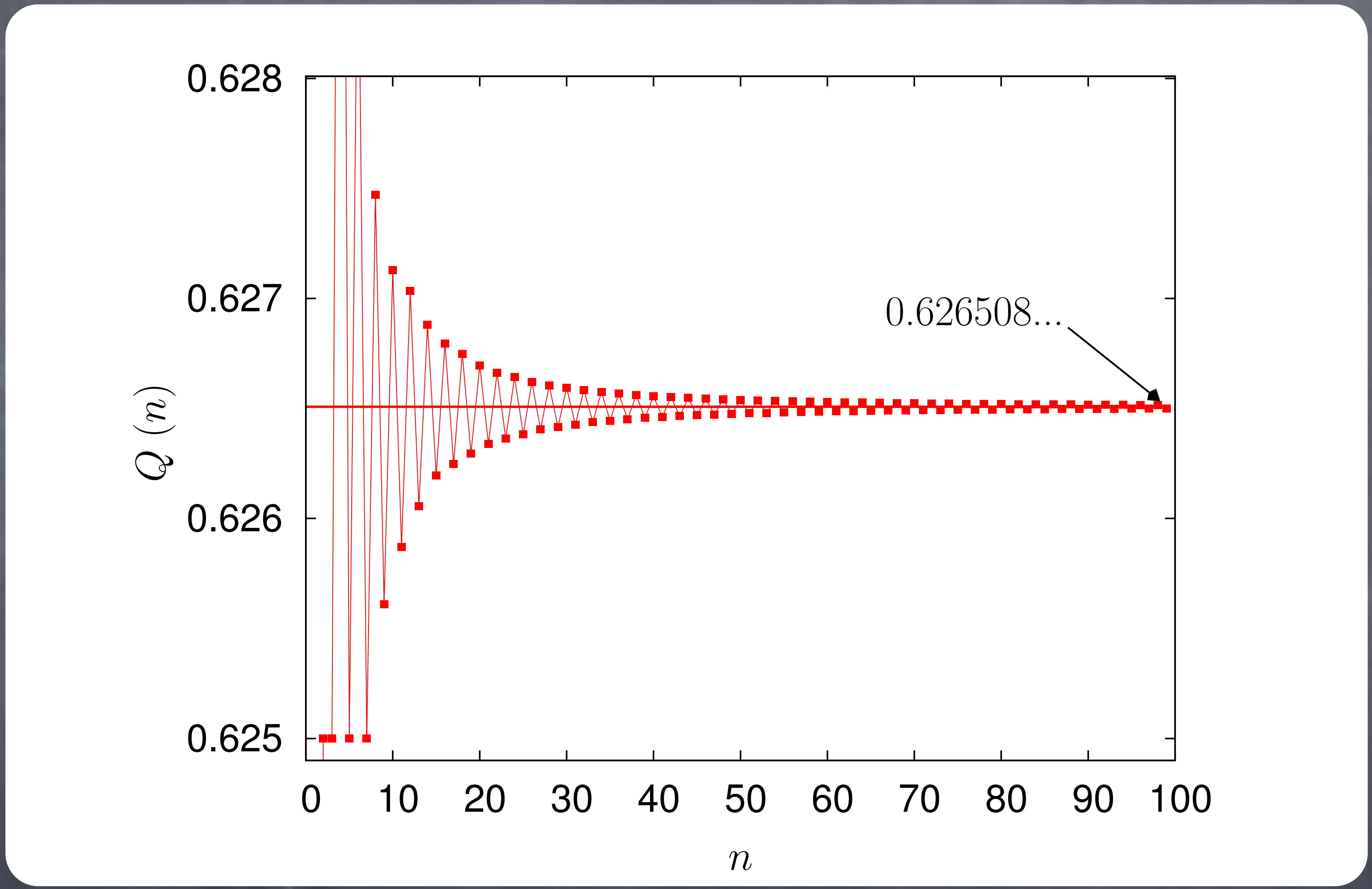


Godrèche, Majumdar, G. S.,`14

$$\lim_{n\to\infty} Q(n) = Q_\infty$$

Statistics of the ages of records

$$Q(n) = \Pr\left[A_n \geq \max(\tau_1, \tau_2, \cdots, \tau_{m-1})\right]$$



