科学カフェ京都　2010年1月9日@京大

経済物理学をやろう

田中美栄子
Mieko Tanaka-Yamawaki
鳥取大学工学研究科
エレクトロニクス専攻

Econophysics Joint Conference

経済物理学2009

～ミクロとマクロの架け橋～

京都大学基礎物理学研究所 研究会

経済物理学とその周辺

統計数理研究所共同研究集会

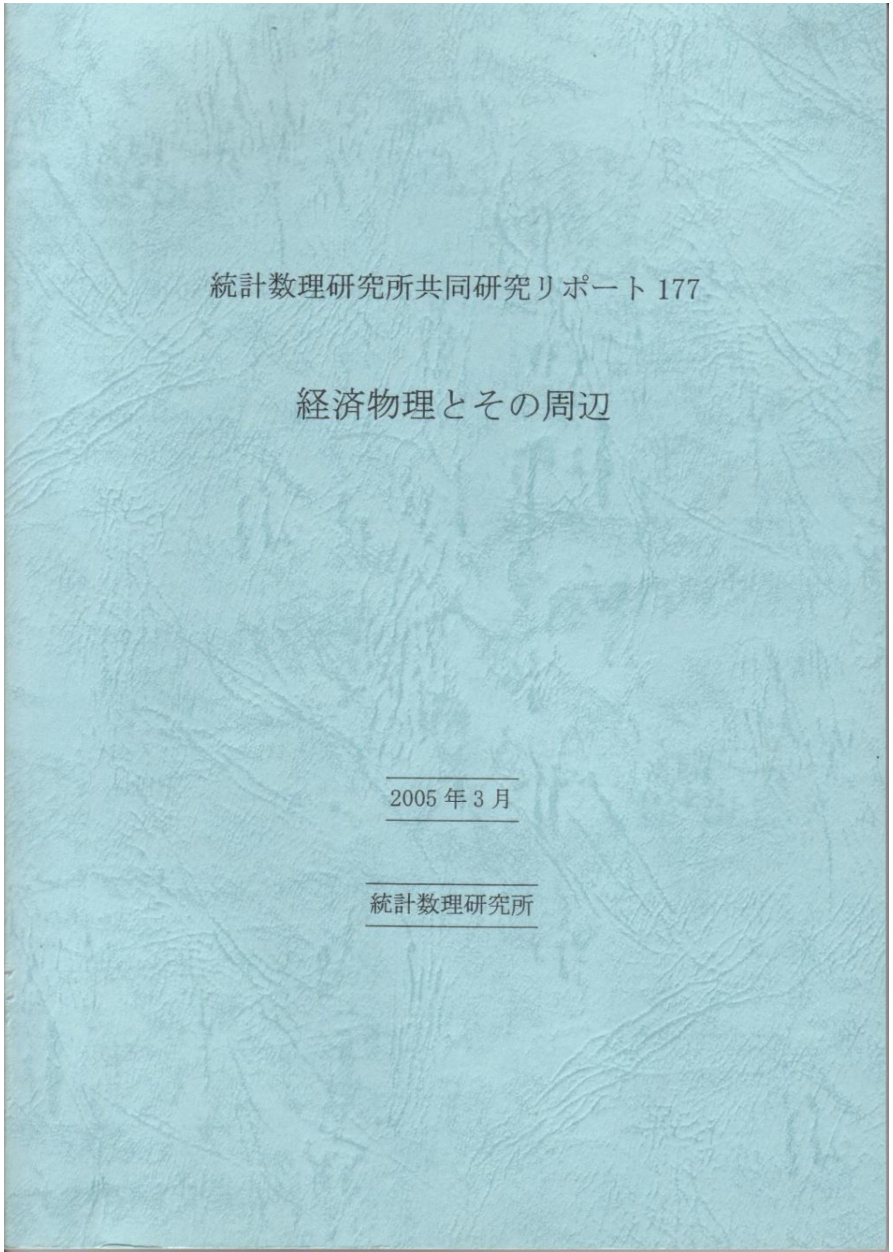
共催：京都大学GCOEプログラム
「知識循環社会のための情報学教育研究拠点」

日時：2009年9月8日（火）～9日（水）

総合研究大学院大学
新分野の開拓1998-2002

統計数理研究所　共同研究集会
経済物理学とその周辺
2003-

統計数理研究所「経済物理とその周辺」2003－

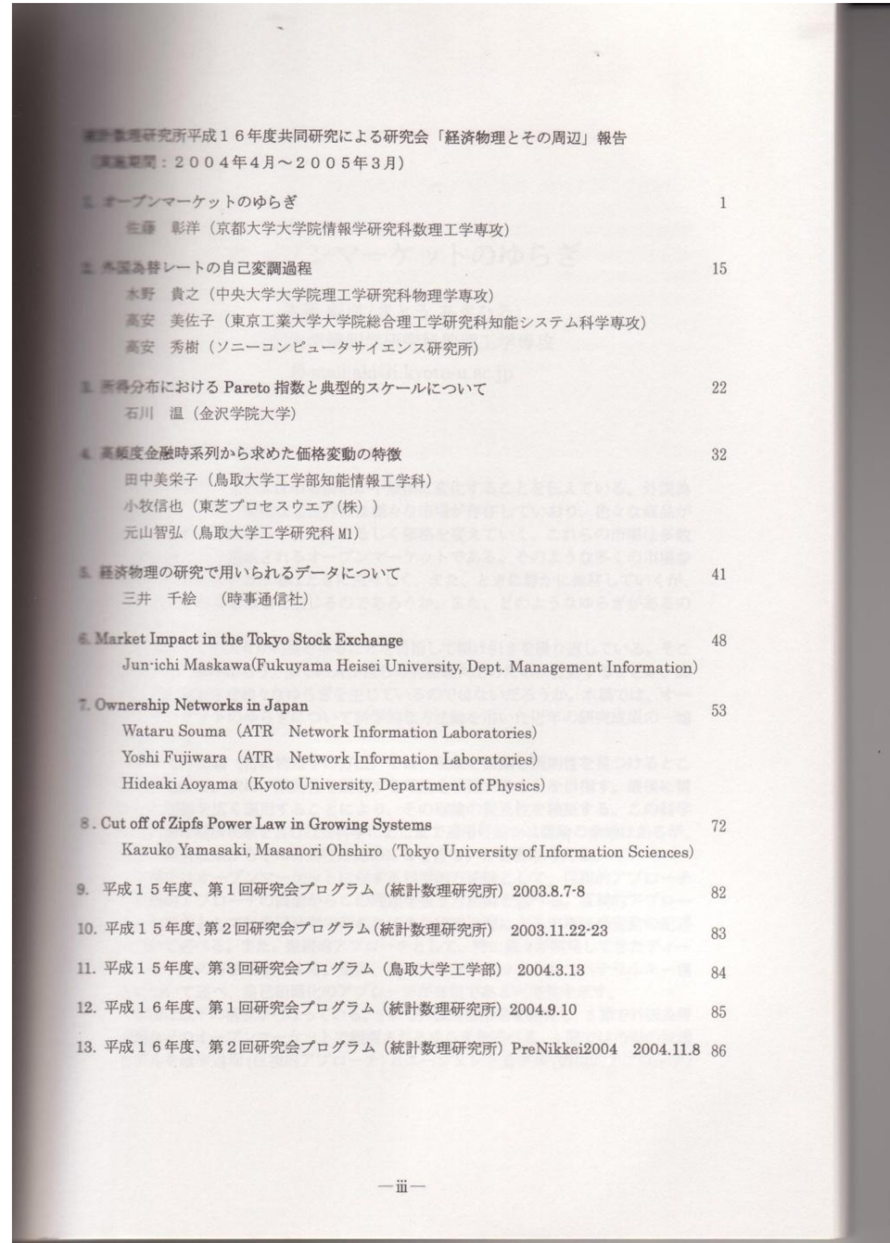


統計数理研究所共同研究リポート 177

経済物理とその周辺

2005 年 3 月

統計数理研究所



統計数理研究所平成１６年度共同研究による研究会「経済物理とその周辺」報告
（実施期間：２００４年４月～２００５年３月）

—iii—

田中美栄子の略歴

京都市中京区の出身
桂高校卒業
京都大学理学部卒業（物理学）
名古屋大学理学研究科理論物理学専攻
ロチェスター大学大学院　Ph.D.（QCD和則）
ニューヨーク市立大学　postdoc
ニューヨーク州立大学 Assitant Prof.
マサチューセッツ州立ノースアダムズ大学
椙山女学園大学、宮崎大学、鳥取大学

アメリカ時代

- ロチェスター大学では

Montroll　と　Mathur　に師事

Montrollが経済物理学に失敗したばかりの頃

間もなくMontrollはMarylandに転出し、死亡

MathurとQCD和則の研究をしてPh.D.

