

# Finding different regimes of Korean stock market and Seoul-si housing market via record statistics with drift

Sejin Lim, Soon-Hyung Yook

Department of Physics, Kyung Hee University

**2021 KPS Fall Meeting**

2021.10.20-22



# Introduction

- Financial market 은 방대한 데이터의 양 덕분에 복잡계의 일부분으로써 통계물리학에서 연구되어 오고 있음
- 일반적으로 경제 시스템에서의 보편적 특징들을 연구 대상으로 함
  - 개인이나 기업의 소득이나 주식 거래 시장과 외환시장의 리턴 에 대한 fat-tailed 분포  
[V. Plerou et al., Phys. Rev. E **60**, 6519 (1999)]
  - 가격 변화에서의 high-order 상관관계  
[Y. Liu et al., Phys. Rev. E **60**, 1390 (1999)]
- 주식 시장과 주택 시장은 GDP, 인플레이션율, 세금 정책 등 다양한 외부 요인들과 상호작용을 하고 다른 금융 시장들과 공통된 보편적 특성들을 공유함 [J. Aizenman, Y. Jinjark, H. Zheng, VoxEU. org **24** (2016)]
- 이런 관점에서 국내의 주식 시장과 주택시장의 record statistics 를 이해하는 것은 글로벌 금융 시스템에 주는 영향을 조사하는 데 중요
- 이 연구에서는 한국 주식 시장과 서울시 주택 시장 데이터를 record statistics 와 inter-record time distribution 으로 분석하고자 함

# Introduction : Record

- 전체 데이터 개수가  $N$  개인 시계열 데이터 :  $\{x_1, x_2, \dots, x_N\}$
- 특정 시점  $m$  에서 발생하는 record 의 정의 :  $x_m > \max\{x_1, x_2, x_3, \dots, x_{m-1}\}$
- Binary record indicator :  $\sigma_m = \begin{cases} 1, & \text{record happens at step } m \\ 0, & \text{otherwise} \end{cases}, \quad \sigma_1 = 1 \text{ (by convention)}$
- Record number ( $N$  시간 동안 record 발생 횟수) :  $R_N = \sum_{m=1}^N \sigma_m$
- 평균 record 발생 횟수 :  $\langle R_N \rangle = \sum_{m=1}^N \langle \sigma_m \rangle$
- Record rate ( $m$  시점의 record 발생 비율) :  $r_m = \langle \sigma_m \rangle$
- Random walk variables  $\{x_m\}$  의 예 :
  - $x_m = x_{m-1} + \eta_m + c$ ,  $\eta_m$  는 대칭적인 분포  $\varphi(\eta)$  를 따르는 random variable,  $c = (\text{constant drift})$
  - 분포  $\varphi(\eta)$  의 종류 [Satya N. Majumdar and Robert M. Ziff., Phys. Rev. Lett. **101**, 505601 (2008), Satya N. Majumdar., Physica A **389**, 4299 (2010)]
    - Uniform distribution
    - Gaussian distribution
    - Cauchy distribution
    - Lévy alpha-stable distribution

# Introduction : Survival prob. and First-passage prob.

- Survival probability :  $q(m) = \text{Prob}(x_1 > 0, x_2 > 0, \dots, x_m > 0 | x_0 = 0)$ 
  - $q(m) = \text{Prob}(x_{m-1} > \max\{x_m\}, \dots, x_2 > \max\{x_m, x_{m-1}, \dots, x_3\}, x_1 > \max\{x_m, x_{m-1}, \dots, x_2\} | x_m = x_m)$
  - 위의 random walk 의 시간 대칭적 성질에 의해  $q(m) = r_m$  [Satya N Majumdar et al, J. Phys. A: Math. Theor. **54** 315002 (2021)]
  
- First-passage probability :  $F(x|0, m)$  : 원점에서 시작하여  $x$  위치에  $m$  시점에 처음으로 도달할 확률
  - $F(x|0, m) = -\frac{\partial}{\partial m} q(m)$
  - $f(m) \equiv F(x|0, m)$
  
- First-passage probability 와 inter-record time distribution :
  - $\tau$  : 인접한 record event 가 발생하는 시간 간격 (inter-record time)
  - $P(\tau)$  : Inter-record time distribution
  - Random variable  $x_m$  의 jump length  $\eta_m$  가 대칭적인 분포  $\varphi(\eta)$  를 따르면,  $F(x|0, m) = F(x|x, m) = F(0|0, m)$
  - $P(\tau)$  는 임의의 시점  $m$  에서 record event가 발생하고  $\tau$  시간 동안 일어나지 않다가  $m + \tau$  시간에 record event가 발생할 확률과 같음
  - $P(\tau) = F(x|x, \tau) = f(\tau)$