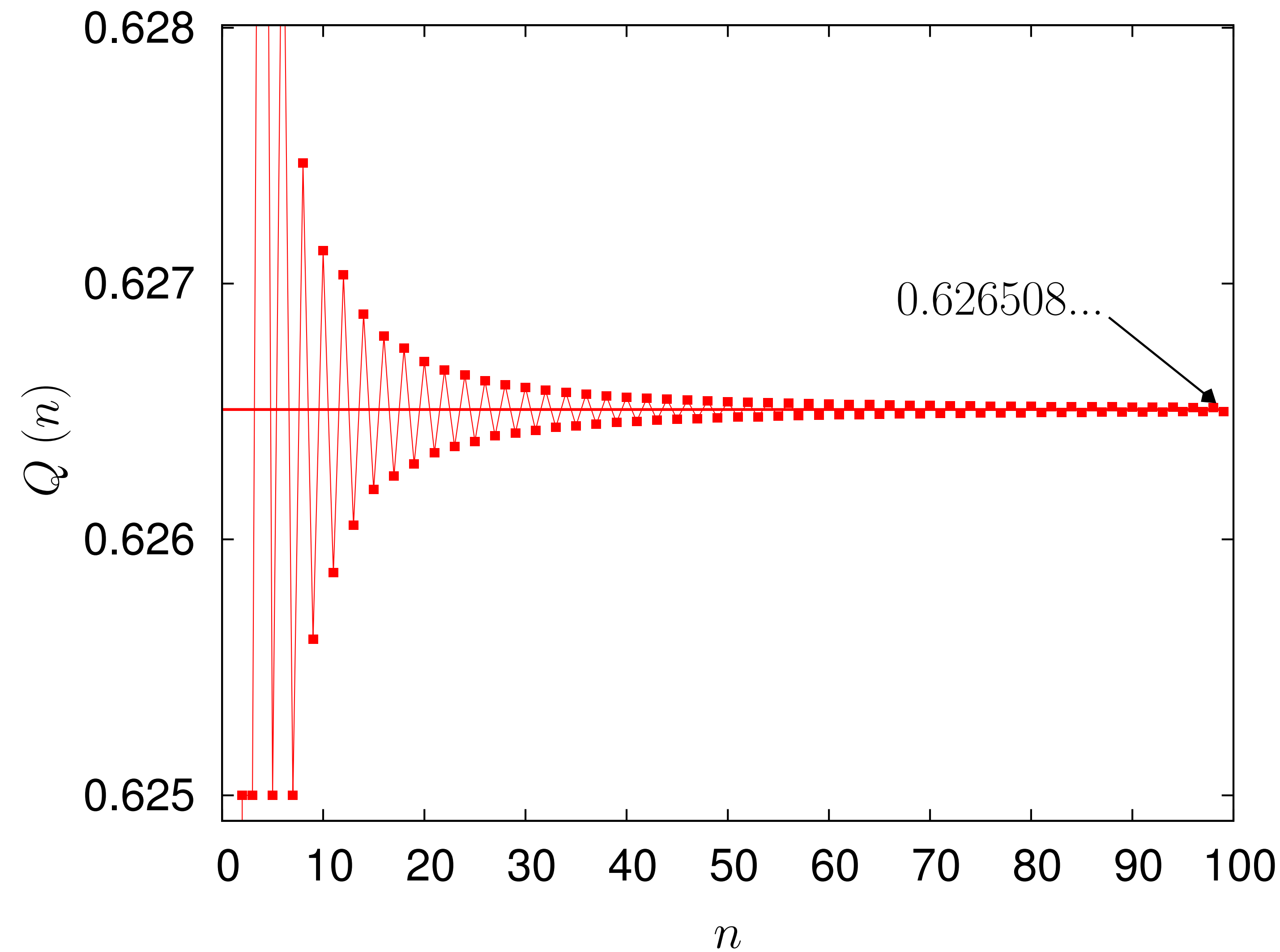
Godrèche, Majumdar, G. S.,`14

$$\lim_{n\to\infty} Q(n) = Q_\infty$$

$$Q_\infty = \int_0^\infty \mathrm{d}x \, \frac{1}{1+\sqrt{\pi x}\,\mathrm{e}^x \mathrm{erf}\sqrt{x}} = 0.626508\ldots$$

