外国為替市場における参加者行動の網羅的計量

佐藤彰洋
京都大学大学院情報学研究科
数理工学専攻

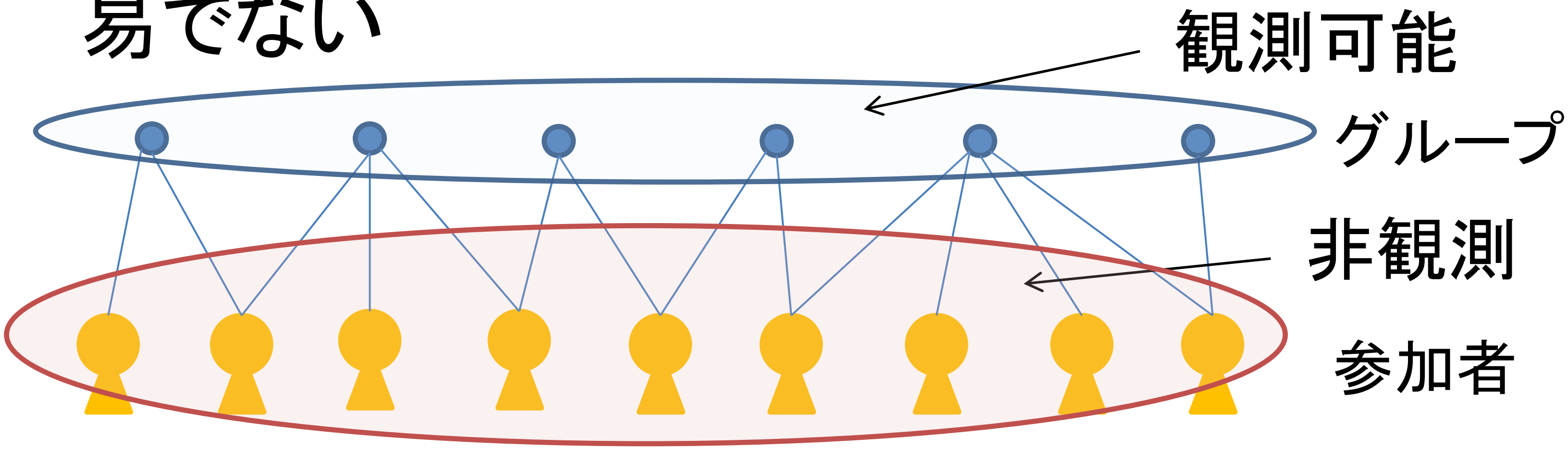
アウトライン

- 動機と目的
- 統計力学的エントロピー
- 有限混合Gumbel分布
- 結論

arXiv:1207.4860v3 [physics.data-an] 25 Jul 2012

動機

- 我々の観測は常に**部分的**である
- 技術的・法律的な理由のため、ネットワーク構造に関するすべての情報を集めることは容易でない



社会グループ：グループノードと参加者ノードからなる二部グラフ

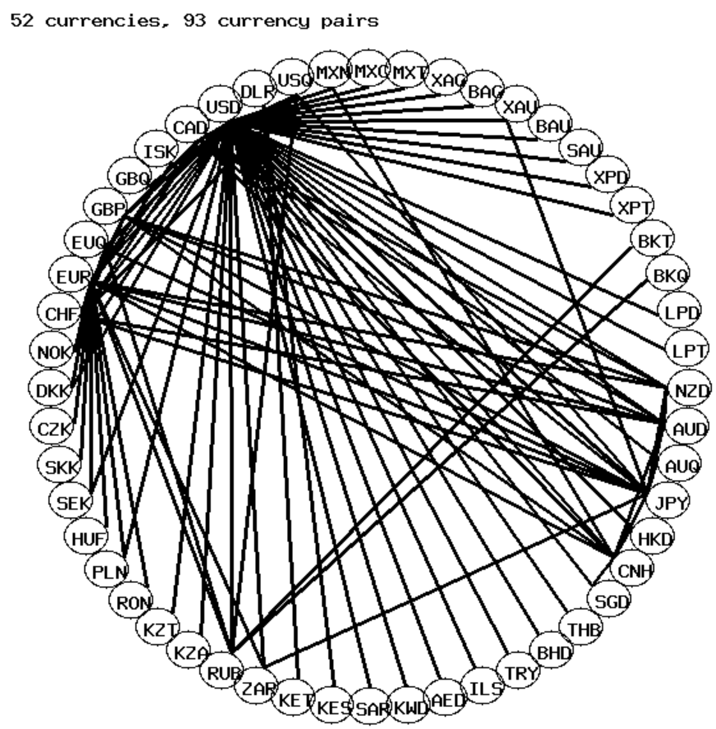
外国為替市場の高解像度データ

- 電子ブローキングシステムに蓄積されているデータ
- ICAP EBS 社 Data Mine Level 1.0
- 各レコードは以下の内容を含む
  - 時間スタンプ（1 秒解像度）
  - 注文 (P)/取引 (D)
  - 通貨ペア
  - 注文価格（ビッド, アスク), 取引価格

ICAP platformで収集されたデータを用いた　https://www.icap.com

データ

Data provider：ICAP EBS
period：1st June 2008-31st July 2012
records：520,023,124



41 currencies, 9 precious metals and 2 basket currency

AED: United Arab Emirates Dirham
AUD : Australia Dollar
AUQ: Australia Dollar (small amount)
BAG: Gold (bank)
BAU: Silver (bank)
BHD: Bahraini Dinar
BKT : Basket of USD/EUR
BKQ: Basket of USD/EUR (small amount)
CAD : Canadian Dollar
CHF : Swiss Fran
CNH: Chinese Yuan
CZK : Czech Koruna
DKK : Danish Krone
EUR : EU Euro
EUQ: EU Euro (small amount)
GBP : British Sterling
GBQ: British Sterling (small amount)
HKD : Hong-Kong Dollar
HUF : Hungarian Forint
ILS: Isreali New Shekel
ISK : Iceland Krona
JPY : Japanese Yen
KES: Kenyan Shilling
KET: Kenyan Shilling (small amount)
KZA: Kazakhstan Tenge (small amount)
KZT: Kazakhstan Tenge
LPD: Palladium (London)
LPT: Platium (London)
MXN : Mexico Peso
MXQ: Mexico Peso (small amount)
MXT: Mexico Peso (special deals)
NOK : Norwegian Krone
NZD : New Zealand Dollar
PLN : Poland Zloty
RON: Romania Leu
RUB : Russian Ruble
SAR: Saudi Arabian Riyal
SGD : Singapore Dollar
SEK :Swedish Krona
SKK : Slovak Koruna
SAU: Silver (small amount)
TRY: Turkish Lira
THB: Thai Baht
USD/DLR : USA Dollar
USQ: USA Dollar (small amount)
ZAR : South African Rand
XAU : Gold
SAU: Gold (small amount)
XAG : Silver
XPD : Palladium
XPT : Platinum