Elliot W. Montroll

http://en.wikipedia.org/wiki/Elliott_Waters_Montroll

(5/4, 1916-12/3, 1983)

Known for

Traffic flow analysis

Notable awards

Lanchester Prize('59)

Kenneth.Wilson,M.Fisher, L.Kadanoff　等に影響.
マンハッタン島トンネル設計

価格変動のスケーリング
所得分布のスケーリング
繰り込み群のアイデア

Montrollに関する本

- C. W. Carey Jr, Elliott Waters Montroll, American National Biography 15 (Oxford, 1999), 717-718.
- Elliott W. Montroll: List of publications, in *The Wonderful World of Stochastics* (North-Holland, Amsterdam, 1985), 17-27.
- D. C. Gazis and R. Herman, "In memory of Elliott W. Montroll," *Transportation Sci*. **18**(2) (1984), 99-100.
- Obituary: Elliott Waters Montroll, *New York Times* (8 December 1983).
- M. F. Shlesinger and G. H. Weiss, Elliott Waters Montroll (May 4, 1916 - December 3, 1983), in The Wonderful World of Stochastics (North-Holland, Amsterdam, 1985), 1-16.
- M. F. Shlesinger and B. J. West, Elliott W. Montroll (1916 - 1983): in memoriam, in *Random Walks and Their Applications in the Physical and Biological Sciences*, Washington, D.C., 1982 (Amer. Inst. Phys., New York, 1 984), vi-viii.
- G. H. Weiss, Elliott Waters Montroll, National Academy of Sciences Biographical Memoirs **63** (1994), 365-380.

1991　帰国

1994-1996　日本は複雑系流行

複雑系研究会（基研）

1997　総研大・新分野の開拓に参加

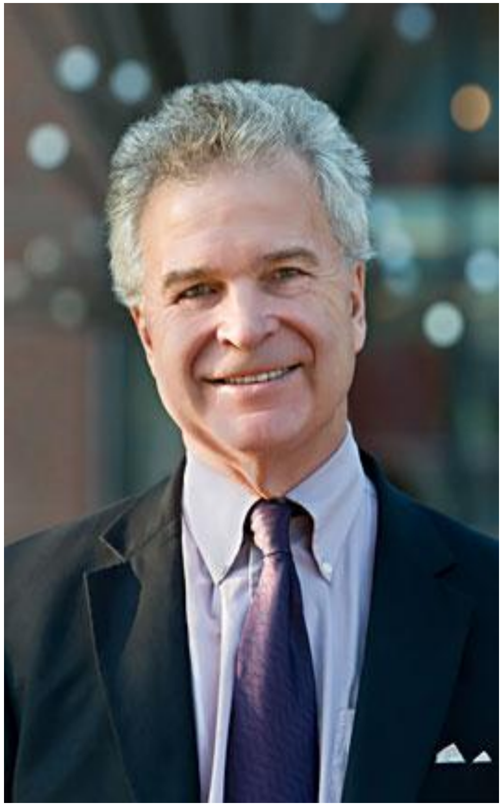
1998-2002　総研大「経済学」小グループ

1997　Stanley等がeconophysicsを提唱

1997　高安秀樹氏が経済物理の可能性を力説（均衡価格に強い疑問を提示.
ダイナミックスの重要性を示唆）

2003-現在　統計数理研共同研究「経済物理とその周辺」

Harry Eugene Stanley 1941-



ジーン･スタンレー
経済物理学の父

Known for Econophysics
statistical physics
Alzheimer's

Notable awards 2004 Boltzmann Medal
2008 Julius Edgar Lilienfeld Prize
Teresiana Medal
Distinguished Teaching Scholar
Director's Award
Nicholson Medal
Memory Ride Award for Alzheimer Research

Econophysicsをググってみると

国際会議

- Nikkei Econophysics
- APFA,
- ECONOPHYS-KOLKATA,
- ESHIA,
- Econophysics Colloquium

書籍

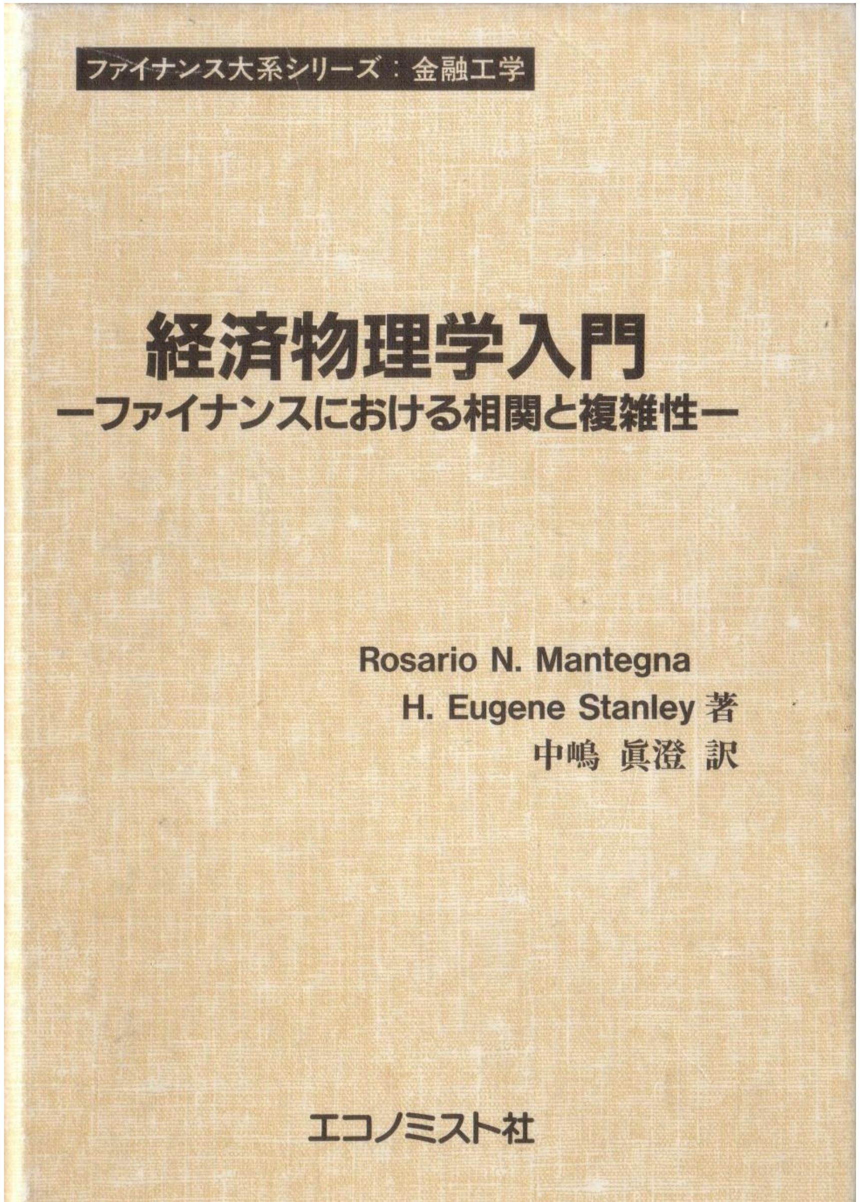
- Mantegna, Stanley, *An Introduction to Econophysics (1999)*
- Bouchaud, Potters(2003)

