Universal record statistics for random walks and Lévy flights with a nonzero staying probability

Journal of Physics A: Mathematical and Theoretical

(Accepted Manuscript)

Satya N.Majumdar, Philippe Mounaix, and Grégory Schehr

Record Statistics

[7월 하순 기록적 폭우...울진 등 하루 강수량 역대 최다 | 연합뉴스 (yna.co.kr)]

Definition of Record Events and Random Variables

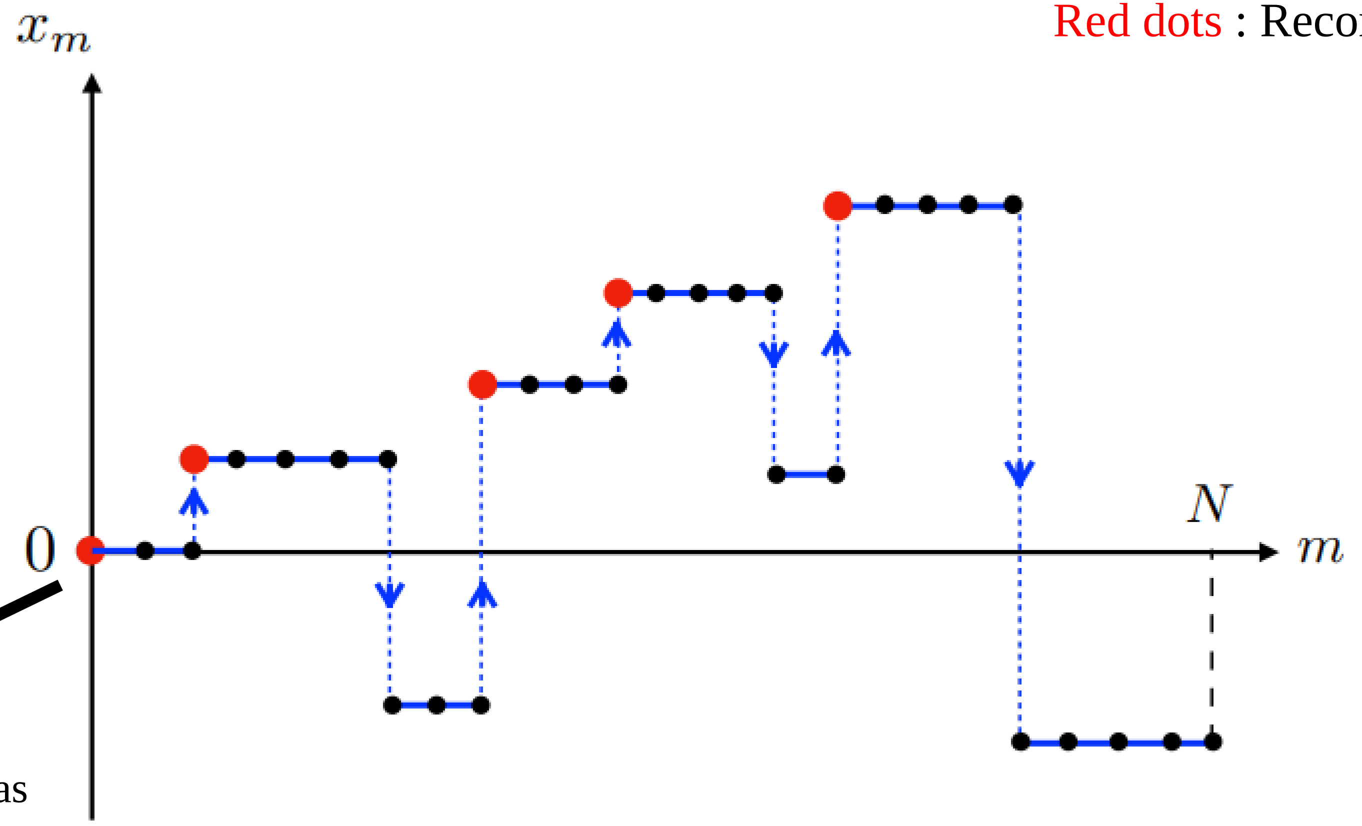
- Time series data (discrete-time with N entries) : $\{x_1, x_2, \ldots, x_N\}$
- The number of Records : R_N (random variable in 'record statistics')

- Record happens at step m : $x_{R_m=k} > max\{x_{R_m=1}, x_{R_m=2}, \ldots, x_{R_m=k-1}\}$

- Three Cases of Random Variables $\{x_i\}$
  1. x_i's are independent, identically and continuously distributed Random Variables (IICD)
  2. x_i from $\varphi(x) = p\delta(x) + (1-p)\varphi_0(x)$, p : the probability that the walker stays at a given position with a zero
  3. $x_i = x_{i-1} + \eta_i$ and η_i from $f(\eta)$: symmetric and continuous distribution.

Definition of Record Events

Red dots : Record Events Occurred Points



By Convention,
Starting Point is also regarded as a Record Event.

FIG 1.

Record Statistics' main concern

- R_N : Record Number
- $\langle R_N \rangle$: mean value of Record Number
- V_N : Variance of Record Number
- $F_N = \frac{V_N}{\langle R_N \rangle}$: Fano Factor
- $C(m_1, m_2) = \langle \sigma_{m_1} \sigma_{m_2} \rangle - \langle \sigma_{m_1} \rangle \langle \sigma_{m_2} \rangle$: Connected Correlation Function

In this paper..

- Dealing with the more general case; Discrete-time version of "Instantaneous Run" model[1],
  - $x_i = x_{i-1} + \eta_i$ and η_i from $f(\eta)$
  - $f(\eta) = p\delta(\eta) + (1-p)f_0(\eta)$, $f_0(\eta)$: symmetric and continuous distribution.
- Main focus
  - For any fixed $0 \le p \le 1$, find record number R_N for all N
  - Compute $\langle R_N \rangle(p)$, $V_N(p)$ exactly for all p and all N and find $F_N(p)$
  - For $p \neq 0$, additional *negative* correlations exist.
  - $C_p(m_1, m_2) = \langle \sigma_{m_1} \sigma_{m_2} \rangle - \langle \sigma_{m_1} \rangle \langle \sigma_{m_2} \rangle$, $C_p(m_1, m_2) < C_0(m_1, m_2)$ for all m_1, m_2 and all $0 < p \le 1$
  - For lattice random walks, it's universality is different from other continuous $f_0(\eta)$s. For simplicity, it is not considered in this paper.

[1] PHYSICAL REVIEW E 102, 042133 (2020)

Previous Research Results

Record Statistics on Three Cases of Random Variables $\{x_i\}$

1. x_i's are IICD, distribution : $\varphi(x)$ case.
2. x_i from $\varphi(x) = p\delta(x) + (1-p)\varphi_0(x)$ case.
3. x_i from $x_m = x_{m-1} + \eta_m$ case. Markov jump process.

1. x_i's are IICD, distribution : $\varphi(x)$ case.

- The average number of Records : $\langle R_N \rangle = 1 + \frac{1}{2} + \cdots + \frac{1}{N}$
  - $\langle R_N \rangle$ is independent of $\varphi(x)$
  - For large N, $\langle R_N \rangle \backsimeq \log N$, $V_N = \langle R_N^2 \rangle - \langle R_N \rangle^2 \backsimeq \log N$
  - Fano Factor : $F_N \equiv \frac{V_N}{\langle R_N \rangle}$, for large N, $F_N \longrightarrow 1$

- R_N is a Poissonian statistics, the Fano factor would be exactly $F_N = 1$ in IICD case.

U. Fano, Phys. Rev. '70, 44 (1946).

1. Analysis of $\langle R_N \rangle$, IICD case.

- Binary variable σ_m (record indicator variable)
  - $\sigma_m = 1$: record happens at step m
  - $\sigma_m = 0$: otherwise

- $R_N = \sum_{m \le N} \sigma_m$

- $\langle R_N \rangle = \sum_{m \le N} \langle \sigma_m \rangle = \sum_{m \le N} \frac{1}{m}$, $\langle \sigma_m \rangle$: record rate
  - Record rate: the probability that a record happens at step m.

1. Why the higher moments of R_N is universal in IICD case?

- *Convolution theorem* for Fourier transforms, (*Random Walks and Random Environments*, Barry D. Hughes)

$$\langle \sigma_{m_1} \sigma_{m_2} \rangle = \begin{cases} \langle \sigma_{m_1} \rangle \langle \sigma_{m_2} \rangle \quad , & m_1 \neq m_2 \\ \langle \sigma_{m_1} \rangle \quad , & m_1 = m_2 \end{cases}$$

- Record-breaking events σ_m : completely uncorrelated in this case.
  - x_i from IICD, $\varphi(x)$ case.