# Lévy alpha-stable distribution with drift

- $x_m = x_{m-1} + \eta_m + c$ ,  $c = (\text{constant drift})$
- jump length  $\eta$  의 분포  $\varphi(\eta)$  이 Lévy alpha-stable distribution 인 경우
- $\varphi(\eta) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi(k) e^{-i\eta k} dk$
- $\varphi(k; \alpha, \beta, \gamma, \delta) = \exp(ik\delta - |\gamma k|^\alpha (1 - i\beta \operatorname{sgn}(k) \Phi))$ , Lévy exponent :  $\alpha$
- $\Phi = \begin{cases} (|\gamma k|^{1-\alpha} - 1) \tan\left(\frac{\pi\alpha}{2}\right) & \alpha \neq 1 \\ -\frac{2}{\pi} \log|\gamma k| & \alpha = 1 \end{cases}$

**Drift regime table**

	Regime 1 $0 < \alpha < 1$ $c$ is arbitrary	Regime 2 $\alpha = 1$ $c$ is arbitrary	Regime 3 $1 < \alpha < 2$ $c > 0$	Regime 4 $\alpha = 2$ $c > 0$	Regime 5 $1 < \alpha \leq 2$ $c < 0$
$q_m$	$B_1 m^{-0.5}$	$B_2 m^{-\theta(c)}$	$B_3 m^{-\alpha}$	$B_4 m^{-1.5} \exp\left(-\frac{c^2}{2\sigma^2} m\right)$	$k_\alpha( c )$
$P(\tau)$	$B_1 \tau^{-1.5}$	$B_2 \tau^{-1-\theta(c)}$	$B_3 \tau^{-1-\alpha}$	$B_4 \exp\left(-\frac{c^2}{2\sigma^2} \tau\right) \left(\frac{c^2}{2\sigma^2} \tau^{-1.5} + 1.5 \tau^{-2.5}\right)$	uniform
$R_m$	$A_1 m^{0.5}$	$\frac{A_2}{\Gamma(1+\theta(c))} m^{\theta(c)}$	$k_\alpha(c) m$	$k_2(c) m$	$1/k_\alpha( c )$

# Lévy alpha-stable distribution with drift

Table of coefficients for each regime

	Regime 1 $0 < \alpha < 1$ $c$ is arbitrary	Regime 2 $\alpha = 1$ $c$ is arbitrary	Regime 3 $1 < \alpha < 2$ $c > 0$	Regime 4 $\alpha = 2$ $c > 0$	Regime 5 $1 < \alpha \leq 2$ $c < 0$
$A_{regime}$	$\frac{2}{\sqrt{\pi}} \exp \left[ \frac{1}{\pi} \int_0^\infty \frac{dk}{k} \arctan \left( \frac{\hat{f}(k) \sin(kc)}{1 - \hat{f}(k) \cos(kc)} \right) \right]$	$\frac{1}{\Gamma(1 - \theta(c)) B_2}$	—	—	—
$B_{regime}$	$\frac{2}{\pi A_1}$	$\frac{e^{-\gamma_0}}{\Gamma(1 - \theta(c))}$	$\frac{\alpha - 1}{2 \cos(\frac{\alpha\pi}{2}) \Gamma(2 - \alpha) c^\alpha} \exp \left[ \sum_{m=1}^\infty \frac{1}{m} \int_{cm}^\infty P_m(x) dx \right]$	$\frac{\sigma}{c\sqrt{2\pi}} \exp \left[ \sum_{m=1}^\infty \frac{e^{\frac{c^2}{2\sigma^2} m}}{m} \int_{cm}^\infty P_m(x) dx \right]$	—
$k_\alpha(c)$	—	—	$\exp \left[ - \sum_{m=1}^\infty \frac{1}{m} \int_{cm}^\infty P_m(x) dx \right]$		

[SN Majumdar, J.Phys. A: Math. Theor. **45**, 355002 (2012)]

$$\theta(c) = \frac{1}{2} + \frac{1}{\pi} \arctan(c)$$

$$\gamma_0 = \sum_{m=1}^\infty \left[ 1 - \theta(c) - \int_{cm}^\infty P_m(x) dx \right], \quad P_m(x) = \text{Prob}[x_m < -cm]$$