目的

- たくさんのネットワーク構造の中から特異的な事象を表現するネットワークを自動的に探し出したい
- 更には、特異的なネットワークが発現する確率はどの程度か推計したい？
- ネットワークの知見から、国際経済環境のリスクファクターを開発したい

統計力学的エントロピー

- 統計力学では、可能な状態の組み合わせの数が“エントロピー”と関係する

- Boltzmann エントロピー

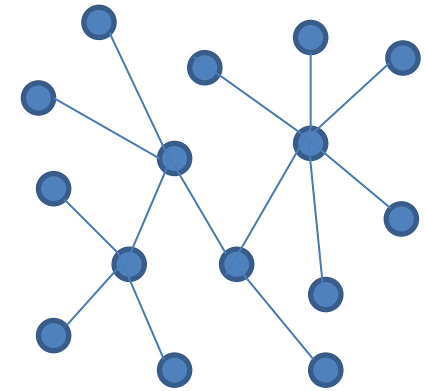
$$S = k_B \ln N$$

Boltzmann定数　　　　可能な状態数

ネットワークエントロピー

- Bianconi ら：いくつかの拘束条件のもとで実現可能なネットワーク構造の組合せ数Nの対数をネットワークエントロピーを定義

$$\Sigma = \ln N$$



G. Bianconi, Physical Revew E, **79** (2009) 036114

K. Anand and G. Bianconi, Physical Review E, **80** (2009) 045102

二部グラフ構造の情報統計力学的ネットワークエントロピー

- 二部グラフ構造として社会システムを表現する場合、しばしば全ての参加者の状態を完全に観測可能できない場合がある
- 統計力学的ネットワークエントロピーを観測可能な量を拘束条件として計算することはネットワーク構造の複雑性を計量することと関係すると期待される
- 二部グラフネットワーク構造のノードAで観測されるリンク数を拘束条件として統計力学的ネットワークエントロピーを計算する

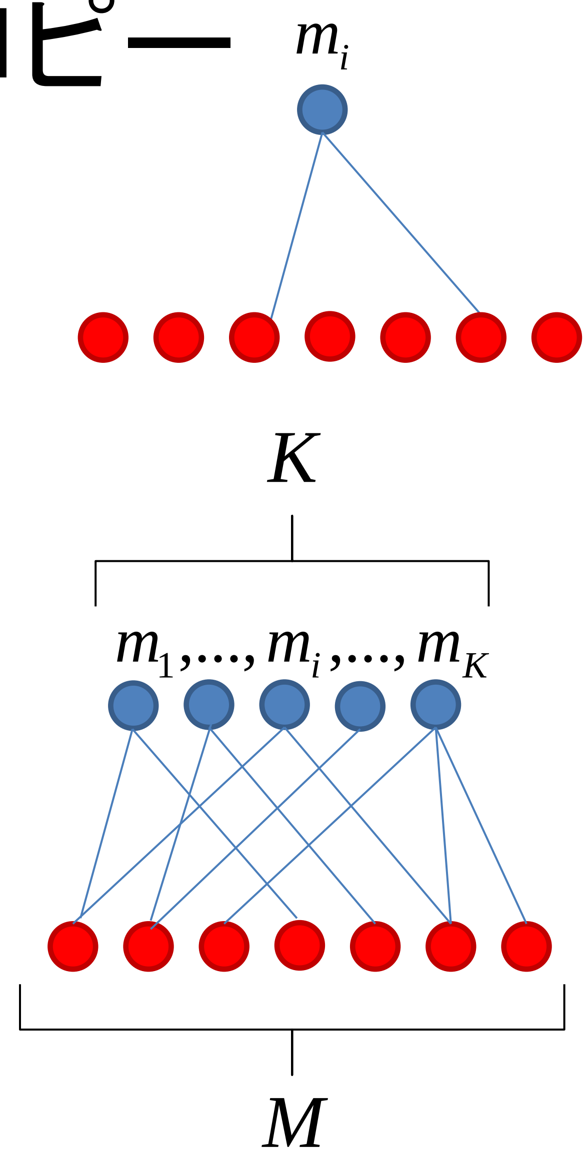
ネットワークエントロピー

- 構成可能なネットワークの数:

$$N=\prod_{i=1}^{K}\boxed{\binom{M}{m_i}}=\prod_{i=1}^{K}\boxed{\frac{M!}{(M-m_i)!\,m_i!}}$$

- ネットワークエントロピー:

$$\begin{aligned}\Sigma &= \ln N\\ &=\ln\left(\prod_{i=1}^{K}\frac{M!}{(M-m_i)!\,m_i!}\right)\\ &=\sum_{i=1}^{K}\ln\frac{M!}{(M-m_i)!\,m_i!}\\ &=K\sum_{n=1}^{M}\ln n-\sum_{i=1}^{K}\sum_{n=1}^{M-m_i}\ln n-\sum_{i=1}^{K}\sum_{n=1}^{m_i}\ln n\end{aligned}$$



単位リンク当たりのネットワークエントロピー

- 1リンクあたりのネットワークエントロピー

$$\sigma = \frac{\Sigma}{\sum_{i=1}^{K} m_i}$$

$$= \frac{1}{\sum_{i=1}^{K} m_i}\left(K\sum_{n=1}^{M} \ln n - \sum_{i=1}^{K} \sum_{n=1}^{M-m_i} \ln n - \sum_{i=1}^{K} \sum_{n=1}^{m_i} \ln n \right)$$