1. Variance & Fano Factor in IICD case.

- $C(m_1, m_2) \equiv \langle \sigma_{m_1} \sigma_{m_2} \rangle - \langle \sigma_{m_1} \rangle \langle \sigma_{m_2} \rangle$

- $\langle R_N^2 \rangle = \sum_{m_1=0}^{N} \sum_{m_2=0}^{N} \langle \sigma_{m_1} \sigma_{m_2} \rangle$

- $V_N = \langle R_N^2 \rangle - \langle R_N \rangle^2 = \sum_{m_1=1}^{N} \sum_{m_2=1}^{N} C(m_1, m_2) = \sum_{m_1=m_2=1}^{N} \langle \sigma_{m_1} \rangle = \langle R_N \rangle$

- $F_N = \frac{V_N}{\langle R_N \rangle} = 1$

2. x_i from $\varphi(x) = p\delta(x) + (1-p)\varphi_0(x)$ case.

- $\Phi(x) = \int_0^x dx'\varphi(x) = p\theta(x) + (1-p)\Phi_0(x)$
  - $\theta(x)$: Heaviside theta function
  - $\Phi_0(x) = \int_0^x dx'\varphi_0(x)$

- $\langle R_N \rangle = \sum_{m=1}^{n} \langle \sigma_m \rangle$
  - $\langle \sigma_m \rangle$: Probability that a record occurs on mth day, Record rate at m step.
  - $\langle \sigma_m \rangle = (1-p)\int_0^\infty dx\varphi_0(x)\Phi(x)^{m-1} = \int_0^\infty (1-p)\varphi_0(x)dx\left(\int_0^x \varphi(x')dx'\right)^{m-1}$

PHYSICAL REVIEW LETTERS 122, 158702 (2019)

2. x_i from $\varphi(x) = p\delta(x) + (1-p)\varphi_0(x)$ case.

- $\langle R_N \rangle = \sum_{m=1}^{n} \langle \sigma_m \rangle$
  - $\langle \sigma_m \rangle$: Probability that a record occurs on mth day, Record rate at m step.
  - $\langle \sigma_m \rangle = (1-p)\int_0^\infty dx\varphi_0(x)\Phi(x)^{m-1} = \int_0^\infty (1-p)\varphi_0(x)dx\left(\int_0^x \varphi(x')dx'\right)^{m-1}$
  - $x \to u = \Phi(x), \quad u \in [p,1] \quad$ and $\quad du = (1-p)\varphi_0(x)dx \quad$ for $\quad x > 0.$
  - $\langle \sigma_m \rangle = \int_p^1 du\, u^{m-1} = \frac{1-p^m}{m}$
  - $\langle R_N \rangle = \sum_{m=1}^{n} \frac{1-p^m}{m}$

2. x_i from $\varphi(x) = p\delta(x) + (1-p)\varphi_0(x)$ case.

- $\langle R_N^2 \rangle = \langle R_N \rangle + 2\sum_{m_1=1}^{n-1}\sum_{m_2=1}^{n-m_1}\langle \sigma_{m_1}\sigma_{m_1+m_2}\rangle$
  - $\langle \sigma_{m_1}\sigma_{m_1+m_2}\rangle = \int_p^1 u_2^{m_2-1}du_2\left(\int_p^{u_2} du_1 u_1^{m_1-1}\right)$
  - $C(m_1, m_1+m_2) \equiv \langle \sigma_{m_1}\sigma_{m_1+m_2}\rangle - \langle \sigma_{m_1}\rangle\langle \sigma_{m_1+m_2}\rangle = -\frac{p^{m_1}}{m_1}\left(\frac{1-p^{m_2}}{m_2} - \frac{1-p^{m_1+m_2}}{m_1+m_2}\right)$

  - For each m_1, m_1+m_2 step, record events is anticorrelated when $0 < p < 1$.
    - Record events is uncorrelated when $p = 0$. (Same as IICD case)

2. x_i from $\varphi(x) = p\delta(x) + (1-p)\varphi_0(x)$ case.

- For large N, $V_N(p) = \langle R_N^2 \rangle - \langle R_N \rangle^2 \to \langle R_N \rangle - \frac{\pi^2}{6} \approx \ln\{(1-p)N\} + \gamma_E - \frac{\pi^2}{6}$
  - γ_E : Euler-Mascheroni constant

- In the scaling limit $(1-p) \to 0, \quad N \to \infty$ at fixed $t = (1-p)N$,
  - (Only considering walkers at zero position)
  - $\langle R_N \rangle \to \mu((1-p)N)$, $\quad \mu(t) = \int_0^t dy \frac{1-e^{-y}}{y} = \ln t + \gamma_E + \int_t^\infty dz \frac{e^{-z}}{z}$, $\quad \gamma_E = -\int_0^\infty dy e^{-y} \ln y$

- $F_N(p) \to F(t = (1-p)N)$

- $F(t) = 1 + \frac{2}{\mu(t)} \int_0^t \frac{dz}{z} e^{-z} [\mu(t) - \mu(z) - \mu(t-z)]$

2. x_i from $\varphi(x) = p\delta(x) + (1-p)\varphi_0(x)$ case.

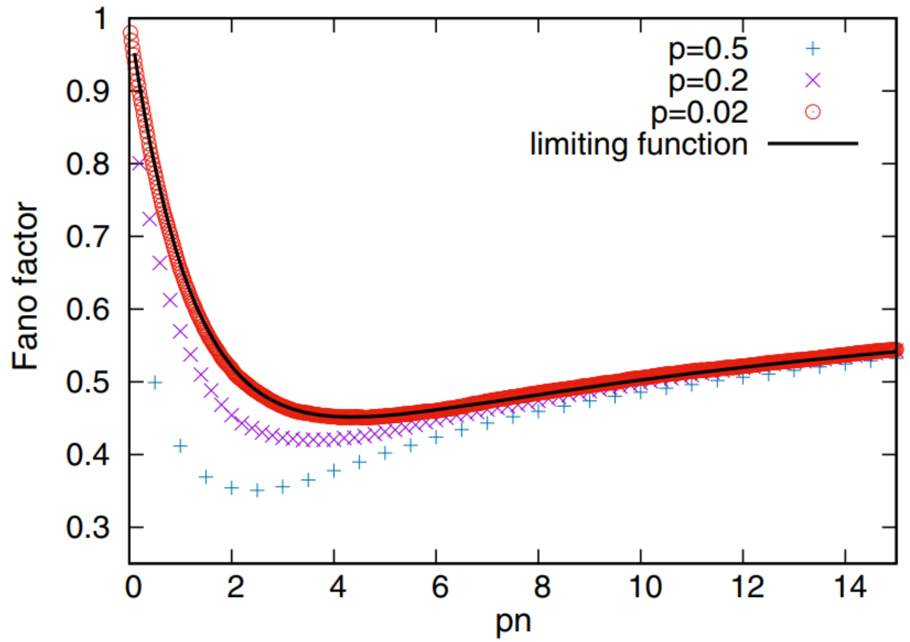


FIG. 1. The Fano factor of the record process obtained from simulations (symbols) is compared to the analytic limit function $F(t)$ in Eq. (14) (full line). Note that the numerical estimates start at $F_1 = 1 - p$.

In FIG 1. $p \to (1-p)$

(Notation is different)

PHYSICAL REVIEW LETTERS 122, 158702 (2019)

3. x_i from $x_m = x_{m-1} + \eta_m$ case. Markov jump process.

- η_m's : jump lengths of x_m
  - IICD; drawn from symmetric & continuous distribution $f(\eta)$

- $f(\eta)$ also includes Lévy flights $f(\eta) \sim |\eta|^{-1-\mu}$ for large $|\eta|$ and $0 < \mu \leq 2$

- By using survival probability(Sparre-Andersen Theroem) and first-passage probability, the number of Record($M \equiv \langle R_N \rangle$)'s moment generating function is followed
  - $\langle M(Z) \rangle = (1 - z)^{-1.5}$

PHYSICAL REVIEW LETTERS 101, 050601 (2008)

3. x_i from $x_m = x_{m-1} + \eta_m$ case. Markov jump process.