Statistics of the ages of records

$$Q(n) = \Pr\left[A_n \geq \max(\tau_1, \tau_2, \cdots, \tau_{m-1})\right]$$



Godrèche, Majumdar, G. S.,`14

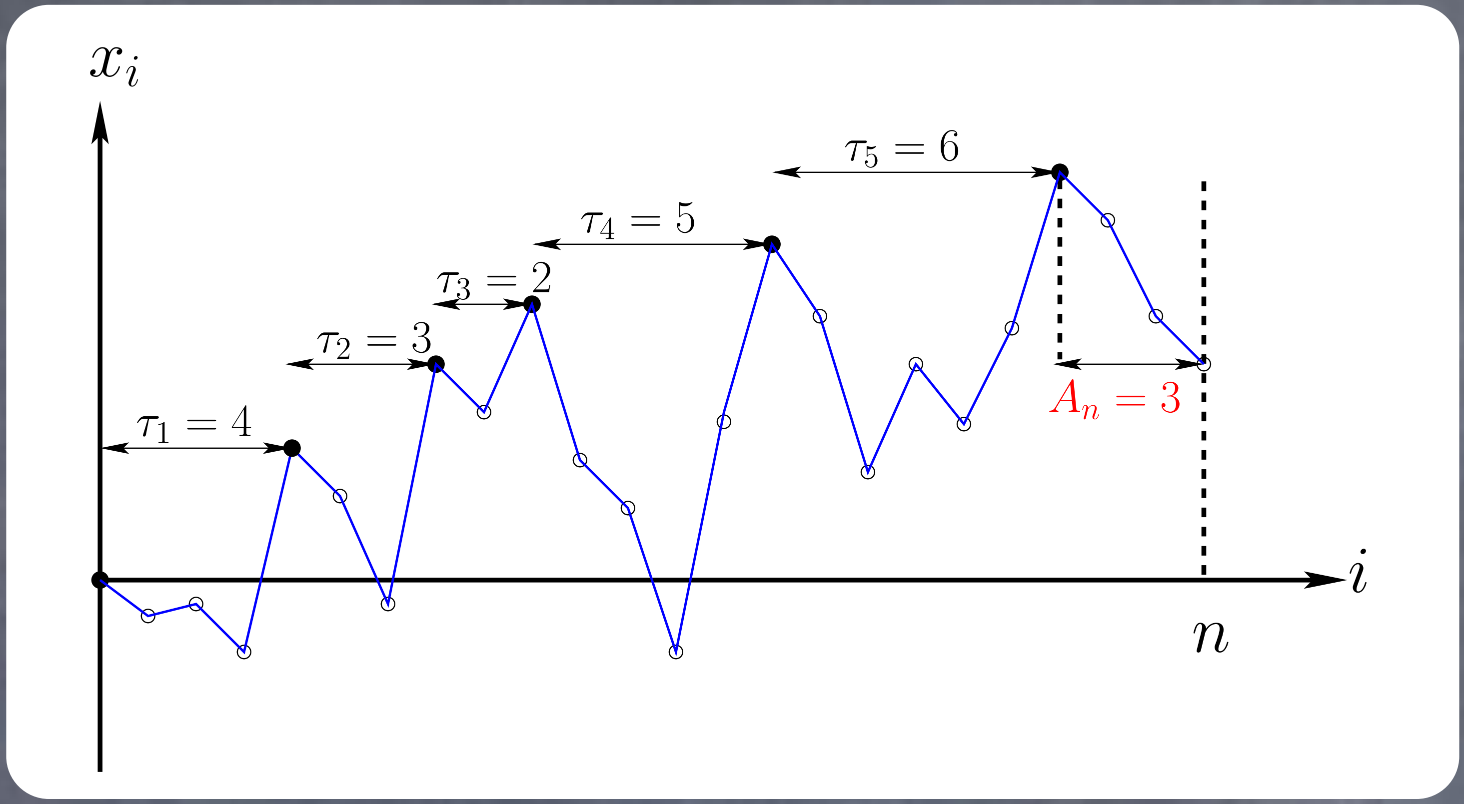
$$\lim_{n\to\infty} Q(n) = Q_\infty$$

Pitman, Yor, `97

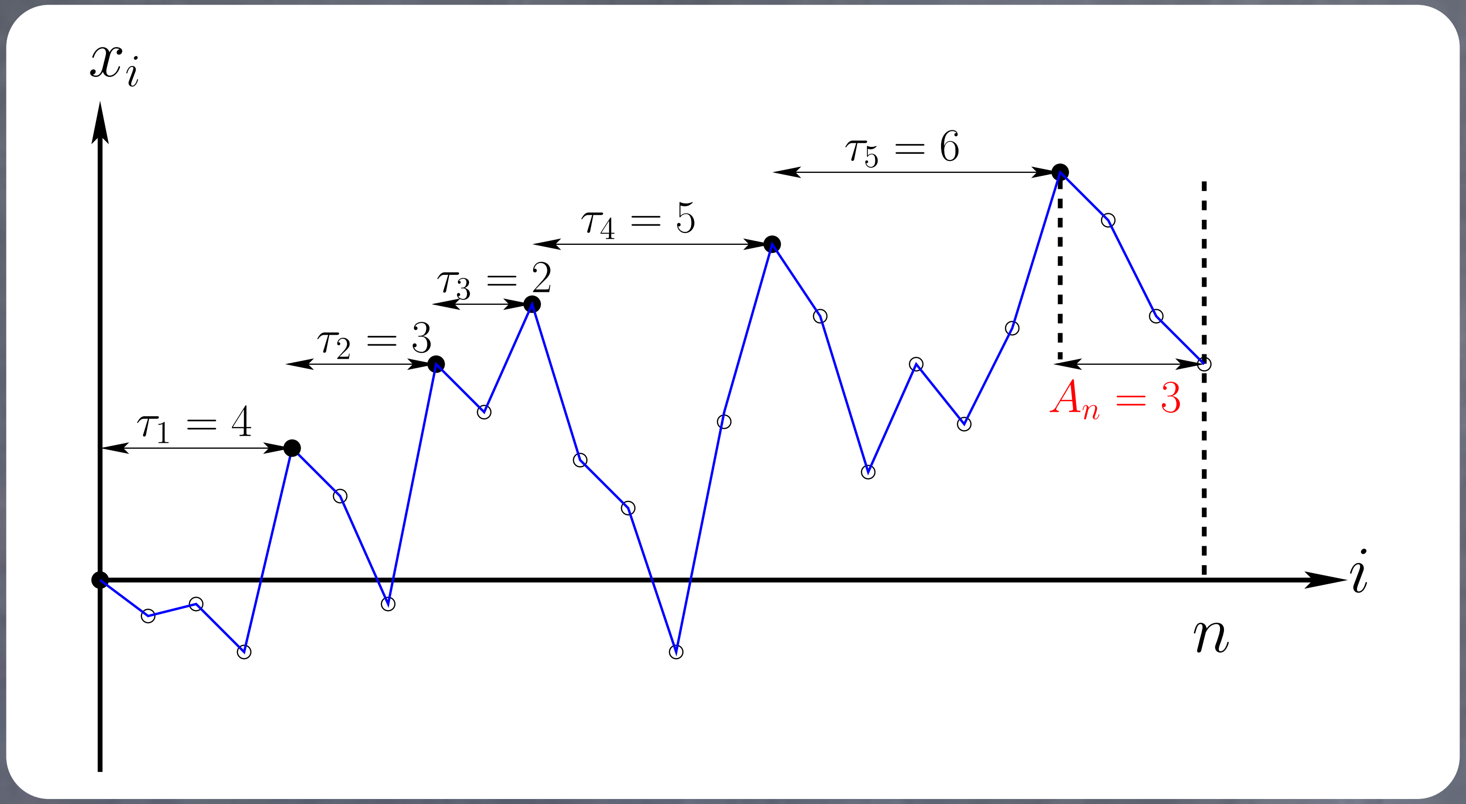
Godrèche, Majumdar, G. S., `09

$$Q_\infty = \int_0^\infty \mathrm{d}x \, \frac{1}{1+\sqrt{\pi x}\, \mathrm{e}^x \operatorname{erf}\sqrt{x}} = 0.626508\ldots$$