σ 大=>二部グラフ構造は一様

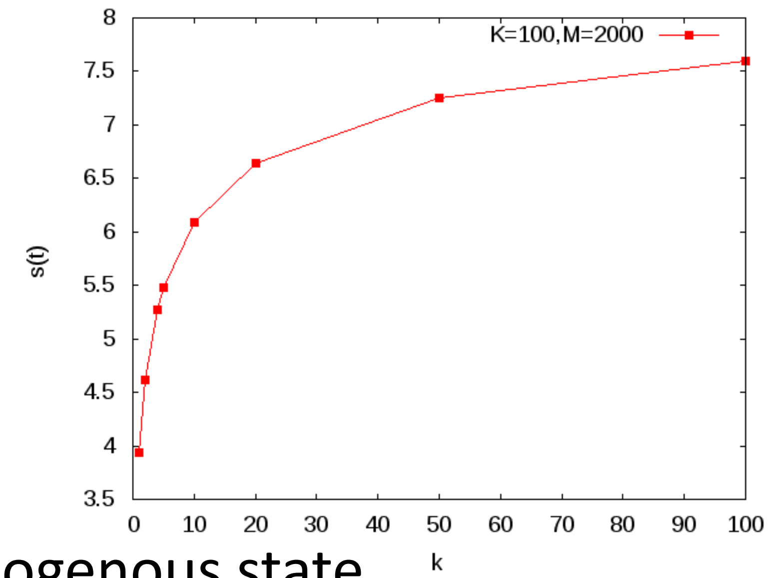
σ 小=>二部グラフ構造は非一様

Some simple examples(1)

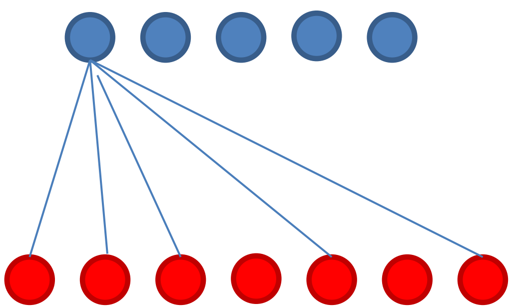
$$K = 100, M = 2000$$

$$m_i = \begin{cases} 100/k & (i = 1, \ldots, k) \\ 0 & (i = k+1, \ldots, K) \end{cases}$$

$$k = 1,2,4,5,10,20,50,100$$

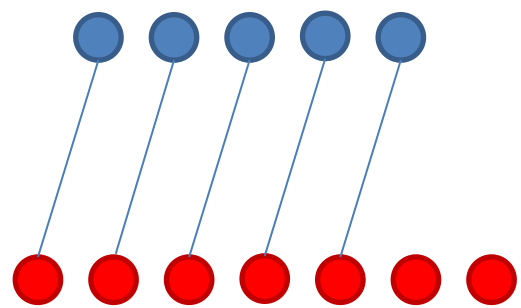


Monopoly state



k=1 ⇔ σ is small

Homogenous state



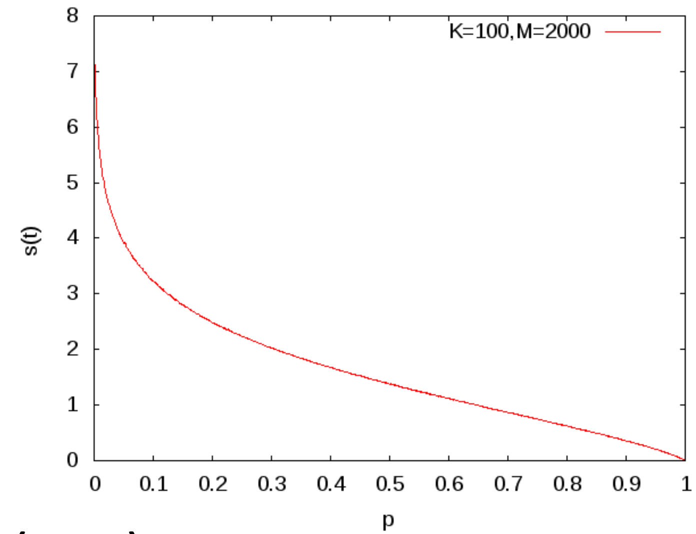
k=100 ⇔ σ is large

Some examples(2)

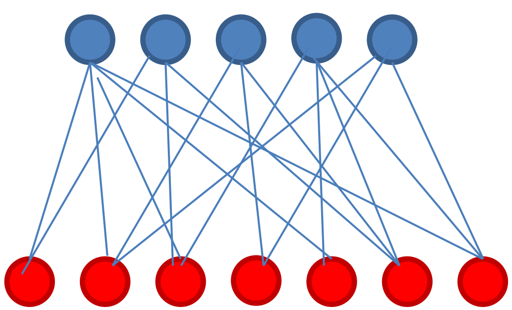
$$K = 100, M = 2000 \ \ (0 < \mathrm{p} \le 1)$$

$$C_{ij} = \begin{cases} 1 & w.p. \quad p \\ 0 & w.p. \quad 1-p \end{cases}$$

$$m_i = \sum_{j=1}^{M} C_{ij}$$

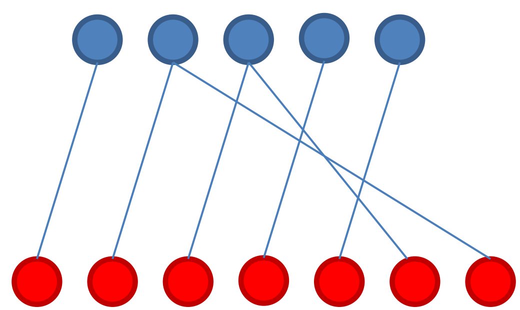


Dense state (p~1)



k=1 ⇔ σ is small

Diluted state (p~0)