- $\langle M \rangle = (2N+1)\binom{2N}{N}2^{-2N} \simeq \sqrt{\frac{4N}{\pi}}$ as $N \to \infty$
- $\langle M^2 \rangle = 2N + 2 - \langle M \rangle$
- $\langle V_N \rangle = \langle M^2 \rangle - \langle M \rangle^2 \simeq 2N + 2 - \sqrt{\frac{4N}{\pi}} - \frac{4N}{\pi} \simeq 2N\left(1 - \frac{2}{\pi}\right)$ as $N \to \infty$
- $F_N = \frac{V_N}{\langle M \rangle} \simeq \sqrt{N}\left(\sqrt{\pi} - \frac{2}{\sqrt{\pi}}\right) \cong 0.644\sqrt{N}$ as $N \to \infty$

Current Research Results

1. Correlation between Record Events: Universal Expression
2. Exact Statistics of Records for Arbitrary $0 \leq p \leq 1$
3. Correlation and Fano Factor in the Continuous Time Limit

Record Statistics' main concern (Remind)

- R_N : Record Number
- $\langle R_N \rangle$: mean value of Record Number
- V_N : Variance of Record Number
- $F_N = \frac{V_N}{\langle R_N \rangle}$: Fano Factor
- $C(m_1, m_2) = \langle \sigma_{m_1} \sigma_{m_2} \rangle - \langle \sigma_{m_1} \rangle \langle \sigma_{m_2} \rangle$: Connected Correlation Function

1. Correlation between Record Events: Universal Expression

- More universal expression..
  - $x_i = x_{i-1} + \eta_i$ and η_i from $f(\eta)$
  - $f(\eta) = p\delta(\eta) + (1-p)f_0(\eta)$, $f_0(\eta)$: symmetric and continuous distribution.

- Mean Record Number
  - $\langle R_N \rangle(p) = \sum_{m=0}^{N} \langle \sigma_m \rangle = \sum_{m=0}^{N} q_p(m)$
    - $q_p(m) = \text{Prob}(x_1 > 0, x_2 > 0, \ldots, x_m > 0 | x_0 = 0)$ for $m \geq 1$: Survival Probability
    - $\langle \sigma_m \rangle = q_p(m)$: shift the origin to the value x_m & reverse the time (Symmetric nature of the walk)

1. Correlation between Record Events: Universal Expression

- Second Moment & Variance of Record Number
  - $\langle R_N^2 \rangle(p) = \sum_{m_2=0}^{N} \sum_{m_1=0}^{N} \langle \sigma_{m_1} \sigma_{m_2} \rangle = -\sum_{m=0}^{N} \langle \sigma_m \rangle + 2 \sum_{m_2=0}^{N} \sum_{m_1=0}^{m_2} \langle \sigma_{m_1} \sigma_{m_2} \rangle$
  - $V_N(p) = \langle R_N^2 \rangle(p) - \left(\langle R_N \rangle(p) \right)^2 = \sum_{m_1=1}^{N} \sum_{m_2=1}^{N} C_p(m_1, m_2)$
  - $\langle \sigma_{m_1} \sigma_{m_2} \rangle = q_p(m_1) q_p(m_2 - m_1)$ for $m_2 \geq m_1$
    - $[0, m_1]$ and $[m_1, m_2]$ statistically independent (Markov property)
  - $V_N(p) = -\sum_{m=0}^{N} \langle \sigma_m \rangle + 2 \sum_{m_2=0}^{N} \sum_{m_1=0}^{m_2} \langle \sigma_{m_1} \sigma_{m_2} \rangle - \left(\sum_{m=0}^{N} \langle \sigma_m \rangle \right)^2$
  - $V_N(p) = -\sum_{m=0}^{N} q_p(m) + 2 \sum_{m_2=0}^{N} \sum_{m_1=0}^{m_2} q_p(m_1) q_p(m_2 - m_1) - \left(\sum_{m=0}^{N} q_p(m) \right)^2$

1. Correlation between Record Events: Universal Expression

- Connected Correlation Function
  - $C_p(m_1, m_2) = \langle \sigma_{m_1} \sigma_{m_2} \rangle - \langle \sigma_{m_1} \rangle \langle \sigma_{m_2} \rangle = q_p(m_1)[q_p(m_2 - m_1) - q_p(m_2)]$ from Markovian property
- Generalised Sparre Andersen Theorem
  - $Q_p(s) = \sum_{m \geq 0} q_p(m) s^m = \exp\left[\sum_{n \geq 1} \frac{s^n}{n} \mathrm{Prob}(x_n < 0)\right]$
  - $\mathrm{Prob}(x_n < 0) = \mathrm{Prob}(x_n > 0)$
  - $2\mathrm{Prob}(x_n < 0) + \mathrm{Prob}(x_n = 0) = 1$
  - $\mathrm{Prob}(x_n < 0) = \frac{1 - \mathrm{Prob}(x_n = 0)}{2}$

1. Correlation between Record Events: Universal Expression

- Generalised Sparre Andersen Theorem
  - $Q_p(s) = \sum_{m\geq 0} q_p(m) s^m = \exp\left[\sum_{n\geq 1} \frac{s^n}{n} \mathrm{Prob}(x_n < 0)\right]$
  - $2\mathrm{Prob}(x_n < 0) + \mathrm{Prob}(x_n = 0) = 1$
  - $\mathrm{Prob}(x_n < 0) = \frac{1-\mathrm{Prob}(x_n=0)}{2} = \frac{1-p^n}{2}$
  - Using $\sum_{n\geq 1} \frac{s^n}{n} = -\ln(1-s)$ and $\exp(a+b) = \exp(a)\exp(b)$ property,
  - $Q_p(s) = \frac{1}{Z_p(s)\sqrt{1-s}}$, where $Z_p(s) = \exp\left[\frac{1}{2}\sum_{n=1}^{\infty} \frac{s^n}{n} \mathrm{Prob}\,(x_n = 0)\right] = \exp\left[\frac{1}{2}\sum_{n=1}^{\infty} \frac{s^n}{n} p^n\right]$

1. Correlation between Record Events: Universal Expression

- From generating function table,
  - $Q_p(s) = \sum_{m\geq 0} q_p(m) s^m = \frac{1}{Z_p(s)\sqrt{1-s}} = \frac{\sqrt{1-sp}}{\sqrt{1-s}}$
- From power series expansions
  - $\sqrt{1-sp} = \sum_{n\geq 0} (-1)^n \binom{1/2}{n} p^n s^n$
  - $\frac{1}{\sqrt{1-s}} = \sum_{n\geq 0} \binom{2n}{n} 2^{-2n} s^n = \sum_{n\geq 0} (-1)^n \binom{-1/2}{n} s^n$

1. Correlation between Record Events: Universal Expression

- Get Survival Probability from generating function of it,
  - $Q_p(s) = \sum_{m\geq 0} q_p(m) s^m = \frac{1}{Z_p(s)\sqrt{1-s}} = \frac{\sqrt{1-sp}}{\sqrt{1-s}} = \sum_{n\geq 0}(-1)^n \binom{1/2}{n} p^n s^n \sum_{n\geq 0}(-1)^n \binom{-1/2}{n} s^n$
  - $q_p(m) = (-1)^m \sum_{k=0}^{m} \binom{1/2}{k} \binom{-1/2}{m-k} p^k = (-1)^m \binom{-1/2}{m} {}_2F_1\left(-\frac{1}{2}, -m; \frac{1}{2} - m; p\right)$
  - $q_p(m) = \binom{2m}{m} 2^{-2m} {}_2F_1\left(-\frac{1}{2}, -m; \frac{1}{2} - m; p\right)$, ${}_2F_1$: standard hypergeometric series
  - When $p \to 0$ limit, using ${}_2F_1(a, b; c; z = 0) = 1$, $q_0(m) = \binom{2m}{m} 2^{-2m}$

1. Correlation between Record Events: Universal Expression

- Additional anticorrelation Exists
  - $\Delta_p(m_1, m_2) = C_p(m_1, m_2) - C_0(m_1, m_2)$
  - In the $m \to +\infty$ limit, because of the dominant contribution of $s = 1$ in $Q_p(s)$.
  - $Q_p(s) \simeq \frac{\sqrt{1-p}}{\sqrt{1-s}}$ as $m \to +\infty$
  - $q_p(m) \simeq \sqrt{1-p}\frac{1}{\sqrt{\pi m}}$ as $m \to +\infty$
  - $C_p(m_1, m_2) \simeq \frac{1-\mathrm{p}}{\pi}\frac{1}{\sqrt{m_1}}\left(\frac{1}{\sqrt{m_2-m_1}} - \frac{1}{\sqrt{m_2}}\right)$ as $m \to +\infty$