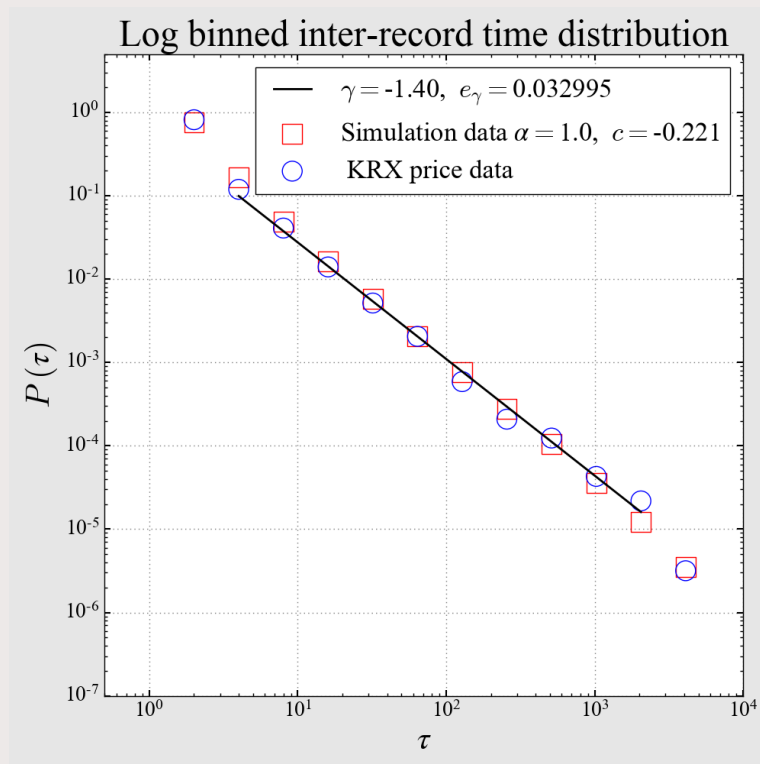
# KRX and 부동산

- 한국 주식 시장(KRX) 일별 고가 데이터
  - 2000년 01월 04일 ~ 2020년 04월 02일 동안 상장되어 있는 주식들의 일별 거래 데이터
- 서울시 부동산 일별 거래 데이터
  - 2006년 01월 01일 ~ 2017년 12월 31일 동안의 행정동 별 일별 거래 데이터
- 해석적 이론과 비교를 위한 수치적 모형
  - $x_m = x_{m-1} + \eta_m + c$
  - $\eta_m$  는  $\varphi(\eta)$  를 따르는 random variable,  $c = (\text{constant drift})$ 
    - $\varphi(\eta) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi(k) e^{-i\eta k} dk$
    - $\varphi(k; \alpha, \beta, \gamma, \delta) = \exp(ik\delta - |\gamma k|^\alpha (1 - i\beta \operatorname{sgn}(k) \Phi))$ , Levy exponent :  $\alpha$
    - $\Phi = \begin{cases} (|\gamma k|^{1-\alpha} - 1) \tan\left(\frac{\pi\alpha}{2}\right) & \alpha \neq 1 \\ -\frac{2}{\pi} \log|\gamma k| & \alpha = 1 \end{cases}$
  - 샘플 개수 10,000개, 전체 step 수  $N = 5,000$
- 실제 시계열 데이터와 시뮬레이션 데이터로부터 record number 와 inter-record time distribution 을 계산 및 해석적 이론 값과의 비교를 통해 regime을 알아냄