Theory of Financial Risk and Derivative Pricing,

- Sornette,

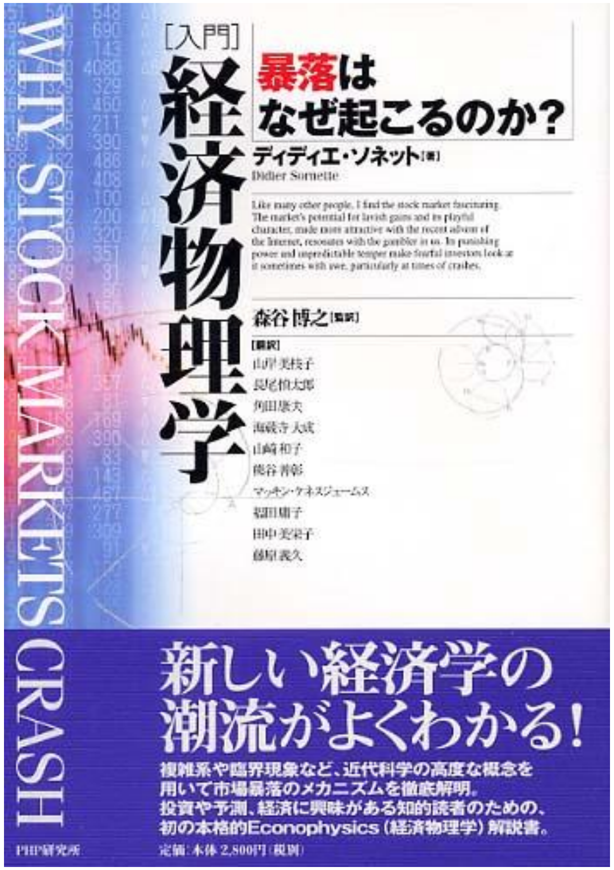
Why Stock Markets Crash (2004).

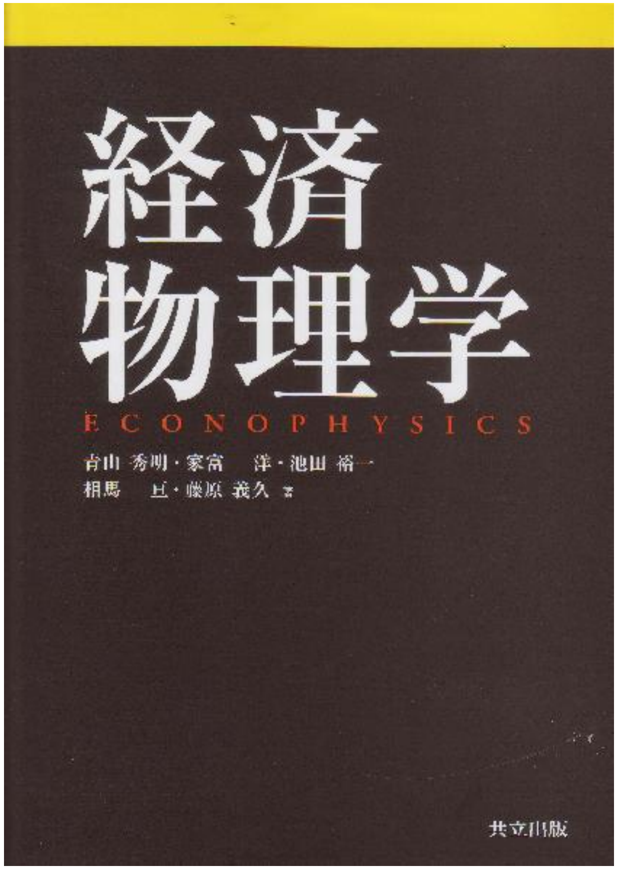
- **Econophysics** is an interdisciplinary research field, applying theories and methods originally developed by physicists in order to solve problems in economics, usually those including uncertainty or stochastic processes and nonlinear dynamics. Its application to the study of financial markets has also been termed statistical finance referring to its roots in statistical physics. Physicists’ interest in the social sciences is not new, Daniel Bernoulli, as an example, was the originator of utility-based preferences. One of the founders of neoclassical economic theory, former Yale University Professor of Economics Irving Fisher, was originally trained under the renowned Yale physicist, Josiah Willard Gibbs.[1


ファイナンス大系シリーズ：金融工学
経済物理学入門
ーファイナンスにおける相関と複雑性ー
Rosario N. Mantegna
H. Eugene Stanley 著
中嶋 眞澄 訳
エコノミスト社


ファイナンス・ライブラリー 6
金融リスクの理論
—— 経済物理からのアプローチ ——
J.-P. ブショー／M. ポッター 著
森平爽一郎 監修　森谷博之／熊谷善彰 訳
朝倉書店

books


[入門]
経済物理学
暴落はなぜ起こるのか？
ディディエ・ソネット[著]
Didier Sornette
森谷博之[監訳]
WHY STOCK MARKETS CRASH
新しい経済学の潮流がよくわかる！
複雑系や臨界現象など、近代科学の高度な概念を用いて市場暴落のメカニズムを徹底解明。
投資や予測、経済に興味がある知的読者のための、初の本格的Econophysics（経済物理学）解説書。
定価：本体2,800円（税別）
PHP研究所


経済物理学
ECONOPHYSICS
青山秀明・家富洋・池田裕一
相馬亘・藤原義久 著
共立出版

価格変動の基本的性質

価格変動 ～ ランダムウオーク(酔歩)

⇒ 予測不可能

⇒ 万人に対して同等の機会を与える投資市場

分単位、秒単位の非常に速い動きは、ランダム性が弱く、先が読める。

価格変動の仕組みを知る

- 価格変動：定量的な経済学の基本
- ティックデータ:

  数秒から数十秒ごとの価格を記録

  1日で数千個

  1年で数十万から数百万個

- 大容量ディスクと高速の計算機が必要
- 計算機科学と経済学の接点

「価格変動＝酔歩」の出典

- 1900年バシュリエの博士論文

「自由競争で決まる価格の変動はブラウン運動（酔歩）」

確率論の創始者アンリ・ポアンカレが推薦

提出先はフランスの最高学府エコール・ノルマル

（インパクト・ファクター高い！）

- その後100年の間、価格変動の基本理念
- 金融工学の出発点
- ブラック・ショールズ・マートンのオプション価格決定式（ブラック・ショールズ公式）導出の出発点として仮定

価格の細かい動きはブラウン運動ではない

- 1960頃、マンデルブロが綿花の投機市場の実データ解析でレヴィ分布になることを発見
- 〈高頻度価格変動データの解析〉

1995年：マンテーニャ・スタンレー

S&P500株価指数の一分毎の価格変動：

確率密度分布がα =1.4のレヴィ分布に従う

1996年：ガシュガイエ等は1分よりも短い、TICK単位の時間間隔の外国為替価格変動データを用いて、外国為替と乱流の類似を指摘

為替の超短期価格変動が、3次元等方乱流とのアナロジーで定式化できるのではと提案

Scaling behaviour in the dynamics of an economic index

Rosario N. Mantegna & H. Eugene Stanley

Center for Polymer Studies and Department of Physics, Boston University, Boston, Massachusetts 02215, USA

THE large-scale dynamical properties of some physical systems depend on the dynamical evolution of a large number of nonlinearly coupled subsystems. Examples include systems that exhibit self-organized criticality[1] and turbulence[2,3]. Such systems tend to exhibit spatial and temporal scaling behaviour—power-law behaviour of a particular observable. Scaling is found in a wide range of systems, from geophysical[4] to biological[5]. Here we explore the possibility that scaling phenomena occur in economic systems—especially when the economic system is one subject to precise rules,