1. Correlation between Record Events: Universal Expression

- Additional anticorrelation Exists
  - $\Delta_p(m_1, m_2) = C_p(m_1, m_2) - C_0(m_1, m_2)$
  - $C_p(m_1, m_2) \simeq \frac{1-\mathrm{p}}{\pi}\frac{1}{\sqrt{m_1}}\left(\frac{1}{\sqrt{m_2 - m_1}} - \frac{1}{\sqrt{m_2}}\right)$ as $m_1 \to +\infty, m_2 \to +\infty$
  - $C_0(m_1, m_2) \simeq \frac{1}{\pi}\frac{1}{\sqrt{m_1}}\left(\frac{1}{\sqrt{m_2 - m_1}} - \frac{1}{\sqrt{m_2}}\right)$ as $m_1 \to +\infty, m_2 \to +\infty$
  - $\Delta_p(m_1, m_2) \simeq -\frac{\mathrm{p}}{\pi}\frac{1}{\sqrt{m_1}}\left(\frac{1}{\sqrt{m_2 - m_1}} - \frac{1}{\sqrt{m_2}}\right) < 0$ as $m_2 \geq m_1$
  - $C_p(m_1, m_2) < C_0(m_1, m_2)$

1. Correlation between Record Events: Universal Expression

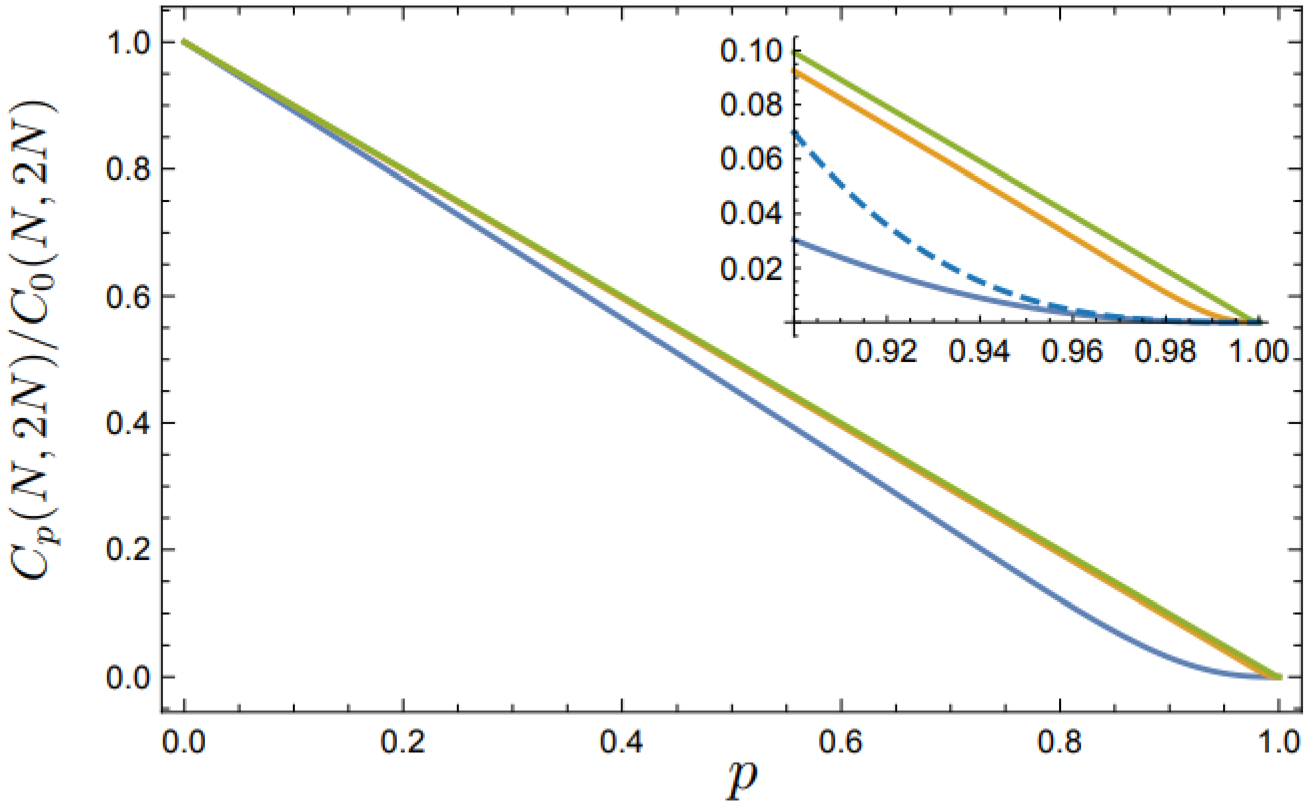


FIG 2.

$N = 10$ (blue)
$N = 100$ (orange)
$N = 1000$ (green)

$m_1 = 10$ (Analytic Line, dashed)

Analytic Line
: $C_p(m_1, m_2) \simeq \frac{m_1(1-p)^3}{16}$

From Taylor Expansion of $q_p(m)$
at $p = 1 - \epsilon$ in powers of ϵ

$q_p(m) \simeq \frac{1-p}{2} - \frac{(m-1)(1-p)^2}{8}$ as $(p \to 1)$

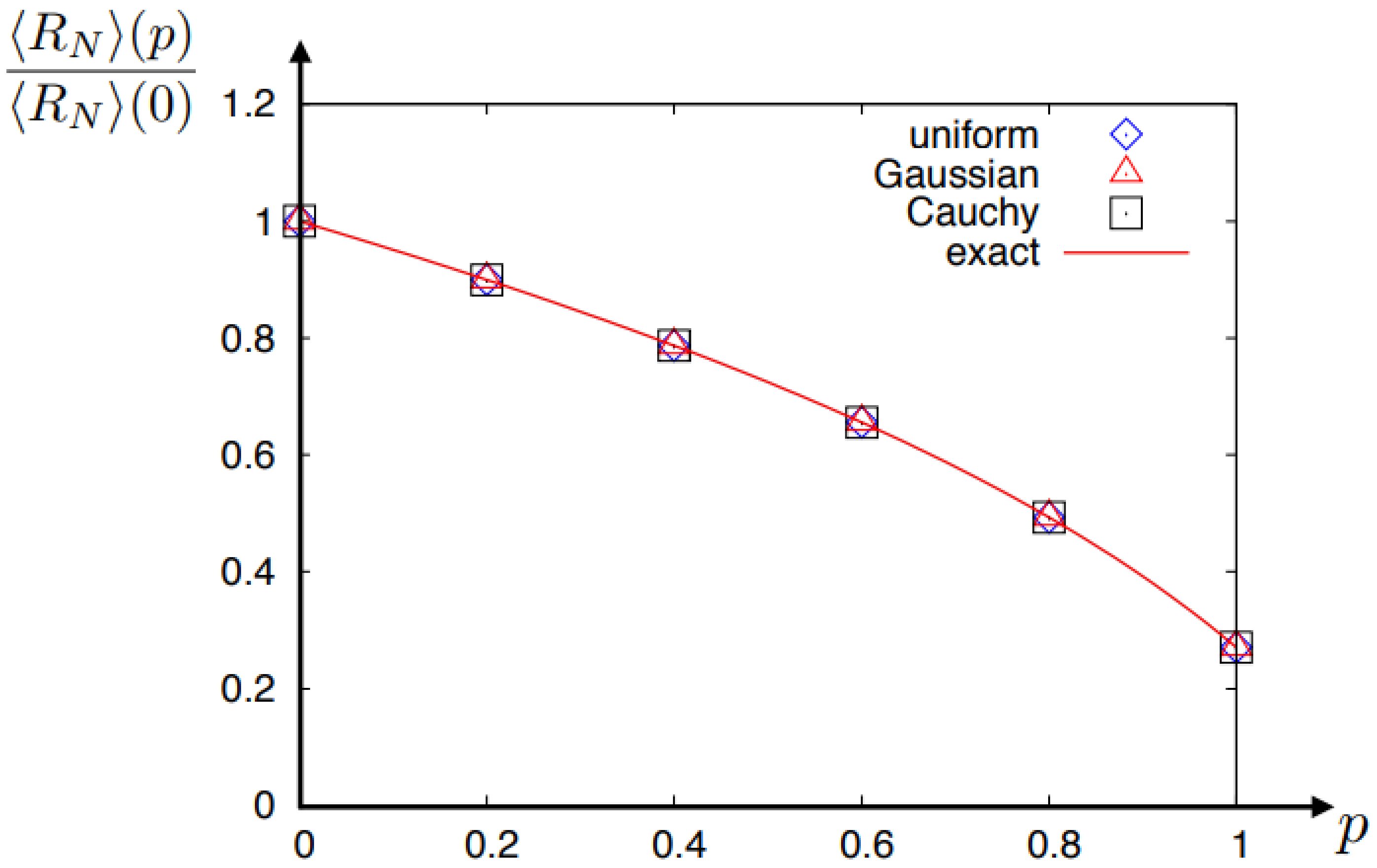
2. Exact Statistics of Records for Arbitrary $0 \leq p \leq 1$