New observable...new universal constant

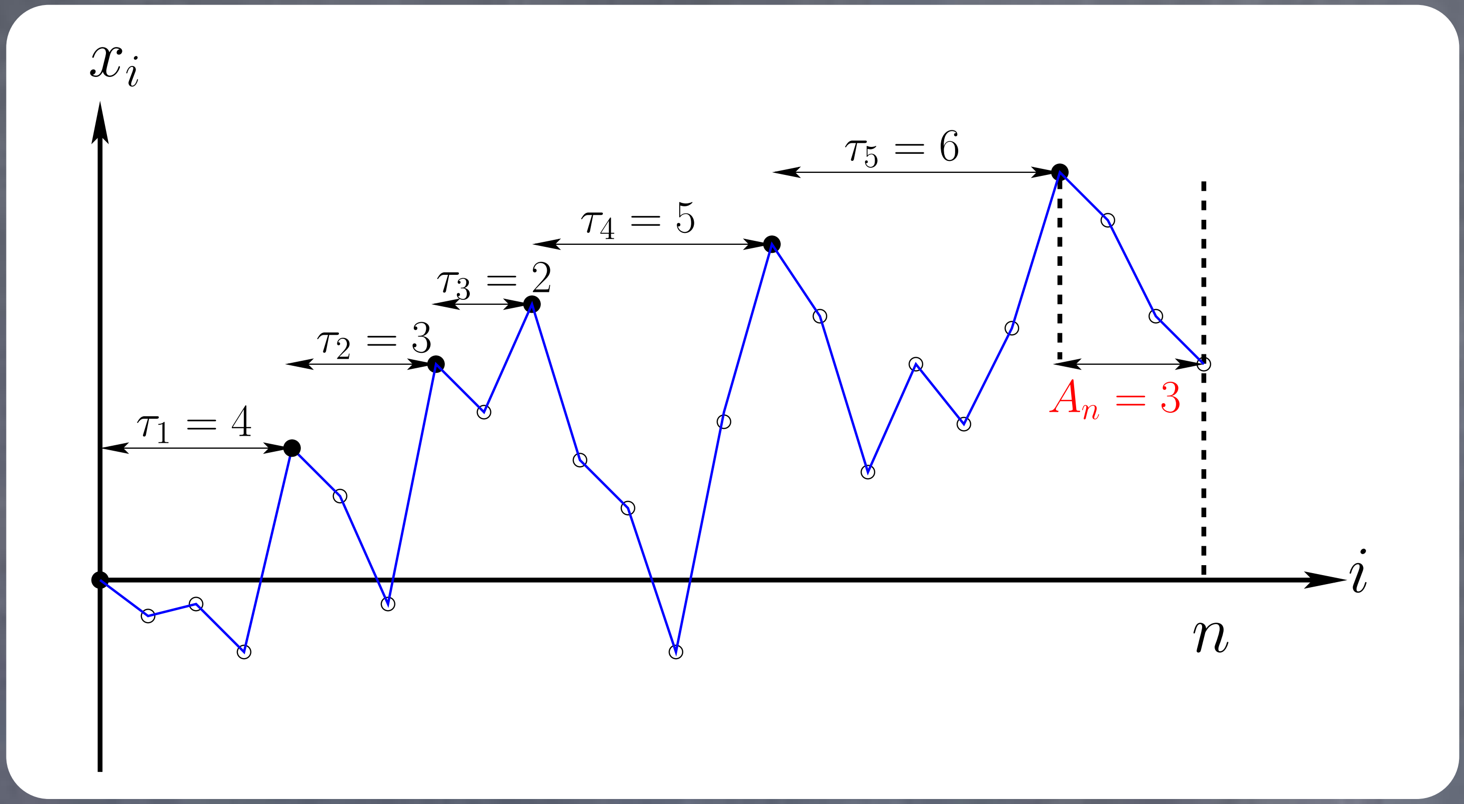


New observable...new universal constant



$$Q_1(n) = \Pr[\tau_1 \geq \max(\tau_2, \cdots, \tau_{m-1}, A_n)]$$

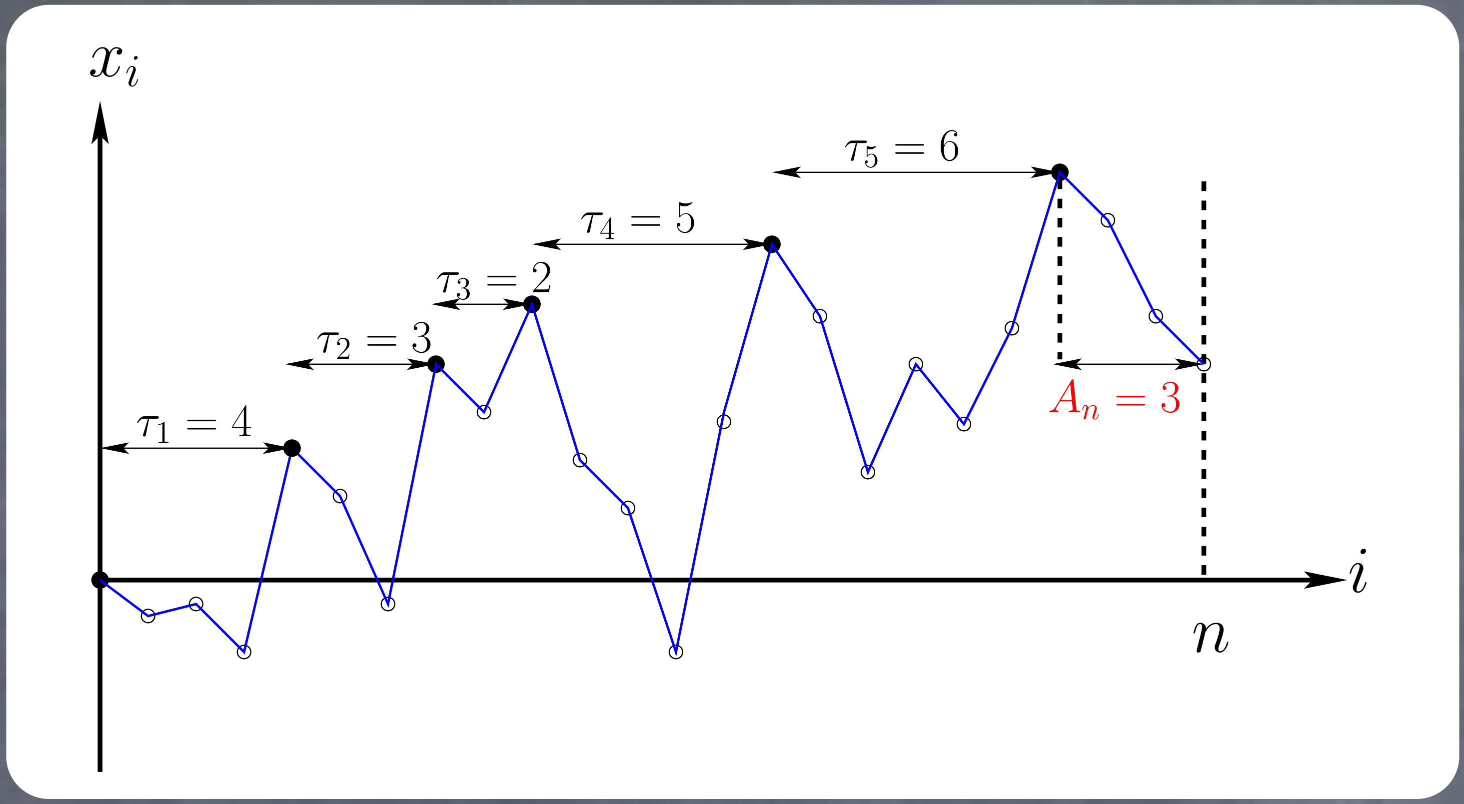
New observable...new universal constant



$$Q_1(n) = \Pr[\tau_1 \geq \max(\tau_2, \cdots, \tau_{m-1}, A_n)]$$

$$\sim \frac{C_1}{\sqrt{n}}$$

New observable...new universal constant



$$Q_1(n) = \Pr[\tau_1 \geq \max(\tau_2, \cdots, \tau_{m-1}, A_n)]$$

$$\sim \frac{C_1}{\sqrt{n}}$$

$$C_1 = \frac{1}{\sqrt{\pi}}\left(1 + \frac{1}{2}\int_0^\infty \frac{\mathrm{d}x}{x}\,\frac{\mathrm{erf}(\sqrt{x})}{1+\sqrt{\pi x}\,\mathrm{e}^x\,\mathrm{erf}(\sqrt{x})}\right) = 0.962641\ldots$$

Godrèche, Majumdar, G. S.,`14

Conclusions

- Exact results for records of strongly correlated time series

  see arXiv:1305.0639 for a short review

- Universal records statistics for (symmetric) RWs

- Extension to multiparticle systems Wergen, Majumdar, G. S. `12

- Extension to Continuous Time Random Walks (CTRWs)