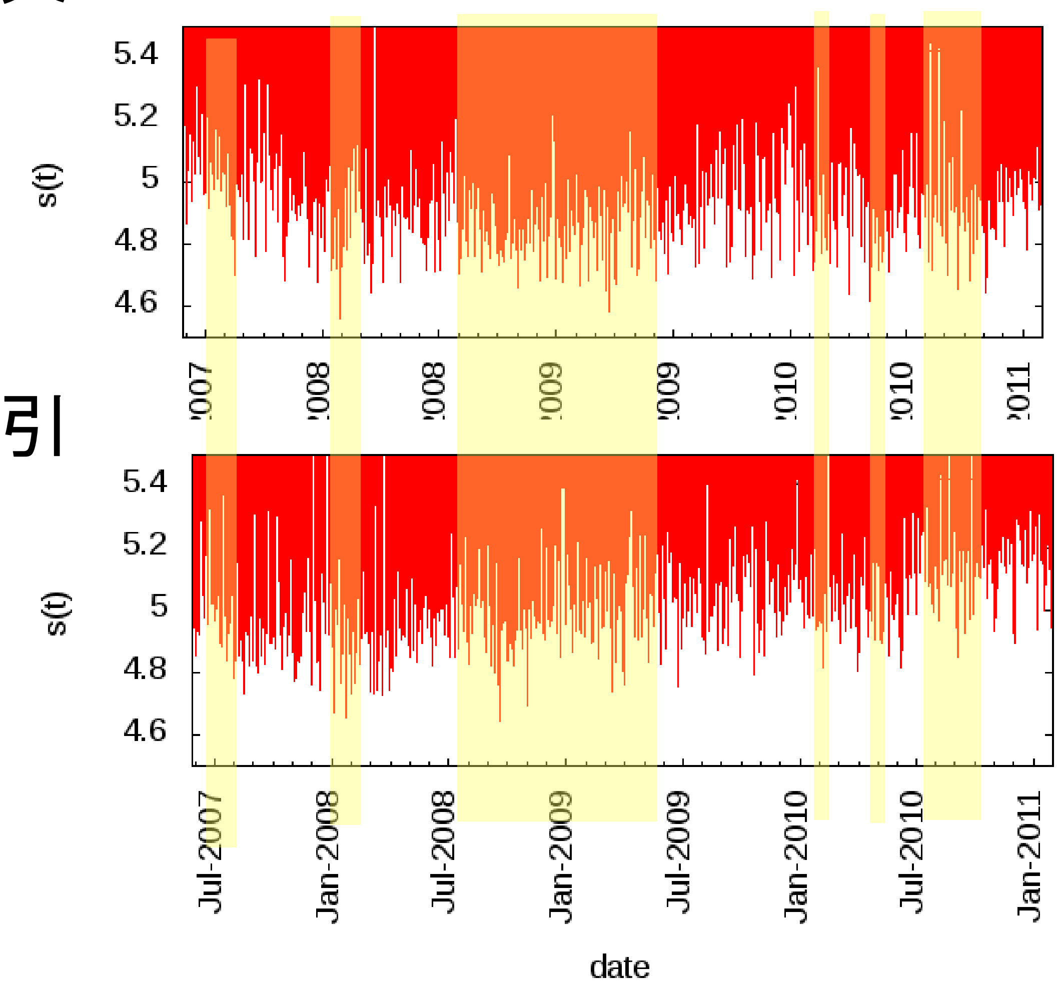


k=100 ⇔ σ is large

Long term investigation

注文

取引



2007年8月:
パリバショック
2008年2月:
ベアースターンズ危機
2008年7月～2009年2月:
世界同時金融危機
2009年11月:
ドバイショック
2010年1月:
ギリシャ危機
2010年4月～5月:
ユーロショック