- Average Number of Records : Exact Universal Expression
  - $\langle R_N \rangle(p) = \sum_{m=0}^{N} q_p(m)$
  - Generating function of Average Number of Records
    - $\sum_{N\geq 0}\langle R_N \rangle(p) s^N = \sum_{N\geq 0}\sum_{m=0}^{N} q_p(m)\, s^N = \frac{1}{1-s} Q_p(s) = \frac{\sqrt{1-sp}}{(1-s)^{1.5}}$, $Q_p(s)$ is generating function of survival probability
  - From power series expansion
    - $\sqrt{1-sp} = \sum_{n\geq 0}(-1)^n \binom{1/2}{n} p^n s^n$
    - $\frac{1}{(1-s)^{1.5}} = \sum_{n\geq 0}(-1)^n \binom{-3/2}{n} s^n$

2. Exact Statistics of Records for Arbitrary $0 \leq p \leq 1$

- Average Number of Records : Exact Universal Expression
  - $\langle R_N \rangle(p) = \sum_{m=0}^{N} q_p(m)$
  - $\langle R_N \rangle(p) = (-1)^N \sum_{m=0}^{N} \binom{1/2}{m} \binom{-3/2}{N-m} p^m = (-1)^N \binom{-\frac{3}{2}}{N} {}_2F_1\left(-\frac{1}{2}, -\mathrm{N}; -\frac{1}{2} - \mathrm{N}; \mathrm{p}\right)$
  - $\langle R_N \rangle(p) = (2\mathrm{N}+1)\binom{2N}{N} 2^{-2N} {}_2F_1\left(-\frac{1}{2}, -\mathrm{N}; -\frac{1}{2} - \mathrm{N}; \mathrm{p}\right)$, ${}_2F_1$: standard hypergeometric series
  - $\frac{\langle R_N \rangle(p)}{\langle R_N \rangle(0)} = {}_2F_1\left(-\frac{1}{2}, -\mathrm{N}; -\frac{1}{2} - \mathrm{N}; \mathrm{p}\right)$

2. Exact Statistics of Records for Arbitrary $0 \leq p \leq 1$



Each distribution : jump distribution ($f_0(\eta)$)
Exactly at total step number $N = 10$

$\langle R_N \rangle(p)$ is only dependent of N & p.
Independent of jump distribution.

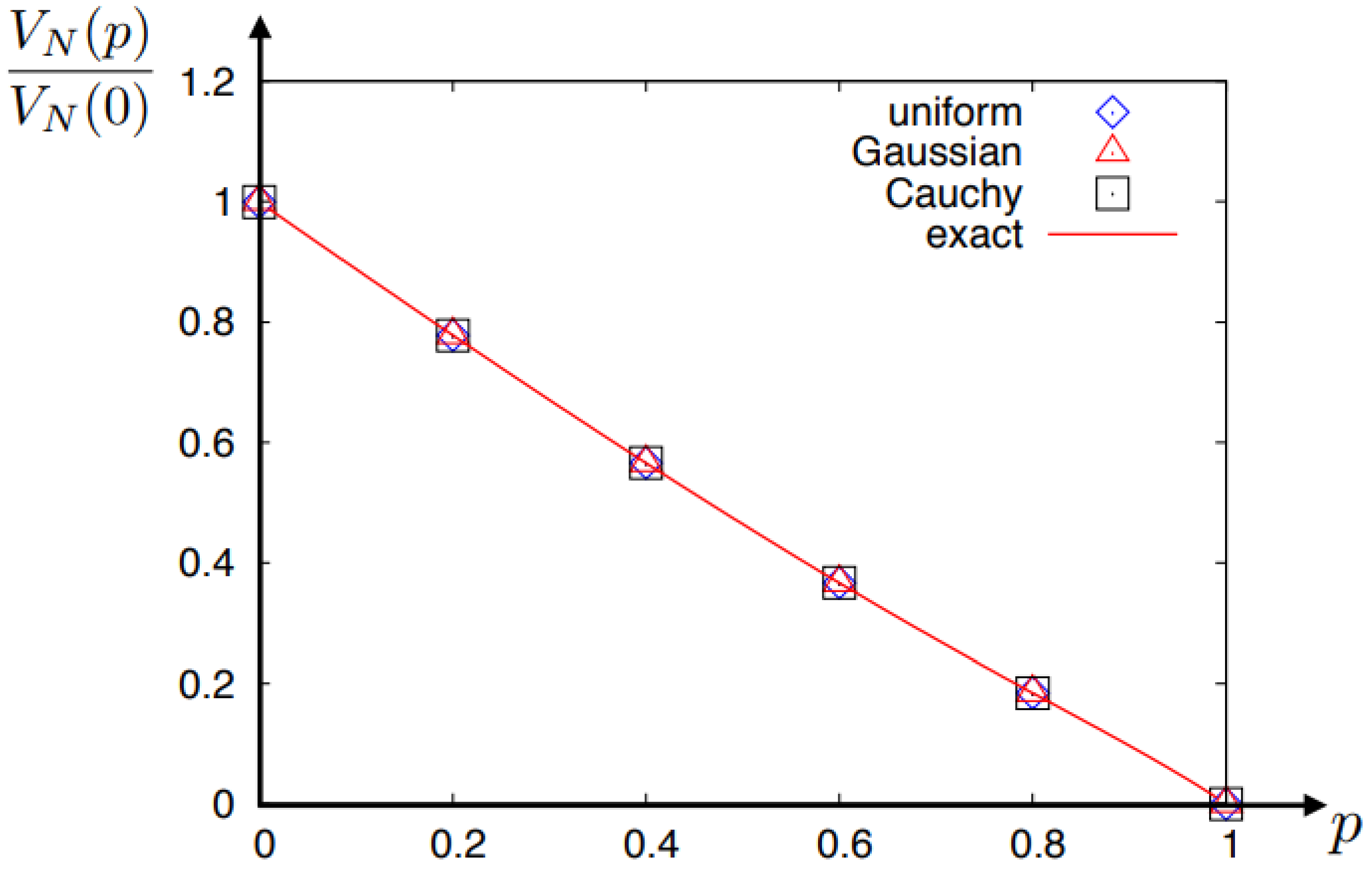
2. Exact Statistics of Records for Arbitrary $0 \leq p \leq 1$

- Variance of the Number of Records : Exact Universal Expression
  - $\langle R_N \rangle(p) = \sum_{m=0}^{N} q_p(m)$
  - $\langle R_N^2 \rangle(p) = -\sum_{m=0}^{N} \langle \sigma_m \rangle + 2 \sum_{m_2=0}^{N} \sum_{m_1=0}^{m_2} \langle \sigma_{m_1} \sigma_{m_2} \rangle$
  - $\sum_{N \geq 0} \langle R_N^2 \rangle(p) s^N = \frac{1}{1-s} \left(-Q_p(s) + 2Q_p^2(s) \right)$
  - $\sum_{N \geq 0} \langle R_N^2 \rangle(p) s^N = -\frac{\sqrt{1-sp}}{(1-s)^{1.5}} + \frac{2(1-sp)}{(1-s)^2}$
  - $\frac{2(1-sp)}{(1-s)^2} = 2 \sum_{n \geq 0} [(1-p)n + 1] s^n$

2. Exact Statistics of Records for Arbitrary $0 \leq p \leq 1$

- Variance of the Number of Records : Exact Universal Expression
  - $\sum_{N\geq 0}\left\langle R_N^2\right\rangle(p)s^N = -\frac{\sqrt{1-sp}}{(1-s)^{1.5}} + \frac{2(1-sp)}{(1-s)^2}$
  - $\frac{2(1-sp)}{(1-s)^2} = 2\sum_{n\geq 0}[(1-p)n+1]s^n$
  - $\left\langle R_N^2\right\rangle(p) = 2[(1-p)n+1] - \langle R_N\rangle(p)$
  - $V_N(p) = \left\langle R_N^2\right\rangle(p) - \langle R_N\rangle(p)^2 = 2[(1-p)n+1] - \langle R_N\rangle(p)[\langle R_N\rangle(p)+1]$

2. Exact Statistics of Records for Arbitrary $0 \leq p \leq 1$



Each distribution : jump distribution ($f_0(\eta)$)
Exactly at total step number $N = 10$