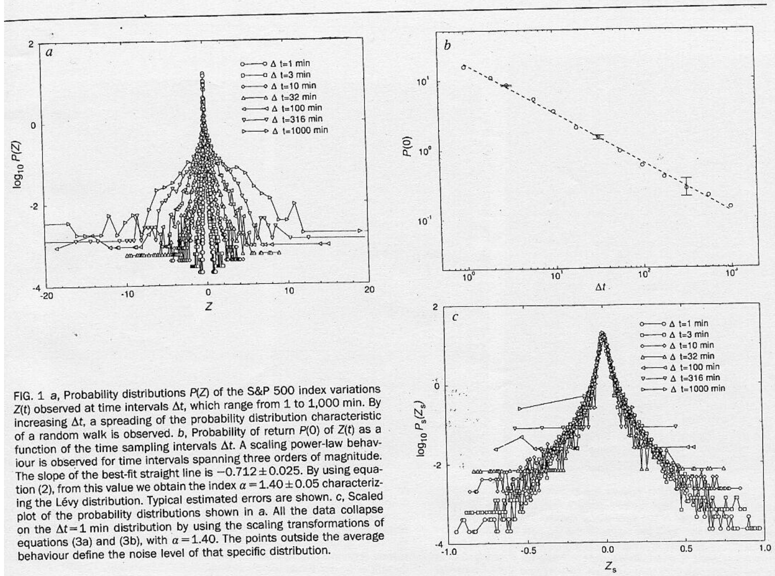


FIG. 1 *a*, Probability distributions $P(Z)$ of the S&P 500 index variations $Z(t)$ observed at time intervals Δt, which range from 1 to 1,000 min. By increasing Δt, a spreading of the probability distribution characteristic of a random walk is observed. *b*, Probability of return $P(0)$ of $Z(t)$ as a function of the time sampling intervals Δt. A scaling power-law behaviour is observed for time intervals spanning three orders of magnitude. The slope of the best-fit straight line is -0.712 ± 0.025. By using equation (2), from this value we obtain the index $\alpha = 1.40 \pm 0.05$ characterizing the Lévy distribution. Typical estimated errors are shown. *c*, Scaled plot of the probability distributions shown in *a*. All the data collapse on the $\Delta t = 1$ min distribution by using the scaling transformations of equations (3a) and (3b), with $\alpha = 1.40$. The points outside the average behaviour define the noise level of that specific distribution.

as is the case in financial markets[6–8]. Specifically, we show that the scaling of the probability distribution of a particular economic index—the Standard & Poor's 500—can be described by a non-gaussian process with dynamics that, for the central part of the distribution, correspond to that predicted for a Lévy stable process[9–11]. Scaling behaviour is observed for time intervals span-

processes have quite similar statistical properties in the high-frequency regime.

We have undertaken a statistical study of timescales as short as 1 min, a value close to the minimum time needed to perform a transaction in the market. Specifically, we investigate the dynamics of a price index of one of the largest financial markets

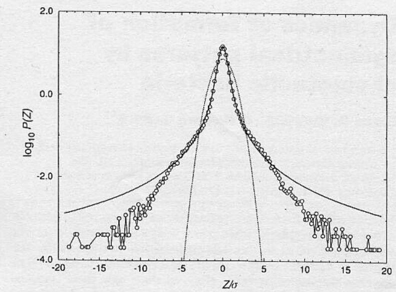


FIG. 2 Comparison of the $\Delta t = 1$ min probability distribution with the symmetrical Lévy stable distribution of index $\alpha = 1.40$ and scale factor $\gamma = 0.00375$ (solid line). The scale factor γ is obtained from equation (2) by using the experimental values of α and $P(0)$. The dotted line shows the gaussian distribution with standard deviation σ equal to the experimental value 0.0508. The variations of price are normalized to this value. Approximately exponential deviations from the Lévy stable profile are observed for $|Z|/\sigma \gtrsim 6$.

1,000 min. The number of data in each set decreases from the maximum value of 493,545 ($\Delta t = 1$ min) to the minimum value of 562 ($\Delta t = 1{,}000$ min).

Figure 1*a* is a semilogarithmic plot of $P(Z)$ obtained for seven different values of Δt. As expected for a random process, the distributions are roughly symmetrical and are spreading when Δt increases. We also note that the distributions are leptokurtic, that is, they have wings larger than expected for a normal process. A determination of the parameters characterizing the distributions is difficult if one uses methods that mainly investigate the wings of distributions, especially because larger values of Δt imply a reduced number of data.

Therefore we use a different approach: we study the 'probability of return' $P(Z=0)$ as a function of Δt. With this choice we investigate the point of each probability distribution that is least affected by the noise introduced by the finiteness of the experimental data set. In Fig. 1*b*, we show $P(0)$ versus Δt in a log–log plot. The data are fitted well by a straight line of slope -0.712 ± 0.025. We observe a non-normal scaling behaviour (slope $\neq -0.5$) in an interval of trading time ranging from 1 to 1,000 min. This experimental finding agrees with the theoretical model of a Lévy walk or Lévy flight[9–11]. In fact, if the central region of the distribution is well described by a Lévy stable symmetrical distribution,

$$L_\alpha(Z, \Delta t) \equiv \frac{1}{\pi}\int_0^\infty \exp(-\gamma \Delta t q^\alpha)\cos(qZ)\,dq \qquad (1)$$

of index α and scale factor γ at $\Delta t = 1$, where $\exp(-\gamma\Delta t|q|^\alpha)$ is the characteristic function of the symmetrical stable process, then the probability of return is given by

$$P(0) \equiv L_\alpha(0, \Delta t) = \frac{\Gamma(1/\alpha)}{\pi\alpha(\gamma\Delta t)^{1/\alpha}} \qquad (2)$$

where Γ is the Gamma function. By using the value -0.712 ± 0.025 from the data of Fig. 1*b* we obtain the index $\alpha = 1.40 \pm 0.05$.

We also check if the scaling extends over the entire probability distribution as well as $Z = 0$. To this end, we first note that Lévy stable symmetrical distributions rescale under the transformations

$$Z_s \equiv \frac{Z}{(\Delta t)^{1/\alpha}} \qquad (3a)$$

and

$$L_\alpha(Z_s, 1) \equiv \frac{L_\alpha(Z, \Delta t)}{(\Delta t)^{-1/\alpha}} \qquad (3b)$$