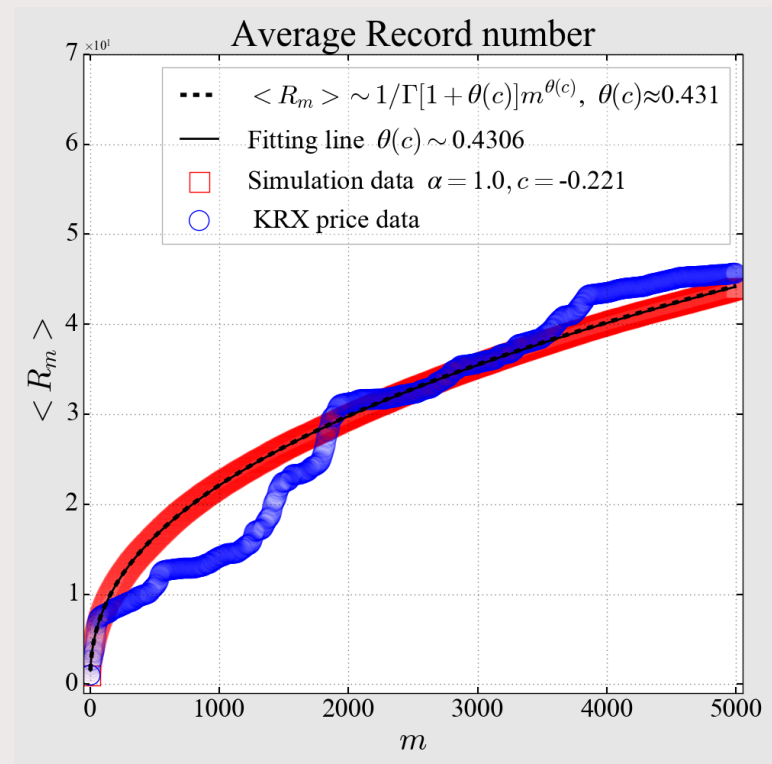
# KRX

## Regime 2

- $\theta(c) = \frac{1}{2} + \frac{1}{\pi} \arctan(c)$
- $\theta(c) \approx 0.4306, c \approx -0.221, \alpha = 1.0$



$P(\tau) \propto \tau^\gamma$ ,  $-\gamma \leq 1.5$  이므로 regime 3, 4, 5 제외 가능



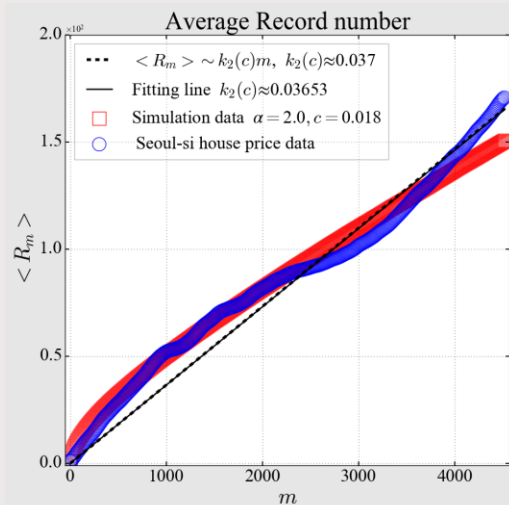
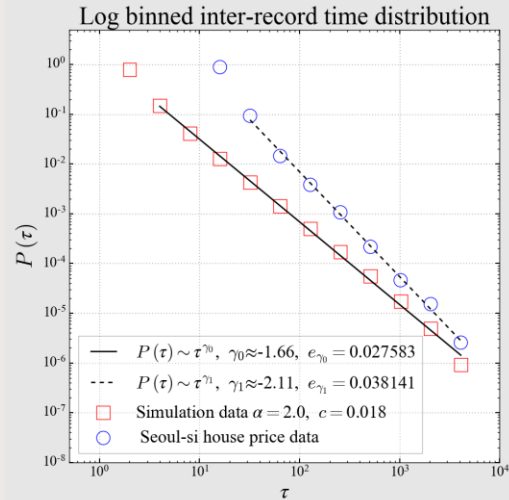
KRX price data를  $\frac{A_2}{\Gamma(1+\theta(c))} m^{\theta(c)}$  에 Fitting 하여  $c$  유추



# 서울시 주택 시장

## Regime 4 로 가정

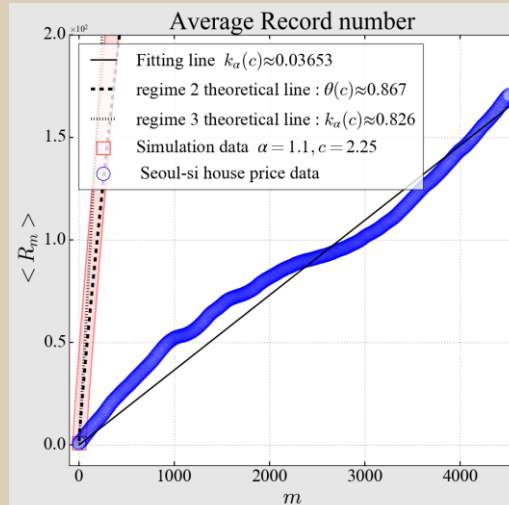
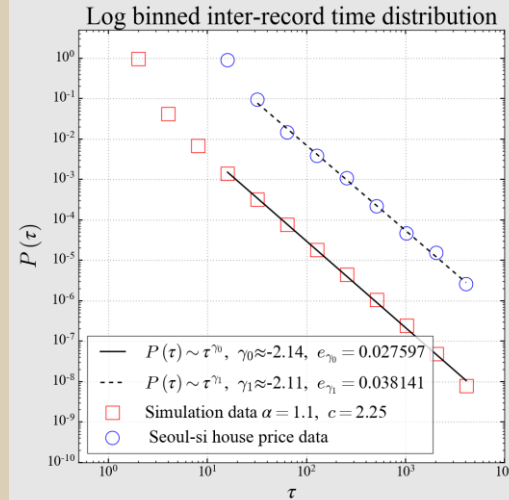
- $k_2(c) \approx 0.03653$ ,  $c \approx 0.018$ ,  $\alpha = 2.0$