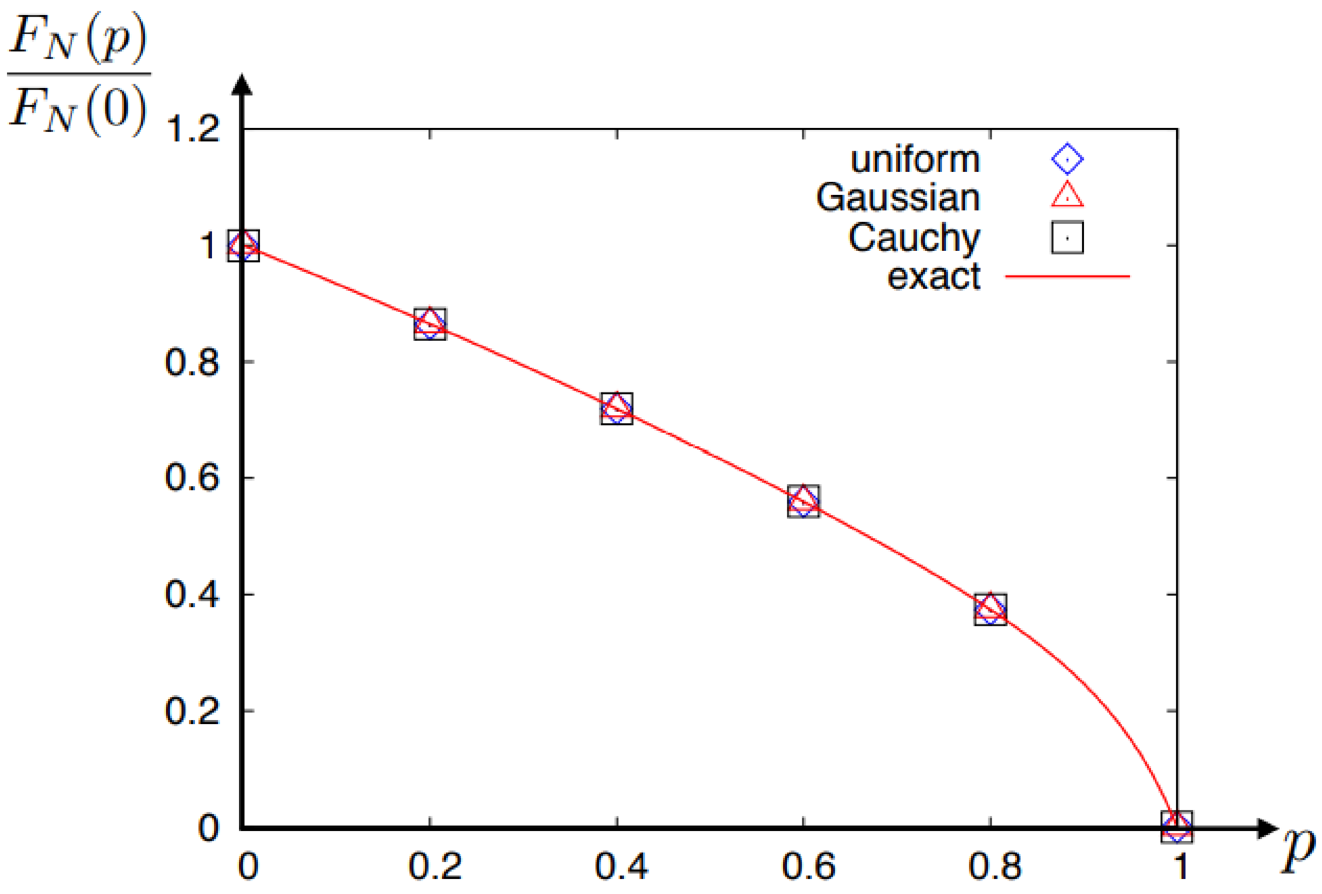
$V_N(p)$ is only dependent of N & p.
Independent of jump distribution.

2. Exact Statistics of Records for Arbitrary $0 \le p \le 1$

- Fano Factor : Exact Universal Expression
  - $\langle R_N \rangle(p) = (2\mathrm{N}+1)\binom{2N}{N}2^{-2N}\ {}_2F_1\left(-\frac{1}{2}, -\mathrm{N}; -\frac{1}{2}-\mathrm{N}; \mathrm{p}\right)$
  - $V_N(p) = 2[(1-p)n+1] - \langle R_N \rangle(p)[\langle R_N \rangle(p)+1]$

  - $F_N(p) = \frac{V_N(p)}{\langle R_N \rangle(p)} = \frac{2[N(1-p)+1]}{\langle R_N \rangle(p)} - \langle R_N \rangle(p) - 1$

2. Exact Statistics of Records for Arbitrary $0 \leq p \leq 1$



Each distribution : jump distribution ($f_0(\eta)$)
Exactly at total step number $N = 10$

$F_N(p)$ is only dependent of N & p.
Independent of jump distribution.

3. Correlation and Fano Factor in the Continuous Time Limit

- Scaling limits
  - $N \to \infty$, $p \to 1$ and keeping $(1-p)N = t$ fixed.
  - In this scaling limits, this model reduces to the continuous time random walk(CTRW) model with exponential waiting-time distribution.

- Record Rate in the Scaling limits
  - $q_p(m) = \frac{1}{2i\pi}\oint \frac{\sqrt{1-sp}}{s^{m+1}\sqrt{1-s}}ds$: generating function of record rate using Cauchy's theorem[2].

[2] Henrici P 1991 Applied and Computational Complex Analysis (Wiley Classics Library vol 2) (New York: Wiley) (Theorem 11.10b: theorem of Darboux)

3. Correlation and Fano Factor in the Continuous Time Limit

- Record Rate in the Scaling limits
  - $q_p(m) = \frac{1}{2i\pi} \oint \frac{\sqrt{1-sp}}{s^{m+1}\sqrt{1-s}} ds$
  - To take the scaling limit, $s = \exp\left(-\frac{\lambda}{m}\right)$
  - In the $m \to +\infty$ limit, because of the dominant contribution of $s = 1$
  - $q_p(m) \simeq \frac{1}{2i\pi m} \int_{\mathcal{B}} \frac{\sqrt{\lambda+(1-p)m}}{\sqrt{\lambda}} e^{\lambda} d\lambda$ as $m \to +\infty, \quad p \to 1$ & $\mathcal{B}$: Bromwich contour
  - $q_p(m) \simeq (1-p)S[(1-p)m], \quad S(t) = \frac{1}{2}\left[I_0\left(\frac{t}{2}\right) + I_1\left(\frac{t}{2}\right)\right] e^{-t/2}$: Scaling Function.
  - $I_\nu(t) = i^{-\nu} J_\nu(it) = \sum_{m=0}^{\infty} \frac{1}{m!\Gamma(m+\nu+1)} \left(\frac{t}{2}\right)^{2m+\nu}$: modified Bessel function of order ν.

Additional Description

- $q_p(m) \simeq \frac{1}{2i\pi m}\int_{\mathcal{B}} \frac{\sqrt{\lambda+(1-p)m}}{\sqrt{\lambda}} e^{\lambda} d\lambda$ as $m \to +\infty,\ \ p \to 1$ & $\mathcal{B}$: Bromwich contour
- Using Bromwich integral,
- Bromwich integral is same with inverse Laplace Transform
- $f(t) = \mathcal{L}^{-1}\{F(s)\}(t) = \frac{1}{2\pi i} \lim_{T\to\infty} \int_{\gamma-iT}^{\gamma+iT} F(s)e^{st} ds$
- $\mathcal{L}\{I_\nu(t)\} = \mathcal{L}\{i^{-\nu}J_\nu(it)\} = \frac{\left(\sqrt{s^2-1}-s\right)^{\nu}}{\sqrt{s^2-1}},\quad \mathcal{L}\{I_0(t)\} = \frac{1}{\sqrt{s^2-1}},\quad \mathcal{L}\{I_1(t)\} = \frac{\sqrt{s^2-1}-s}{\sqrt{s^2-1}}$
- $q_p(m) \simeq (1-p)S[(1-p)m],\qquad S(t) = \frac{1}{2}\left[I_0\left(\frac{t}{2}\right) + I_1\left(\frac{t}{2}\right)\right]e^{-t/2}$: Scaling Function.