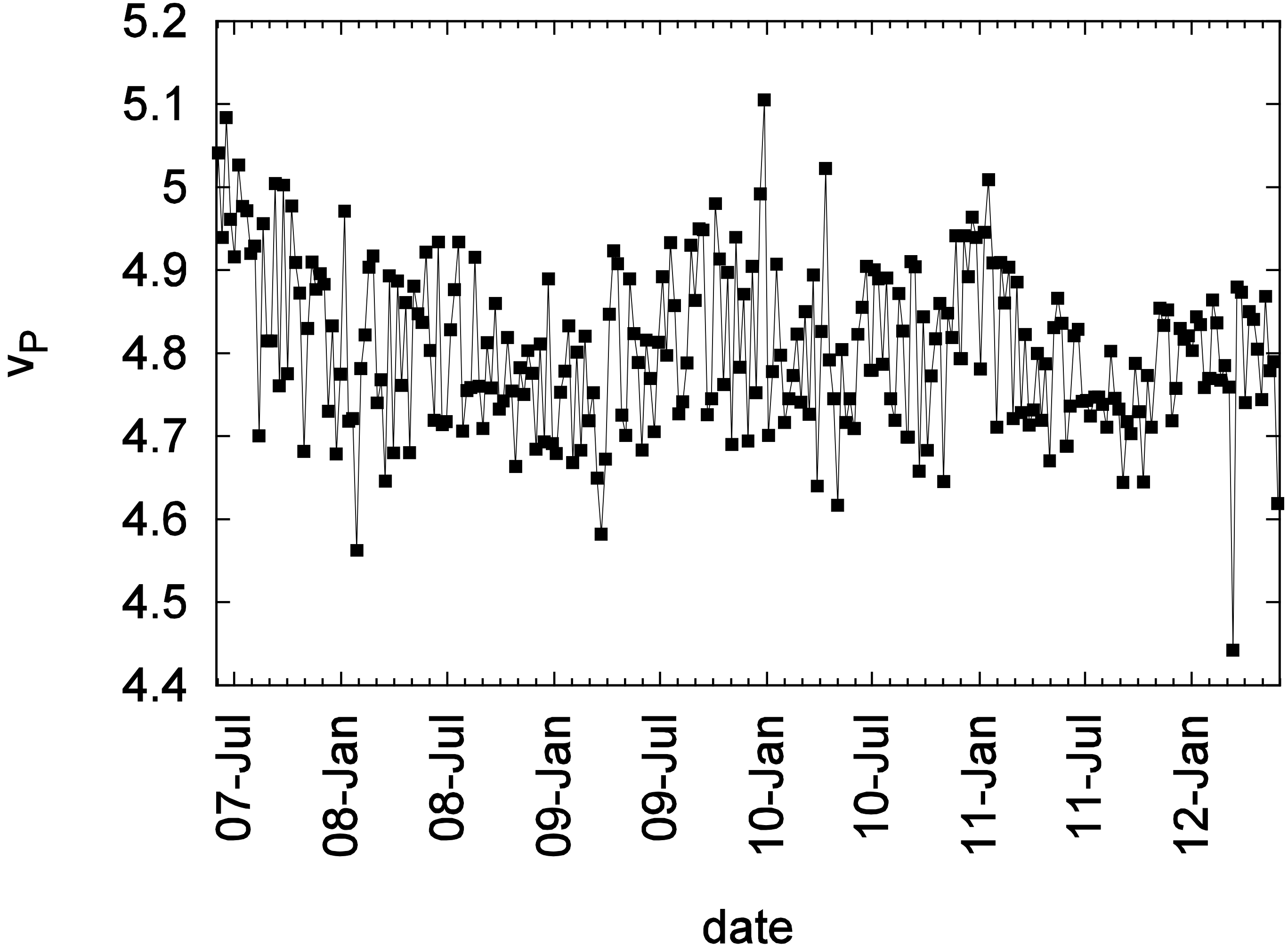
最小値

- 一週間ごとの最小値を考える

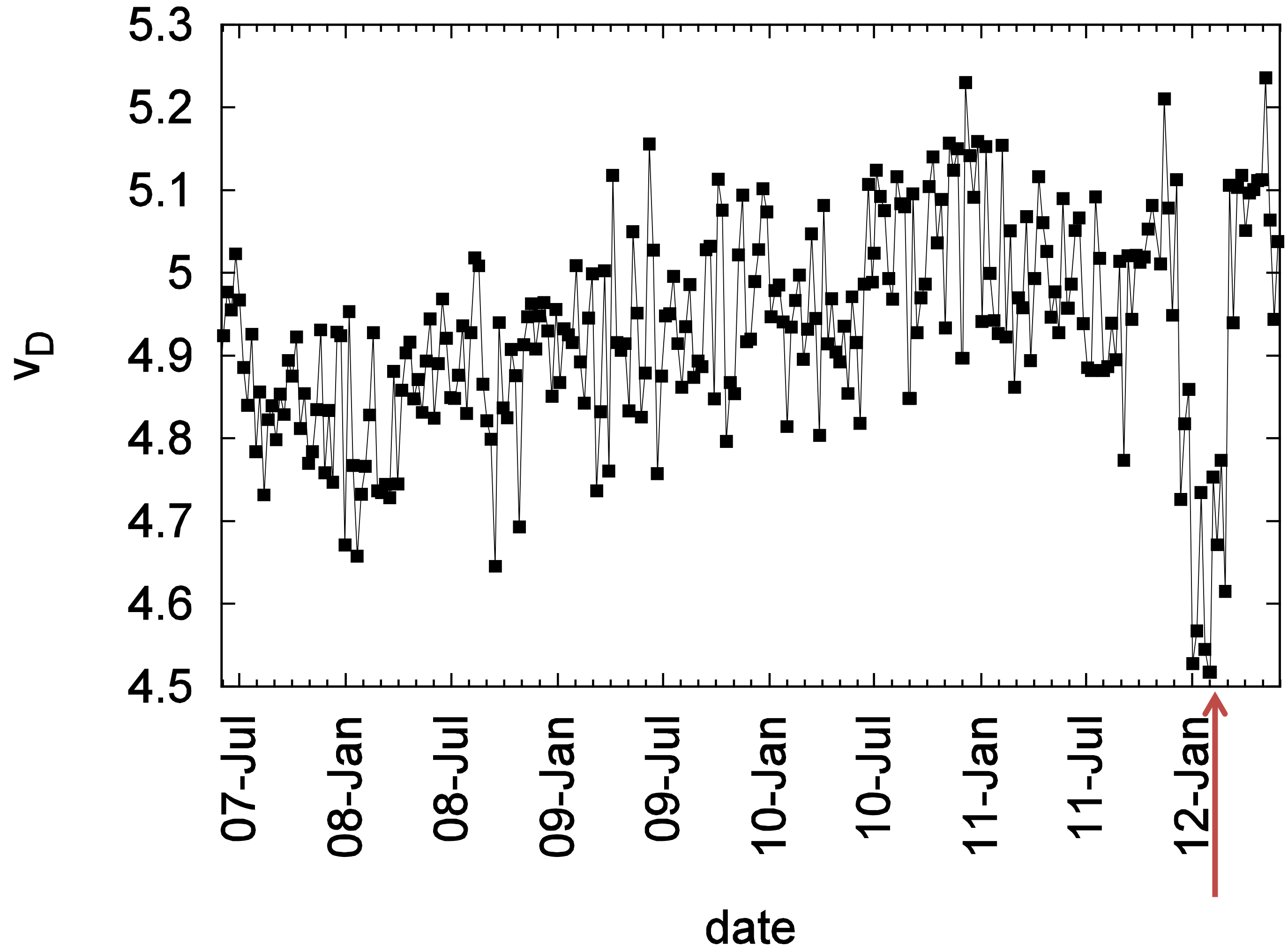
$$v_X(s) = \min_{t \in W(s)} \left\{ \sigma_X(t) \right\}$$

ここで、X=P (注文), X=D(取引)とする

注文


5.2
5.1
5
4.9
4.8
4.7
4.6
4.5
4.4
v_P
07-Jul
08-Jan
08-Jul
09-Jan
09-Jul
10-Jan
10-Jul
11-Jan
11-Jul
12-Jan
date

取引



5.3
5.2
5.1
5
4.9
4.8
4.7
4.6
4.5
v_D
07-Jul
08-Jan
08-Jul
09-Jan
09-Jul
10-Jan
10-Jul
11-Jan
11-Jul
12-Jan
date

まとめ

- 統計力学的エントロピーに基づいた二部グラフ構造を特徴づけるための方法を提案
- 外国為替市場の二部グラフ構造は金融危機やショックの前後において小さい値となる
- 特に、ベアスターンズショック、リーマンショック、ユーロ危機において注文または取引の単位リンクあたりのエントロピーの値は小さい
- <u>2012年2月14日に決定された日銀金融緩和以前に極端なネットワーク構造が連続して取引系列で確認される</u>

極値分布

- 最小値のGumbel分布

$$\Pr\left[V_X \le v_X\right] = 1 - \exp\left[-e^{\frac{v_X+\mu_X}{\rho_X}}\right]$$

$$P(v_X;\mu_X,\rho_X) = \frac{1}{\rho_X}\exp\left[\frac{v_X+\mu_X}{\rho_X} - \exp\left(\frac{v_X+\mu_X}{\rho_X}\right)\right]$$