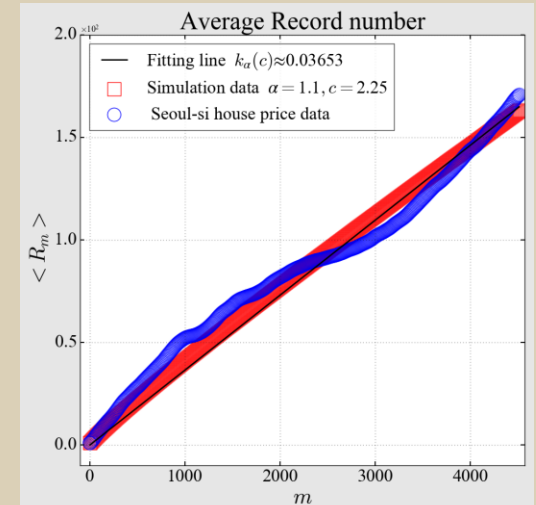
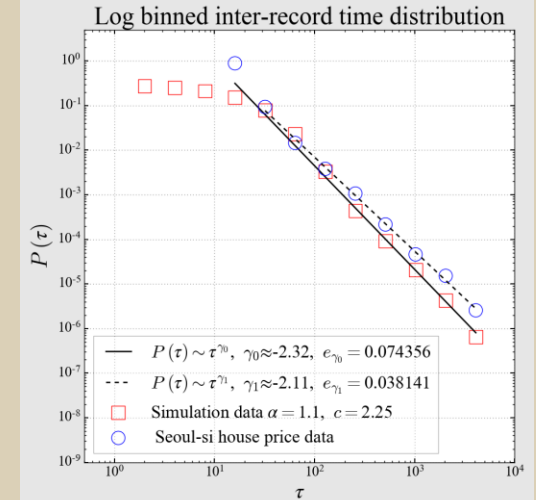
[Satya N Majumdar et al, J. Phys. A: Math. Theor. **54** 315002 (2021)]

## Regime 3

- $c = 2.250$ ,  $\alpha = 1.10$



- $x_m = x_{m-1} + \eta_m + c$ ,
- $\varphi(\eta) = p_0 \delta(\eta) + (1 - p_0) \varphi_0(\eta)$ ,
- $\varphi_0(\eta)$  : Lévy  $\alpha$ -stable distribution
- $p_0$  : jump length  $\eta$  가 0일 확률
- $c = 2.250$ ,  $\alpha = 1.10$ ,  $p_0 = 0.939$



# 요약 및 결론

- 한국 주식 시장과 서울시 주택 시장의 drift regime에 차이가 있음을 발견됨
- 한국 주식 시장은  $c = -0.221$ ,  $\alpha = 1.0$  인 Regime 2
- 서울시 주택 시장은  $c = 2.250$ ,  $\alpha = 1.10$  Regime 3
  - Regime 4 로 짐작 했으나 inter-record time distribution 의 exponent 값이 regime 4 와 상이함
  - Regime 3 에서 record number 가 잘 맞지 않는 이유는 서울시 주택 시장의 거래 빈도 수가 적기 때문으로 여겨 짐
  - Regime 3 에서 jump length 가 0일 확률  $p_0$  가 일자 별 거래할 확률 0.939와 같은 경우, record number 가 비슷한 경향성을 보임을 확인함
- Regime 2 에서는 drift 값이 record statistics에 영향을 크게 미치고 Regime 3 에서는 jump length의 분포에 대한 정보인  $\alpha$  값이 영향을 크게 미침