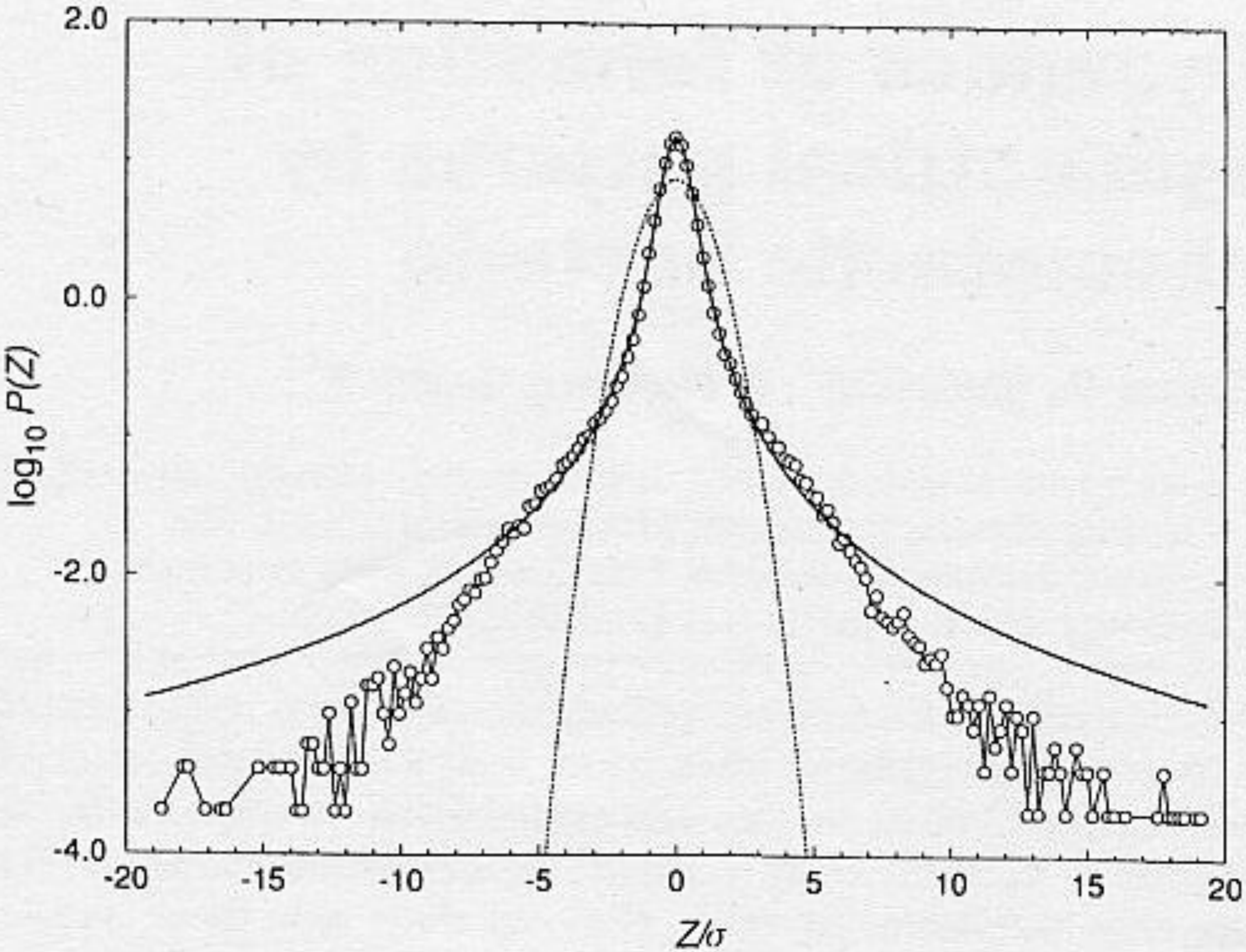
Figure 1*c* shows the distributions of Fig. 1*a* plotted in scaled variables. All the data collapse on the $\Delta t = 1$ min distribution when we use equations (3a) and (3b) with $\alpha = 1.40$. From Fig. 1*c* we conclude that a Lévy distribution describes well the dynamics of the probability distribution $P(Z)$ of the random process over time intervals spanning three orders of magnitude.

In Fig. 2, we compare the probability distribution observed for $\Delta t = 1$ min with the Lévy stable distribution of index $\alpha = 1.40$. Note that the solid line is not simply a 'fit' to the data; rather, the appropriate scale factor $\gamma = 0.00375$ is obtained by using the experimental value of $P(0)$ and equation (2). A good agreement with the Lévy (non-gaussian) profile is observed for almost three orders of magnitude when $|Z|/\sigma \lesssim 6$ and an approximately exponential fall-off from the stable distribution is observed for $|Z|/\sigma \gtrsim 6$; here $\sigma = 0.0508$ is the standard deviation. Our results show a clear deviation of the tails of the distribution from the Lévy profile.

The Lévy distribution has an infinite second moment (if $\alpha < 2$). But our experimental finding of an exponential (or stretched exponential) fall-off implies that the second moment is finite, thereby resolving the question of how one could get around the problem of an infinite variance if the Lévy distribution is used to describe the price difference distribution[24]. This conclusion might at first sight seem to contradict our observation of Lévy scaling of the central part of the price difference distribution over fully three orders of magnitude. However, there is no contradiction, because (for example), a recent study[27] finds that Lévy scaling may hold over a long period of time for the dynamics of 'quasi-stable' stochastic processes having a finite variance.

By using the Berry–Esseen theorem[28,29], we can estimate that the maximal time needed to observe convergence for the price differences to a gaussian process is of the order of 1 month. This estimate is obtained by using the experimental values of $\langle |Z|^3 \rangle$ and $\langle Z^2 \rangle$ observed in the high-frequency regime (for example at $\Delta t = 1$ minute). For our data set, we measure $\langle |Z|^3 \rangle = 0.605 \times 10^{-3}$ and $\langle Z^2 \rangle = 0.00257$, and we set an upper bound of the difference between the integral of the distribution and the corresponding asymptotic normal process of 0.1. Our estimate of roughly 1 month is in agreement with an independent empirical study of the distribution of daily, weekly and monthly returns, for which a progressive convergence to a gaussian process is found[30].

We also investigate the scaling properties of $P(0)$ within each year to determine if the scale index α is highly fluctuating from year to year. The results of this analysis are shown in Fig. 3. We also show the line best fitting $P(0)$ of the entire set of data. In different years, the graph of $P(0)$ is always parallel to the overall behaviour (dotted line) in a log–log plot. This implies that the index α is roughly constant over the years. The scale factor γ (related to the vertical position of data in Fig. 3) varies somewhat from year to year. It is higher (lower positions of the experimental points in the figure) for periods of higher 'volatility' (an economic term indicating higher values of the variance of the price variations). When the same analysis is performed monthly, similar conclusions are obtained: α is roughly constant ($\alpha =$


2.0
0.0
-2.0
-4.0
log10 P(Z)
-20
-15
-10
-5
0
5
10
15
20
Z/σ

為替と乱流の類似

- Ghashgaie,et.al、NATURE　381、1996
- 為替のモーメントが発達した３次元等方乱流のKolmogorovスケーリング則と同じ形になる
- 情報流／エネルギー流が、カスケード構造に従って大きなスケールから小さなスケールへと移動する、と解釈

Turbulent cascades in foreign exchange markets

S. Ghashghaie*, W. Breymann†, J. Peinke‡, P. Talkner§ & Y. Dodge‖

* Fürstensteinerstrasse 4, 4053 Basel, Switzerland
† Institute für Physik der Universität Basel, 4056 Basel, Switzerland
‡ Experimentalphysik II, Universität Bayreuth, 95440 Bayreuth, Germany
§ Paul Scherrer Institut, 5232 Villigen, Switzerland
‖ Groupe de Statistiques, Université de Neuchâtel, 2000 Neuchâtel, Switzerland

The availability of high-frequency data for financial markets has made it possible to study market dynamics on timescales of less than a day[1]. For foreign exchange (FX) rates Müller *et al.*[2] have shown that there is a net flow of information from long to short timescales: the behaviour of long-term traders (who watch the markets only from time to time) influences the behaviour of short-term traders (who watch the markets continuously). Motivated by this hierarchical feature, we have studied FX market

homogeneous isotropic turbulence in three spatial dimensions[24]. It provides a mechanism for dissipating large amounts of energy in a viscous fluid. Energy is pumped into the system at large scales of the order of, say, metres (by a moving car or a flying aeroplane) or kilometres (by meteorological events), transferred to smaller scales through a hierarchy of eddies of decreasing sizes, and dissipated at the smallest scale—of the order of millimetres in the above examples. This cascade of kinetic energy extending over several orders of magnitude generates a scaling behaviour of the eddies and manifests itself in a scaling of the moments $\langle(\Delta v)^n\rangle$ of Δv as $(\Delta r)^{\zeta_n}$ (refs 5,6). Here the angle brackets $\langle\,\rangle$ denote the mean value of the enclosed quantity and Δv is the difference of the velocity component in the direction of the spatial separation of length Δr. Under the assumption that the eddies of each size are space-filling and that the downward energy flow is homogeneous, $\zeta_n = n/3$ (ref. 8). The probability densities $P_{\Delta r}(\Delta v)$ are then scale invariant. This means that if the velocity differences are normalized by their respective standard deviation, the resulting standardized probability densities do not depend on Δr. But for $n > 3$, experimentally determined values of ζ_n follow a concave curve definitely below the ($\zeta_n = n/3$)-line (Fig. 2*b*). The dependence of the standardized probability density on Δr also provides evidence that eddies of a given size are not space-filling but rather fluctuating in space and time in a typical intermittent way (Fig. 1*b*).