最尤法

- Gumbel分布のパラメータを最尤法により推定

$$L_1 = \sum_{s'=1}^{T} \log P\left(v_X(s'); \mu_X, \rho_X\right)$$

$$= -T \log \rho_X + \sum_{s'=1}^{T} \frac{v_X(s') + \mu_X}{\rho_X} - \sum_{s'=1}^{T} \exp\left[\frac{v_X(s') + \mu_X}{\rho_X}\right]$$

$$\left\{\hat{\mu}_X, \hat{\rho}_X\right\} = \arg\max_{\mu,\rho} L_1$$

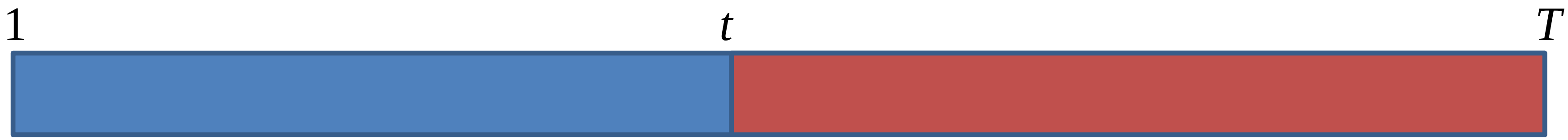
T=268 [weeks]

推定値

X	μ_X	ρ_X	KS-value	p-value
P(注文)	-4.85	0.101	1.72	0.005
D(取引)	-4.99	0.117	1.47	0.026

5%有意水準でKolmogorov-Smirnov検定を合格しない

時系列分割



$$L_1 = \sum_{s'=1}^{T} \log P\left(v_X(s'); \hat{\mu}, \hat{\rho}\right)$$

$$L_2(s) = \sum_{s'=1}^{s} \log P\left(v_X(s'); \hat{\mu}_L, \hat{\rho}_L\right) + \sum_{s'=s+1}^{T} \log P\left(v_X(s'); \hat{\mu}_R, \hat{\rho}_R\right)$$

$$\Delta(s) = L_2(s) - L_1 \geq 0$$

$$s^* = \arg\max_s \Delta(s)$$

S.A. Cheong et al. (2012), S.M. Goldfeld and R.E. Quandt (1973)

対数尤度差とJensen-Shannon divergence

$$L_1 = \sum_{s'=1}^{T} \log P\left(v_X(s'); \mu, \rho\right)$$

$$L_2(s) = \sum_{s'=1}^{s} \log P\left(v_X(s'); \mu_L, \rho_L\right) + \sum_{s'=s+1}^{T} \log P\left(v_X(s'); \mu_R, \rho_R\right)$$

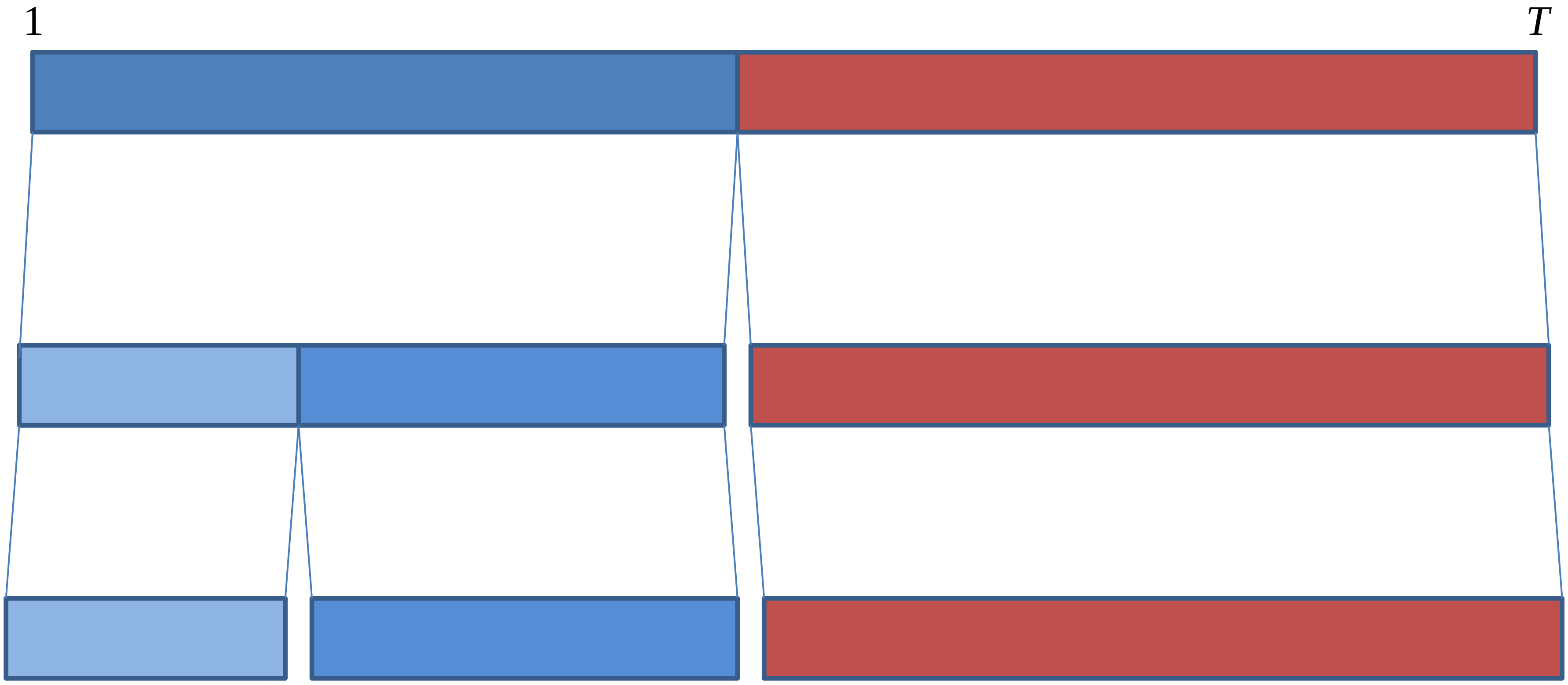
$$\begin{aligned}
\Delta(s) &= L_2(s) - L_1 \\
&\approx TH\left[P(v_X; \mu, \rho)\right] - tH\left[P(v_X; \mu_L, \rho_L)\right] - (T-t)H\left[P(v_X; \mu_R, \rho_R)\right] \\
&= TJS_{\frac{s}{T}, \frac{T-s}{T}}\left[P\left(v_X; \mu_L, \rho_L\right), P\left(v_X; \mu_R, \rho_R\right)\right]
\end{aligned}$$