3. Correlation and Fano Factor in the Continuous Time Limit

- Record Rate in the Scaling limits
  - $q_p(m) \simeq (1-p)S[(1-p)m],$ $\quad S(t) = \frac{1}{2}\left[I_0\left(\frac{t}{2}\right) + I_1\left(\frac{t}{2}\right)\right]e^{-t/2}$: Scaling Function
  - $S(t) \simeq \begin{cases} \frac{1}{2} - \frac{t}{8}, & \text{as} \quad t \to 0 \\ \frac{1}{\sqrt{\pi t}}, & \text{as} \quad t \to +\infty \end{cases}$

3. Correlation and Fano Factor in the Continuous Time Limit

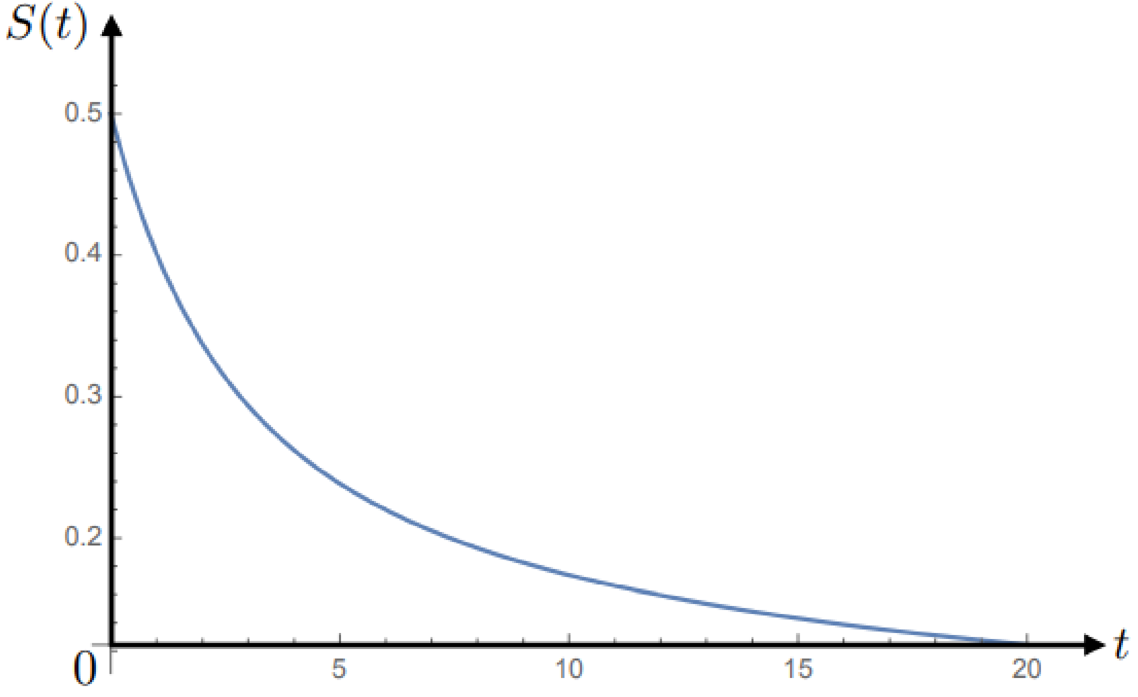


FIG. 6: Plot of $S(t)$ vs t as given in Eq. (66).

Plot of $S(t) = \frac{1}{2}\left[I_0\left(\frac{t}{2}\right) + I_1\left(\frac{t}{2}\right)\right]e^{-t/2}$

Scaling form : $q_p(m) \simeq (1-p)S[(1-p)m]$

It coincides with

1) $q_p(m) \simeq \sqrt{1-p}\frac{1}{\sqrt{\pi m}}$ as $m(1-p) \to +\infty$

2) $q_p(m) \simeq \frac{1-p}{2} - \frac{(m-1)(1-p)^2}{8}$ as $m(1-p) \to +\infty$

3. Correlation and Fano Factor in the Continuous Time Limit

- Connected Correlation Function in the Scaling limits
  - From $q_p(m) \simeq (1-p)S[(1-p)m]$ and $C_p(m_1, m_2) = q_p(m_1)[q_p(m_2 - m_1) - q_p(m_2)]$
  - $C_p(m_1, m_2) \simeq (1-p)^2 \mathcal{C}[(1-p)m_1, (1-p)m_2]$
  - $\mathcal{C}(t_1, t_2) = S(t_1)[S(t_2 - t_1) - S(t_2)]$: Scaling Function of Connected Correlation Function

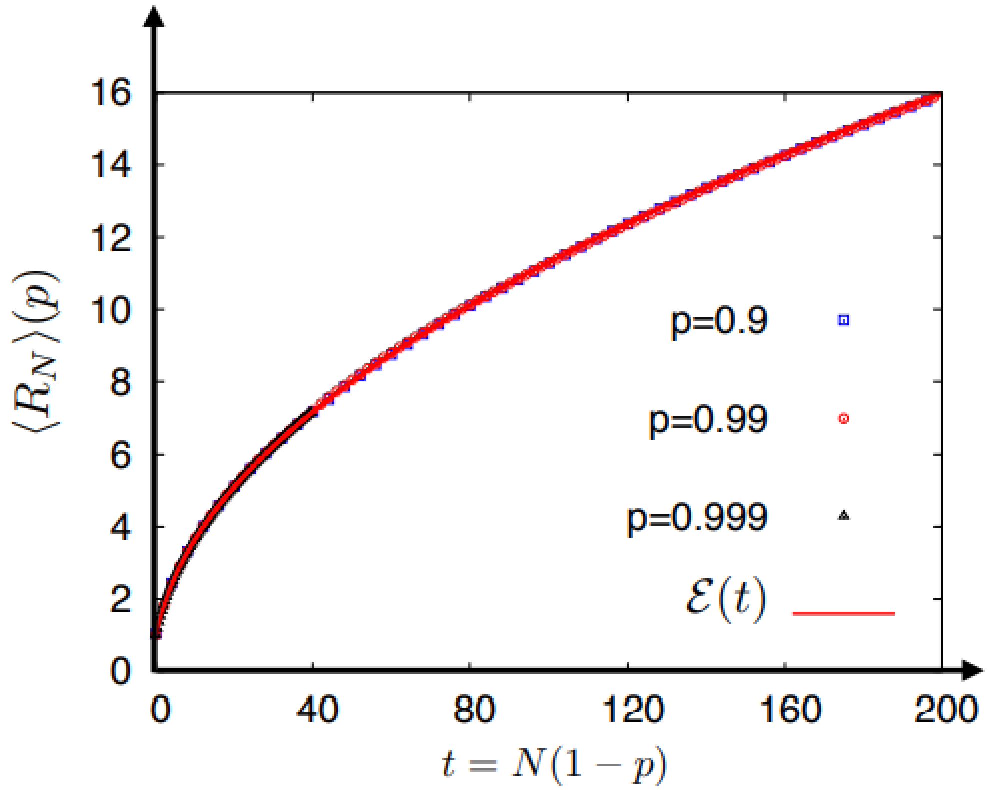
3. Correlation and Fano Factor in the Continuous Time Limit

- Average number of Records in the Scaling limits
  - $\langle R_N \rangle(p) = \frac{1}{2i\pi} \oint \frac{\sqrt{1-sp}}{s^{N+1}(1-s)^{1.5}} ds$
  - To take the scaling limit, $s = \exp\left(-\frac{\lambda}{m}\right)$
  - In the $m \to +\infty$ limit, because of the dominant contribution of $s = 1$
  - $\langle R_N \rangle(p) \simeq \frac{1}{2i\pi} \int_{\mathcal{B}} \frac{\sqrt{\lambda+(1-p)N}}{\lambda^{1.5}} e^{\lambda} d\lambda$ as $N \to +\infty$, $p \to 1$ & $\mathcal{B}$: Bromwich contour

3. Correlation and Fano Factor in the Continuous Time Limit

- Average number of Records in the Scaling limits
  - $\langle R_N \rangle(p) \simeq \frac{1}{2i\pi} \int_{\mathcal{B}} \frac{\sqrt{\lambda + (1-p)N}}{\lambda^{1.5}} e^{\lambda} d\lambda$ as $N \to +\infty, \;\; p \to 1$ & $\mathcal{B}$: Bromwich contour
  - From Bromwich Integral..
  - $\langle R_N \rangle(p) \simeq \mathcal{E}[(1-p)N]$ as $\mathcal{E}(t) = \left[(1+t) I_0\left(\frac{t}{2}\right) + t I_1\left(\frac{t}{2}\right)\right] e^{-t/2}$: Scaling Function
  - $\mathcal{E}(t) \simeq \begin{cases} 1 + \frac{t}{2} \quad , & \text{as} \quad t \to 0 \\ 2\sqrt{\frac{t}{\pi}} \quad , & \text{as} \quad t \to +\infty \end{cases}$