Our analyses of FX markets are based on a data set provided by Olsen and Associates containing all worldwide 1,472,241 bid–ask quotes for US dollar–German mark exchange rates which have emanated from the interbank Reuters network from 1 October 1992 until 30 September 1993. From these data we have determined the probability densities of price changes $P_{\Delta t}(\Delta x)$ with time delays Δt varying from five minutes up to approximately two days, which are displayed in Fig. 1*a*. In comparison, Fig. 1*b* shows the analogous turbulent probability densities $P_{\Delta r}(\Delta v)$, which exhibit the same characteristic features. Using the probability density $P_{\Delta t}(\Delta x)$, the information acquired by observing the market after a time Δt can be quantified as $I(\Delta t) = -\int P_{\Delta t}(\Delta x)\log P_{\Delta t}(\Delta x)\,d(\Delta x)$. It turns out that the dependence of this information on Δt is directly related to the scaling of the variance of Δx with Δt. In turbulence, on the other hand, the variance of the velocity differences at a distance Δr is proportional to the mean energy which is contained in an eddy of size Δr. This further supports the proposed analogy between energy and information.

Given the analogy between turbulence and FX market dynamics, we expect the moments of FX price changes to scale with the time delay as $\langle(\Delta x)^n\rangle \propto (\Delta t)^{\xi_n}$. Scaling has already been reported for the mean absolute values of FX returns in ref. 9 and by Evertsz in ref. 1, and for the second moments of the variations of the Standard & Poor's 500 economic index in ref. 10. For FX

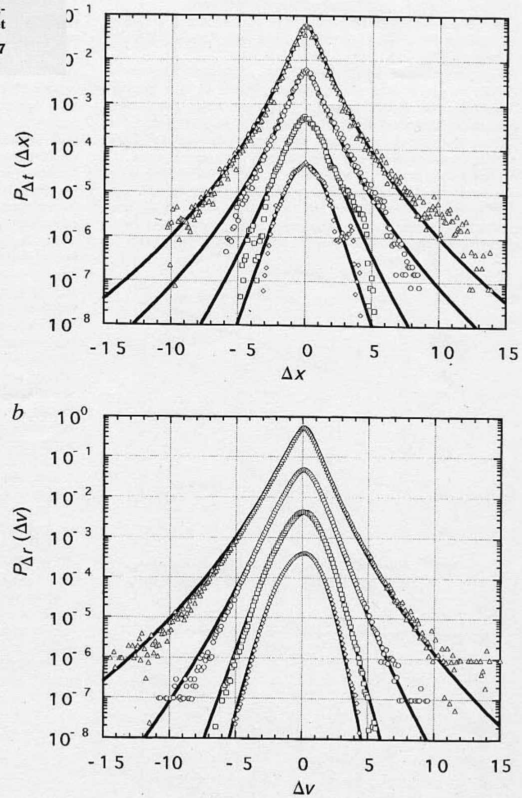


FIG. 1 *a*, Data points: standardized probability density $P_{\Delta t}(\Delta x)$ of price changes $\Delta x = x(t) - x(t+\Delta t)$ for time delays Δt = 640 s, 5,120 s, 40,960 s, 163,840 s (from top to bottom). The middle prices $x(t) = (x_{bid}(t) + x_{ask}(t))/2$ have been used (data provided by Olsen & Associates (see text)). The probability density has been calculated in a similar way as ref. 10. Full lines: results of (least-squares) fits carried out according to ref. 7; λ^2 = 0.25, 0.23, 0.13, 0.06 (from top to bottom). For better visibility, the curves have been vertically shifted with respect to each other. Note the systematic change in shape of the densities. *b*, Data points: standardized probability density $P_{\Delta v}$ of velocity differences $P_{\Delta r}(\Delta v)$ for a turbulent flow with $\Delta r = 3.3\eta$, 18.5η, 138η, 325η (data taken from ref. 13, $R_\lambda = 598$. Here η is the Kolmogorov scale, where viscous dissipation occurs. Full lines: results of (least-squares) fits carried out according to ref. 7; λ^2 = 0.19, 0.10, 0.04, 0.01.

market data, it has also been observed that the kurtosis, which is the ratio of the fourth moment and the squared variance, decreases with increasing time delay, that is, the tails of the probability density lose weight[9,11]. Figure 2*a* shows the higher-order moments of the FX price changes. They scale for time delays Δt varying from about five minutes up to several hours. As in turbulence, the scaling exponents ξ_n depend on the order of the moments in a nonlinear way. Moreover, the ξ_n of the FX data are close to the ζ_n of turbulent data (Fig. 2*b*). This is also in agreement with the observation that the shapes of the turbulent and FX probability densities, $P_{\Delta r}(\Delta v)$ and $P_{\Delta t}(\Delta x)$, depend on their respective scale parameters in a similar way: both show a decrease of kurtosis with increasing scale parameter[12] (Fig. 1*a* and 1*b*).

Even though the probability densities $P_{\Delta t}(\Delta x)$ and $P_{\Delta r}(\Delta v)$ are in principle determined by their moments, this is of little help in practice because errors drastically increase when calculating higher-order moments from observed data. In turbulence, for different experimental situations, the change in shape of $P_{\Delta r}(\Delta v)$ has been successfully parametrized using a method motivated by the energy cascade[7,13,14]. The standardized probability density is approximated by a superposition of gaussian densities with log-normally distributed variances. The variance λ^2 of this log–normal distribution is a measurable form parameter containing information on the energy cascade[15]. It is the only parameter that must be adjusted to the data. Applying this method to the FX market data yields a surprisingly good fit (Fig. 1*a*). Although the data in the centre dominate the fit, the agreement is reasonable also as regards the tails of the probability densities as long as the time delay is shorter than about two days. This is in contrast to fits with Levy distributions which significantly deviate in the tails[10]. The parameter λ^2, which measures the spread of the variances of the superimposed gaussians (that is, the variance of the log–normal distribution), decreases linearly with $\log \Delta t$ (Fig. 3). This result further confirms the similarity of the statistical behaviour of FX markets with the classical picture of turbulence as given by Kolmogorov[5–7]. In particular, it verifies the scaling of the moments, $\langle(\Delta x)^n\rangle \propto (\Delta t)^{\xi_n}$, in an independent way.

An important aspect of turbulent flows is their intermittent behaviour, that is, the typical occurrence of laminar periods which are interrupted by turbulent bursts. In FX markets this corresponds to clusters of high and low volatility[12], which give rise to relatively high values of the probability densities $P_{\Delta t}(\Delta x)$ both in the centre and the tails. This particular aspect is well reproduced for time delays smaller than two days by the proposed log–normal superpositions of gaussians. However, long interruptions of the trading process (particularly during weekends) affect the probability densities and lead to deviations from the expected shape. Hence, we conclude that the range which is governed by the information cascade has an upper limit. We note that in turbulence also there is an upper length scale, where the energy cascade sets in and beyond which scaling fails.

Different types of models have been proposed to describe the statistical characteristics of price differences of financial indices as, for example, the behaviour of the volatility. Prominent approaches are models using mixtures of distributions[16–18] and ARCH/GARCH-type models[19,20] (see, for example, the overview in ref. 21). But these studies do not address the scaling behaviour of the probability distributions of FX price changes.