Jensen-Shannon divergence

$$H\left[P\left(v_X; \mu, \rho\right)\right] = -\int_{-\infty}^{\infty} P(v_X; \mu, \rho) \log P(v_X; \mu, \rho) \mathrm{d}v_X$$

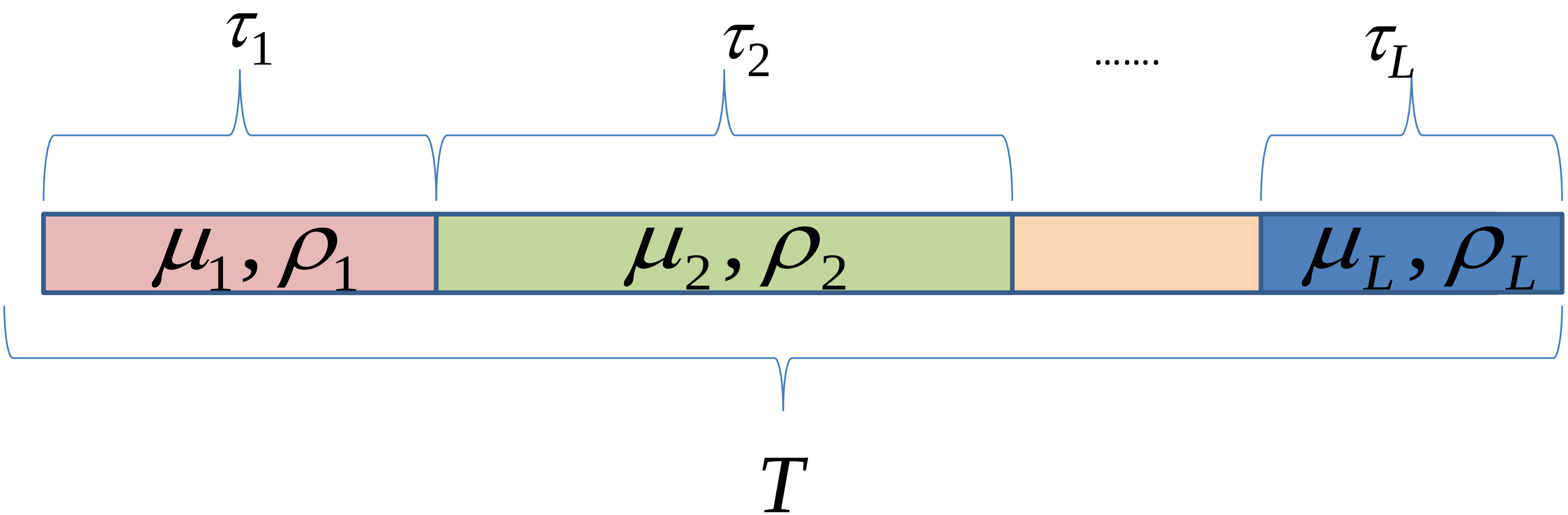
Shannon entropy

再帰的時系列分割

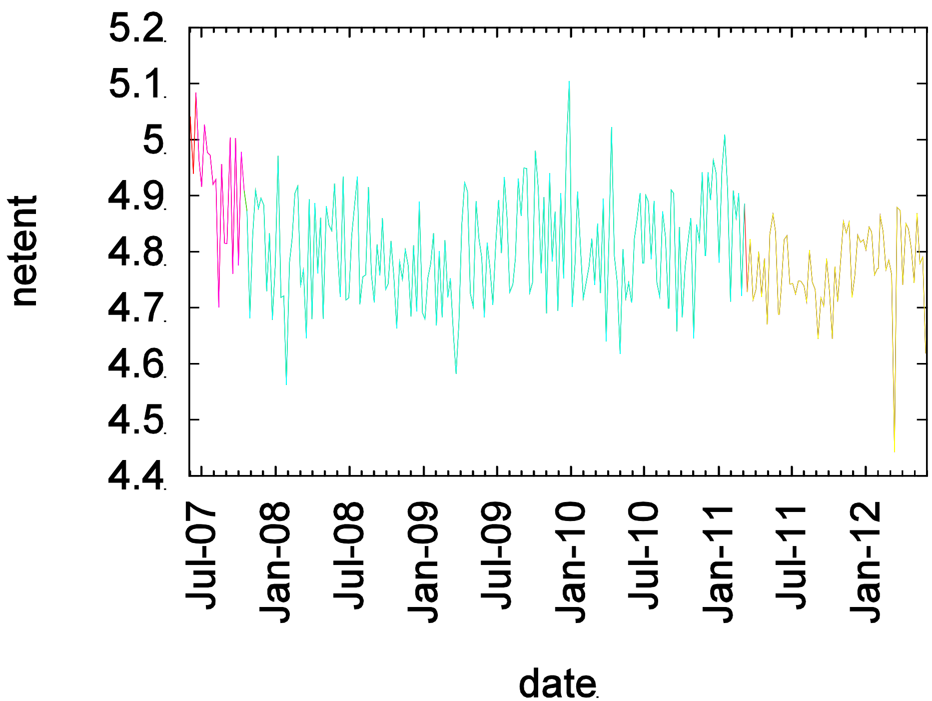


非定常性と時系列分割

- 時系列をパラメータの異なるGumbel分布からの抽出からなるブロックにより構成されていると考える



分割結果(注文)