3. Correlation and Fano Factor in the Continuous Time Limit



Plot of $\langle R_N \rangle (p)$

Symbols from simulation with different $p \to 1$ and $N = 40000$

Solid Line is $\langle R_N \rangle (p) \simeq \mathcal{E}[(1-p)N]$

3. Correlation and Fano Factor in the Continuous Time Limit

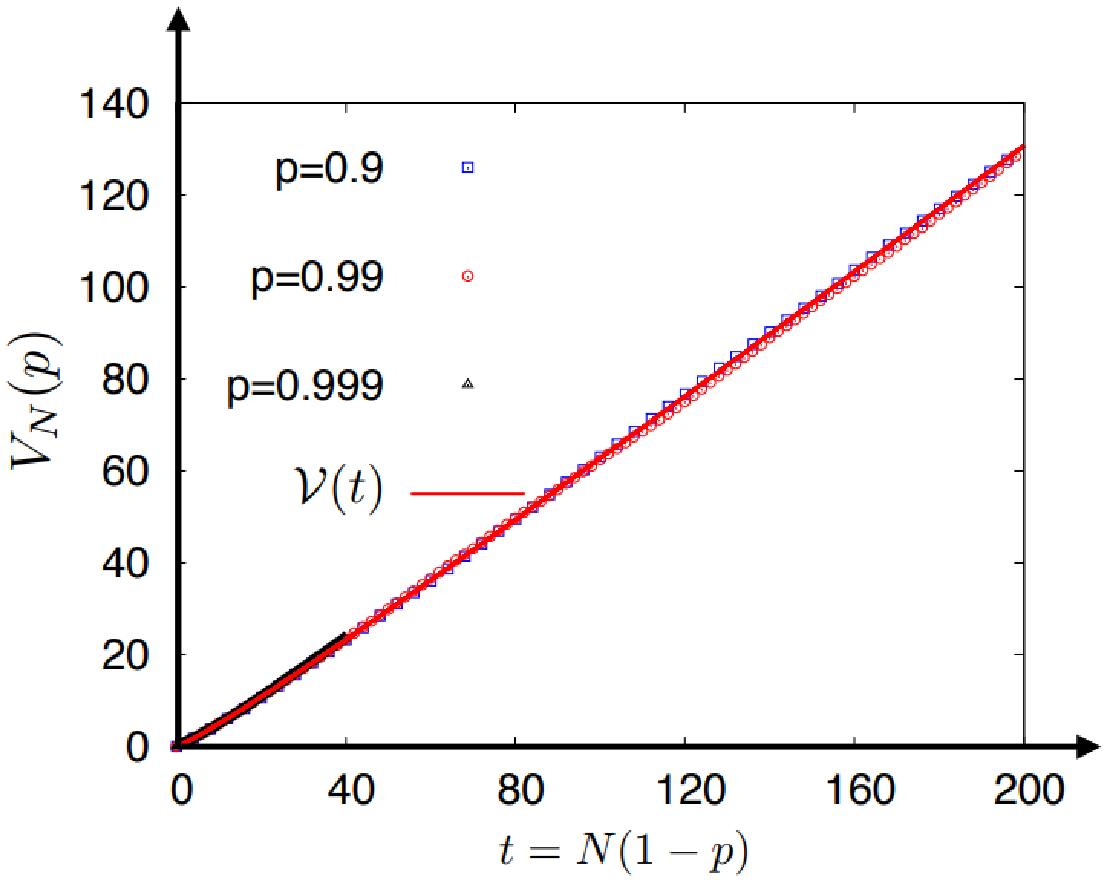
- Relation between Scaling functions $\mathcal{E}(t), S(t)$
  - $\mathcal{E}(t) = 1 + \int_0^t S(\tau)d\tau$
  - Exact relation : $\langle R_N \rangle(p) - 1 = \sum_{m=1}^{n} q_p(m)$

- Variance and the Fano Factor in the Scaling limits
  - $V_N(p) \simeq \mathcal{V}[N(1-p)]$ as $\mathcal{V}(t) = 2(t+1) - \mathcal{E}(t)(\mathcal{E}(t)+1)$

3. Correlation and Fano Factor in the Continuous Time Limit

- Variance and the Fano Factor in the Scaling limits
  - $V_N(p) \simeq \mathcal{V}[N(1-p)]$ as $\mathcal{V}(t) = 2(t+1) - \mathcal{E}(t)(\mathcal{E}(t)+1)$
  - $\mathcal{V}(t) = \begin{cases} \frac{t}{2} \quad , & \text{as} \quad t \to 0 \\ 2\left(1 - \frac{2}{\pi}\right)t \quad , & \text{as} \quad t \to +\infty \end{cases}$

3. Correlation and Fano Factor in the Continuous Time Limit



Plot of $\langle V_N \rangle(p)$

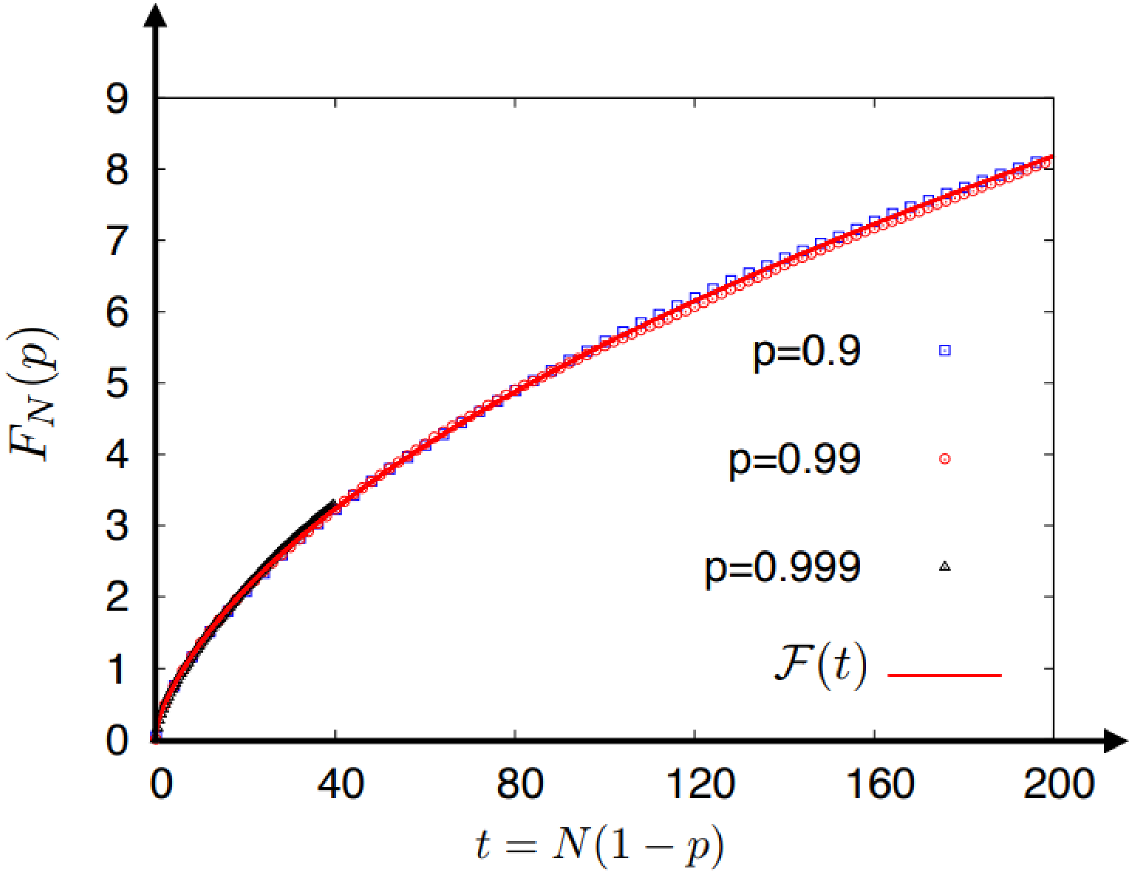
Symbols from simulation with different $p \to 1$ and $N = 40000$

Solid Line is $V_N(p) \simeq \mathcal{V}[N(1-p)]$

3. Correlation and Fano Factor in the Continuous Time Limit

- Variance and the Fano Factor in the Scaling limits
  - $F_N(p) \simeq \mathcal{F}[N(1-p)]$ as $\mathcal{F}(t) = \frac{2(t+1)}{\mathcal{E}(t)} - \mathcal{E}(t) - 1$
  - $$\mathcal{F}(t) = \begin{cases} \frac{t}{2} \quad , & \text{as} \quad t \to 0 \\ \left(1 - \frac{2}{\pi}\right)\sqrt{\pi t} \quad , & \text{as} \quad t \to +\infty \end{cases}$$

3. Correlation and Fano Factor in the Continuous Time Limit



Plot of $\langle F_N \rangle(p)$

Symbols from simulation with different $p \to 1$ and $N = 40000$

Solid Line is $F_N(p) \simeq \mathcal{F}[N(1-p)]$

Summary and Conclusion

- For zero position probability $p \neq 0$, additional *negative* correlations exist.
- In the limit $p \to 0$, well known random walk model (Markov Jump Process).
- In the limit $p \to 1$, continuous time random walk, CTRW model with exponential waiting-time distribution.

Q&A