It is unlikely that there is a set of a few partial differential equations (like the Navier Stokes equations in hydrodynamics) which might serve as a model of FX market dynamics. More striking is the similarity of the two phenomena, which is accounted for by the existence of a cascade in both cases. This has prompted us to introduce concepts of turbulence in the description of FX

TABLE 1 Correspondence between fully developed three-dimensional turbulence and FX markets

Hydrodynamic turbulence	FX markets
Energy	Information
Spatial distance	Time delay
Laminar periods interrupted by turbulent bursts (intermittency)	Clusters of low and high volatility
Energy cascade in space hierarchy $\langle(\Delta v)^n\rangle \propto (\Delta r)^{\zeta_n}$	Information cascade in time hierarchy $\langle(\Delta x)^n\rangle \propto (\Delta t)^{\xi_n}$

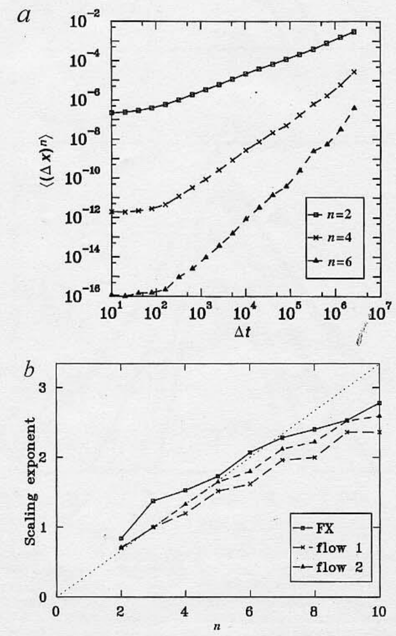


FIG. 2 *a*, Scaling of moments $\langle(\Delta x)^n\rangle$ with n = 2, 4 and 6. *b*, n-dependence of the scaling factors ξ_n and ζ_n for the nth moments of the probability densities of FX price changes (squares) and hydrodynamic velocity differences taken from ref. 22 (crosses) and ref. 23 (triangles). Note the same qualitative deviation of all curves from a straight line.

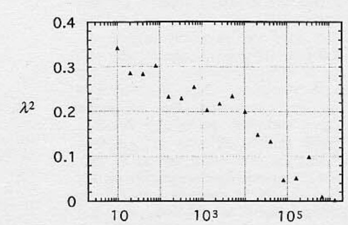
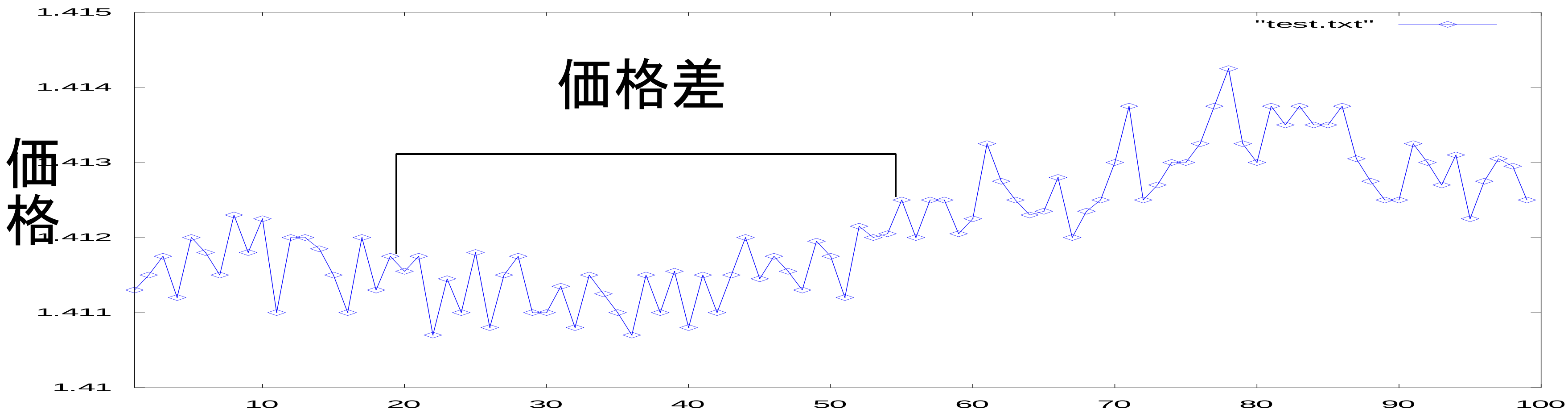
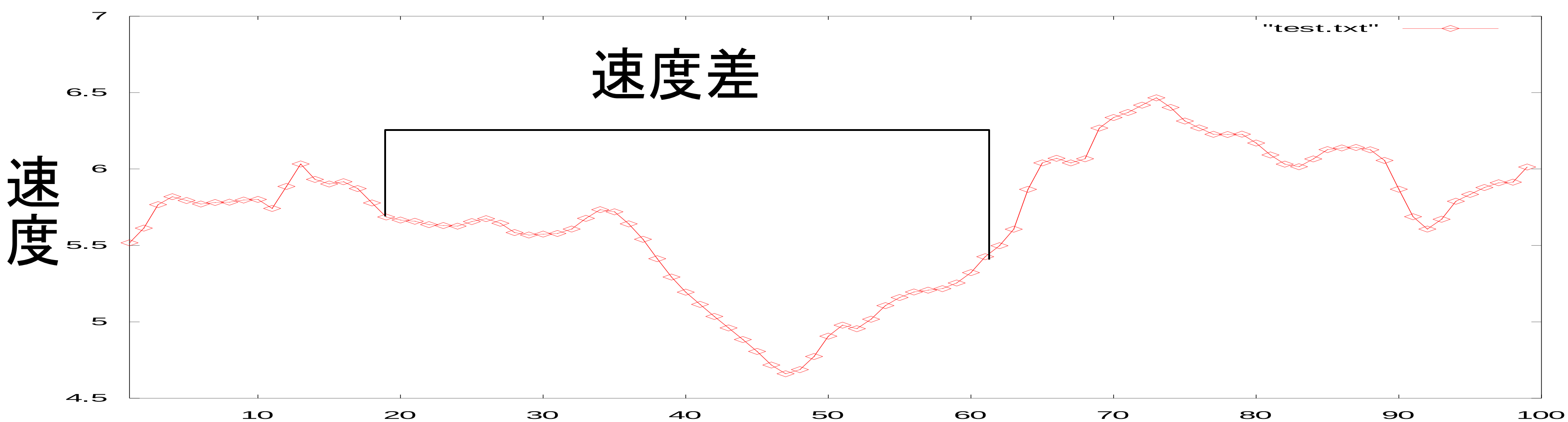


FIG. 3 Dependence of the form parameter λ^2 on Δt (in seconds).