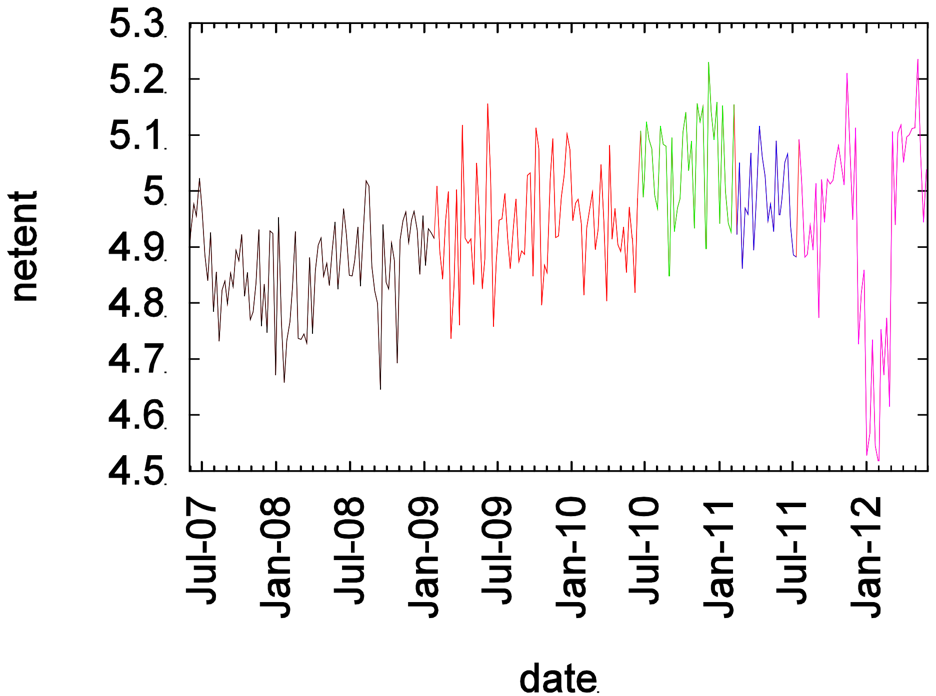


Segment:j	τ_j/T	μ_j	ρ_j
1	0.074906	-4.968606	0.080422
2	0.681648	-4.852616	0.096405
3	0.048689	-4.792295	0.058829
4	0.074906	-4.764064	0.04439
5	0.11985	-4.82097	0.051438

$L_P=5$

T=268 [weeks]

分割結果



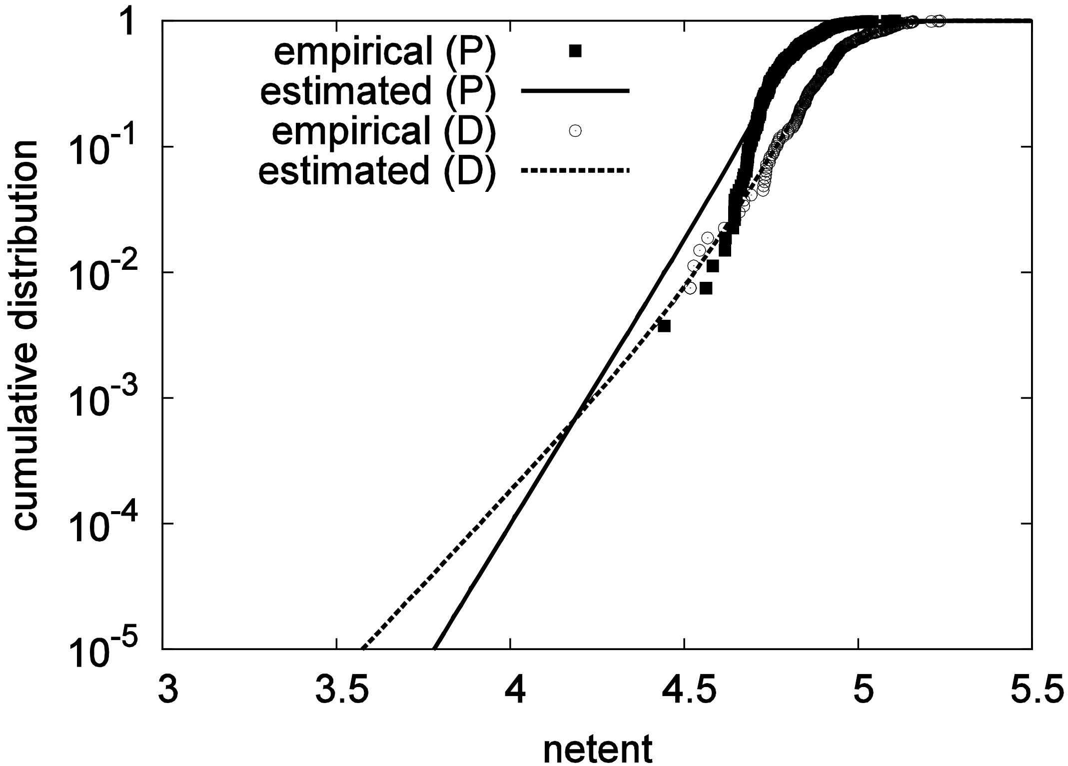
Segment:j	τ_j/T	μ_j	ρ_j
1	0.329588	-4.900974	0.071269
2	0.273408	-4.985201	0.089563
3	0.138577	-5.090311	0.083715
4	0.086142	-5.017903	0.067083
5	0.172285	-5.016248	0.147701

L_D=5

T=268 [weeks]

有限混合Gumbel分布

$$\Pr\left[V_X \le v_X\right] = \sum_{j=1}^{L_X} \frac{\tau_{X,j}}{T} \left\{1 - \exp\left[-e^{\frac{v_X + \mu_{X,j}}{\rho_{X,j}}}\right]\right\}$$



X	KS-value	p-value
P	1.196	0.11
D	0.691	0.72

結論

- ２部グラフネットワークアンサンブルに基づいたネットワークエントロピーを定義
- １リンク当たりのネットワークエントロピーをネットワーク構造の特徴量として利用し大規模同期現象を定量化
- １リンク当たりのネットワークエントロピーの各週最小値の時系列をGumbel分布を用いて分割した
- 有限混合Gumbel分布は大規模同期現象の実際の発生確率をよく説明する

ご清聴ありがとうございました

佐藤彰洋
京都大学大学院情報学研究科数理工学専攻
sato.akihiro.5m@kyoto-u.ac.jp

本研究は日本学術振興会若手研究(B)(#23760074)の財政支援を受けておこなわれた