  S. Sabhapandit `12

Conclusions

- Exact results for records of strongly correlated time series

  see arXiv:1305.0639 for a short review

- Universal records statistics for (symmetric) RWs

- Extension to multiparticle systems Wergen, Majumdar, G. S. `12

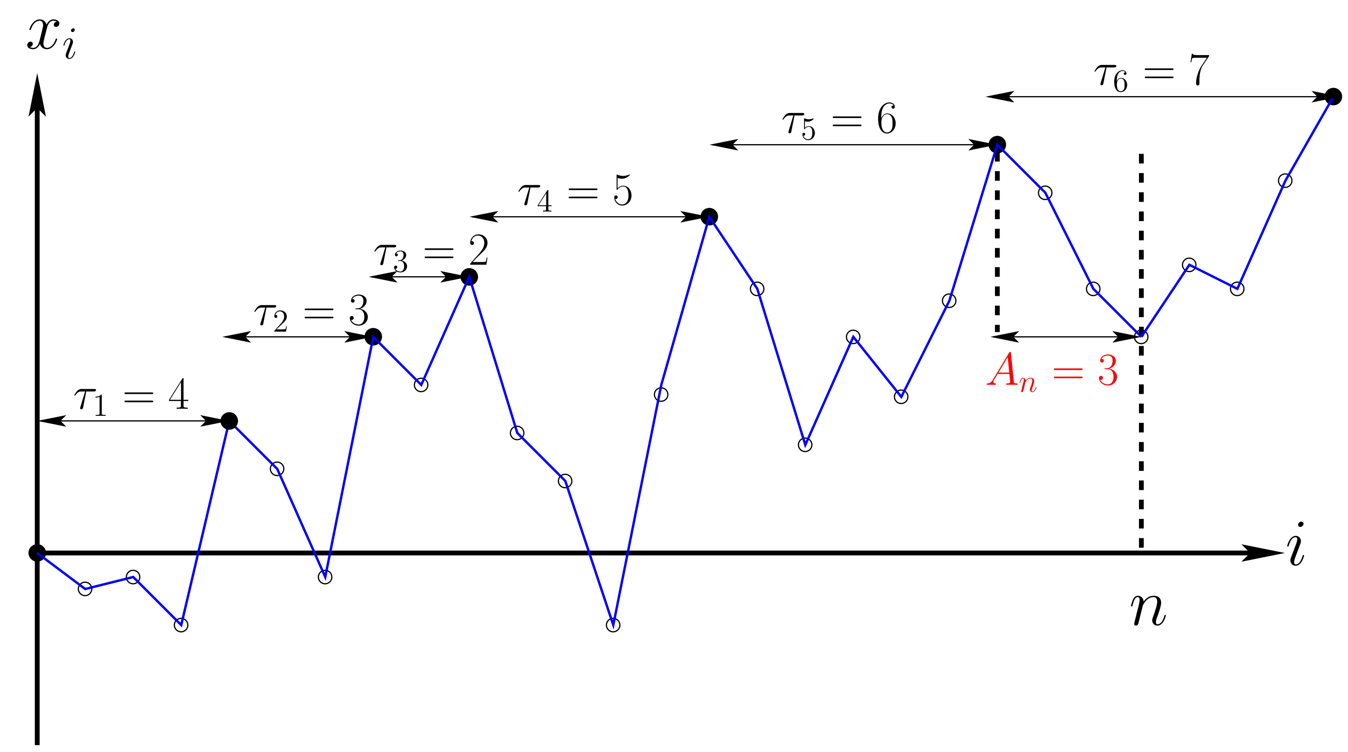
- Extension to Continuous Time Random Walks (CTRWs)

  S. Sabhapandit `12

- High sensitivity to the definition of the age of the last record

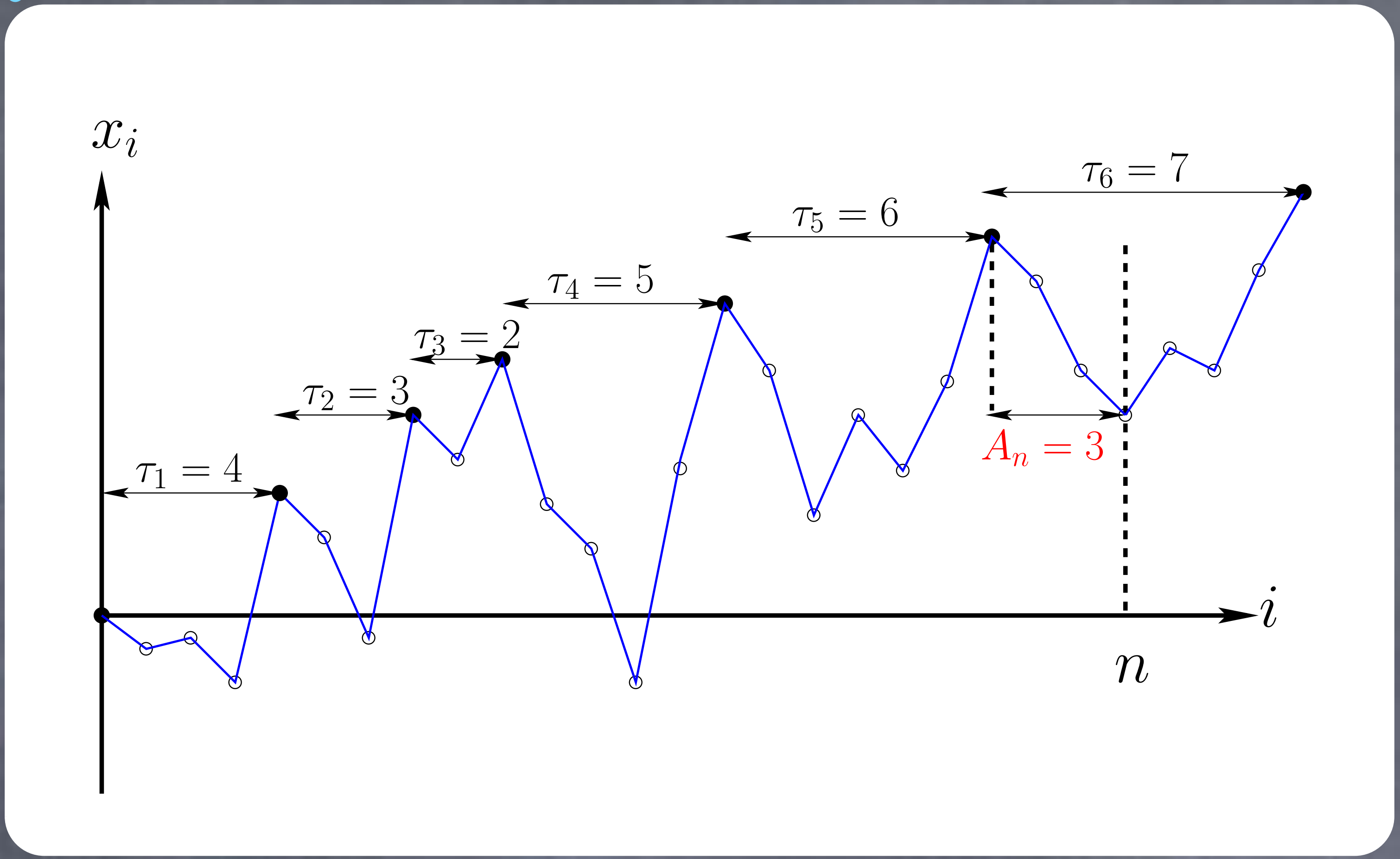
Sensitivity to the definition of the age of the last record

Godrèche, Majumdar, G. S., `14



Sensitivity to the definition of the age of the last record

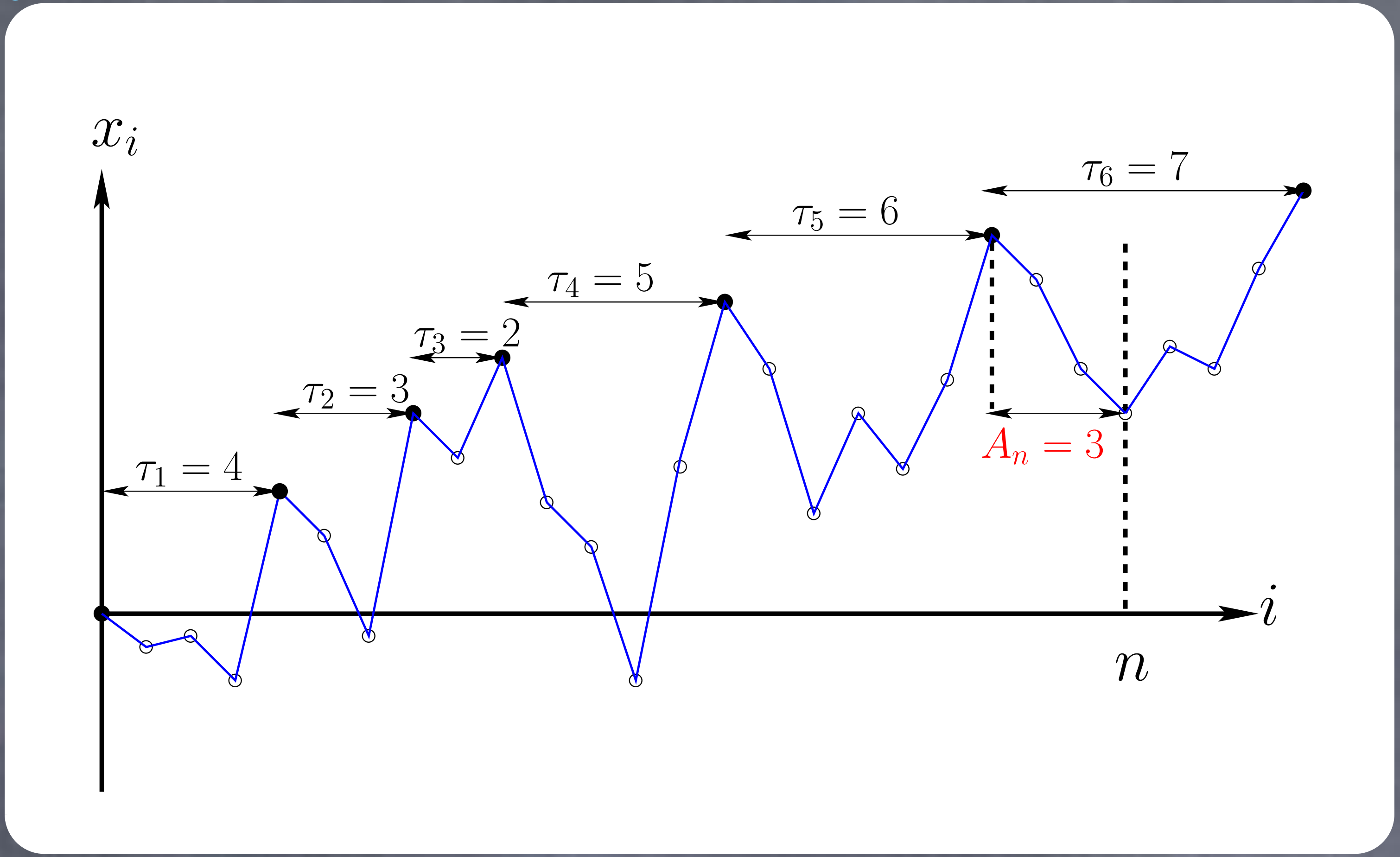
Godrèche, Majumdar, G. S., `14



$$Q^{\mathrm{II}}(n) = \Pr[\tau_m \geq \max(\tau_1, \cdots, \tau_{m-1})]$$

Sensitivity to the definition of the age of the last record

Godrèche, Majumdar, G. S., `14

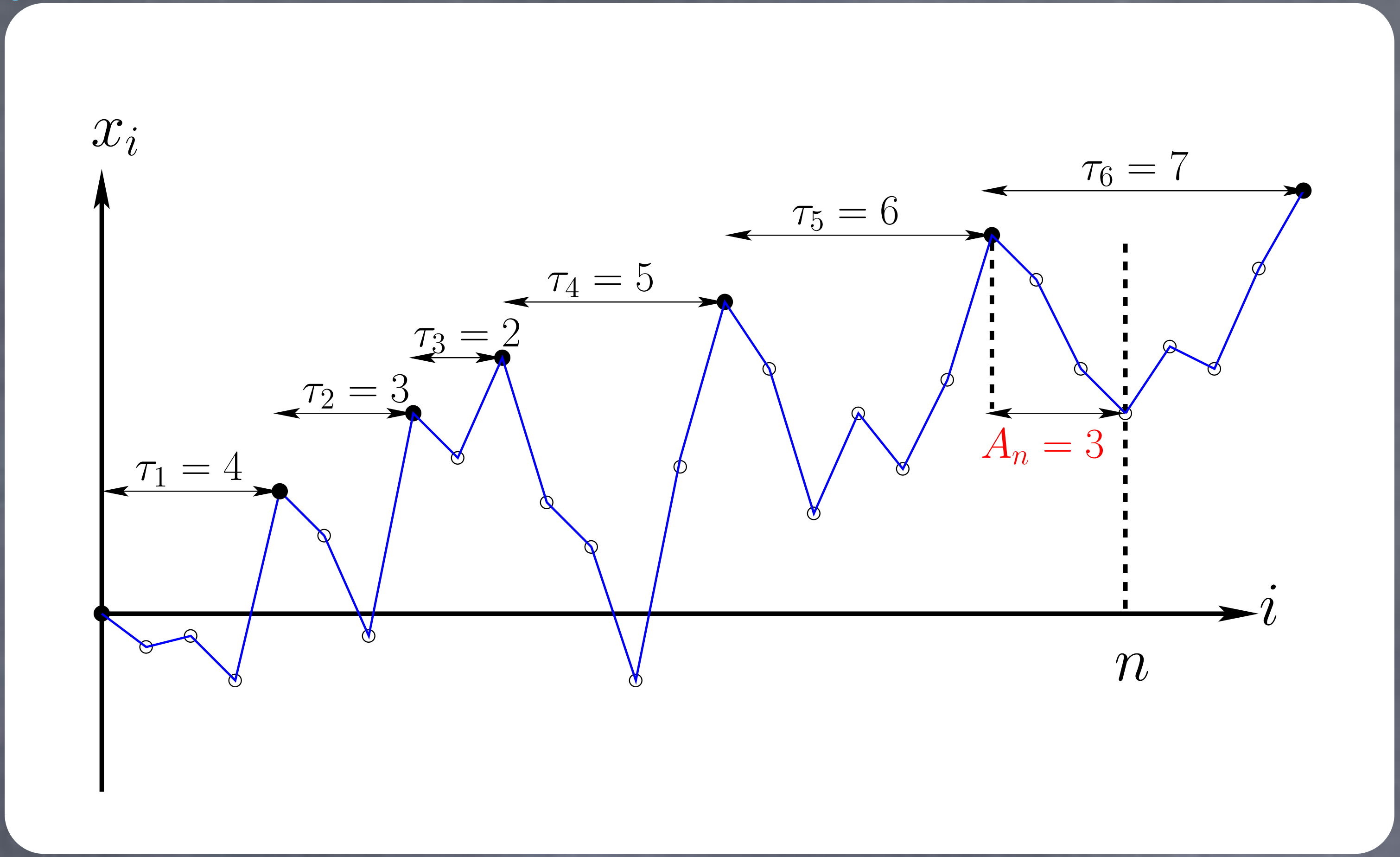


$$Q^{\mathrm{II}}(n) = \Pr[\tau_m \geq \max(\tau_1, \cdots, \tau_{m-1})]$$

$$\lim_{n\to\infty} Q^{\mathrm{II}}(n) = Q^{\mathrm{II}}(\infty)$$

Sensitivity to the definition of the age of the last record

Godrèche, Majumdar, G. S., `14



$$Q^{\mathrm{II}}(n) = \Pr[\tau_m \geq \max(\tau_1, \cdots, \tau_{m-1})]$$

$$\lim_{n\to\infty} Q^{\mathrm{II}}(n) = Q^{\mathrm{II}}(\infty)$$

$$Q^{\mathrm{II}}(\infty) = \frac{1}{2}\int_0^\infty \mathrm{d}x \, \frac{\mathrm{e}^x - 1}{x + \sqrt{\pi} x^{3/2} \, \mathrm{e}^x \, \mathrm{erf}(\sqrt{x})} = 0.800310\ldots \neq 0.626508\ldots$$