外国為替と乱流の実データ

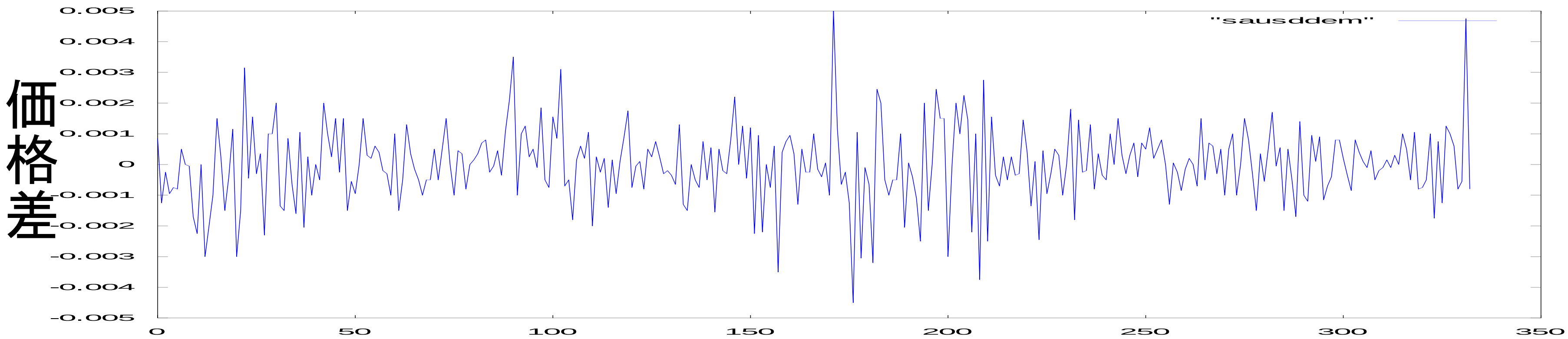


外国為替

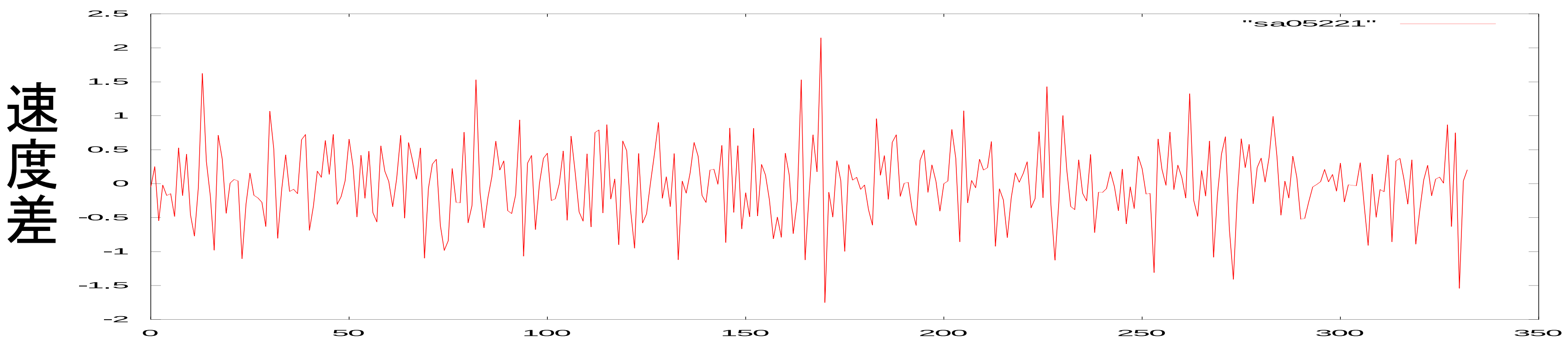


乱流

価格差、速度差は類似

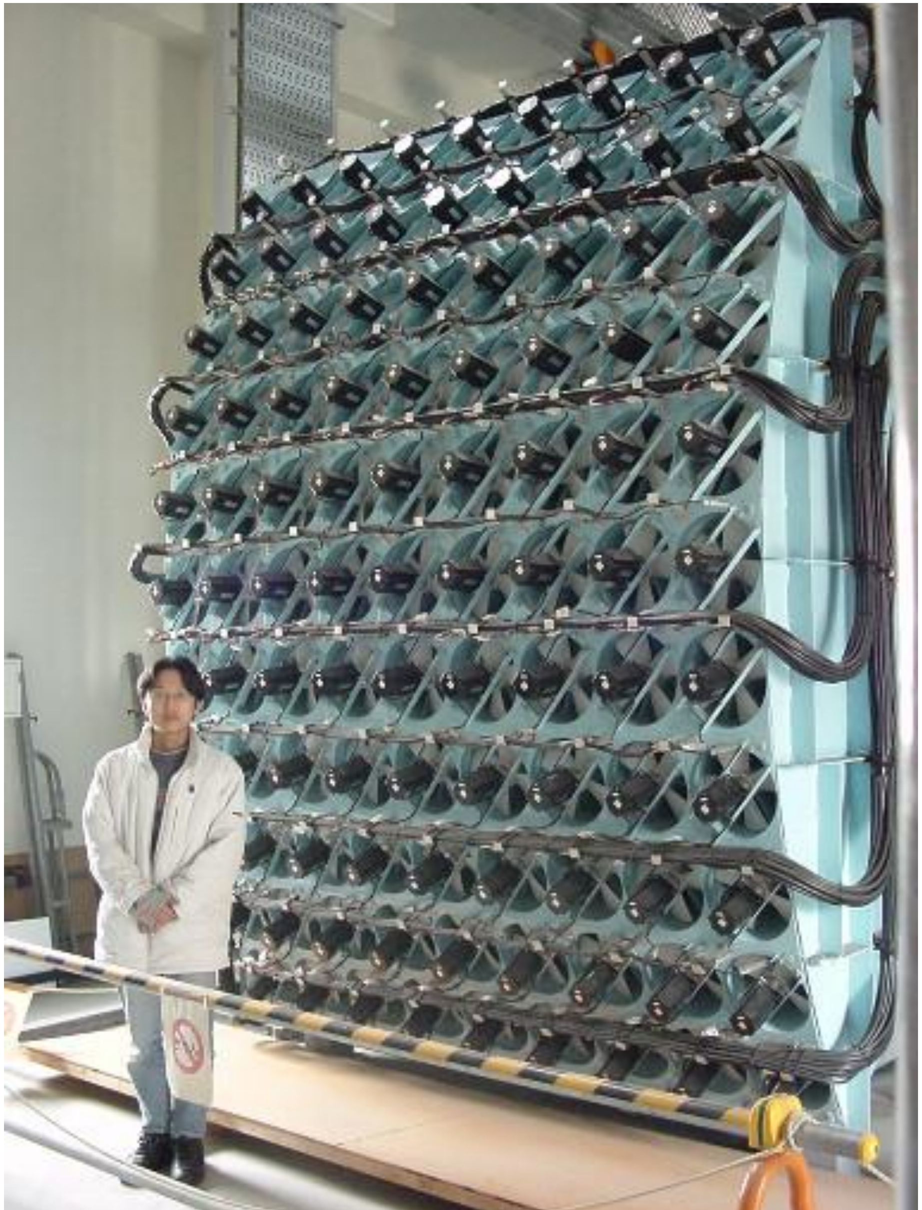


外国為替



乱流

マルチファン乱流風洞＠宮崎大学



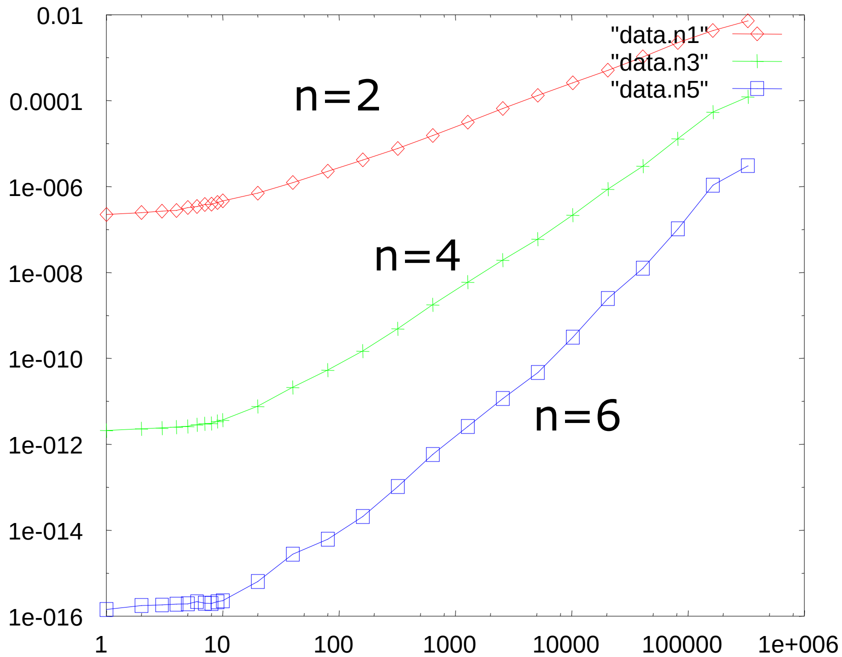
筒は15メートル

実験は2001年

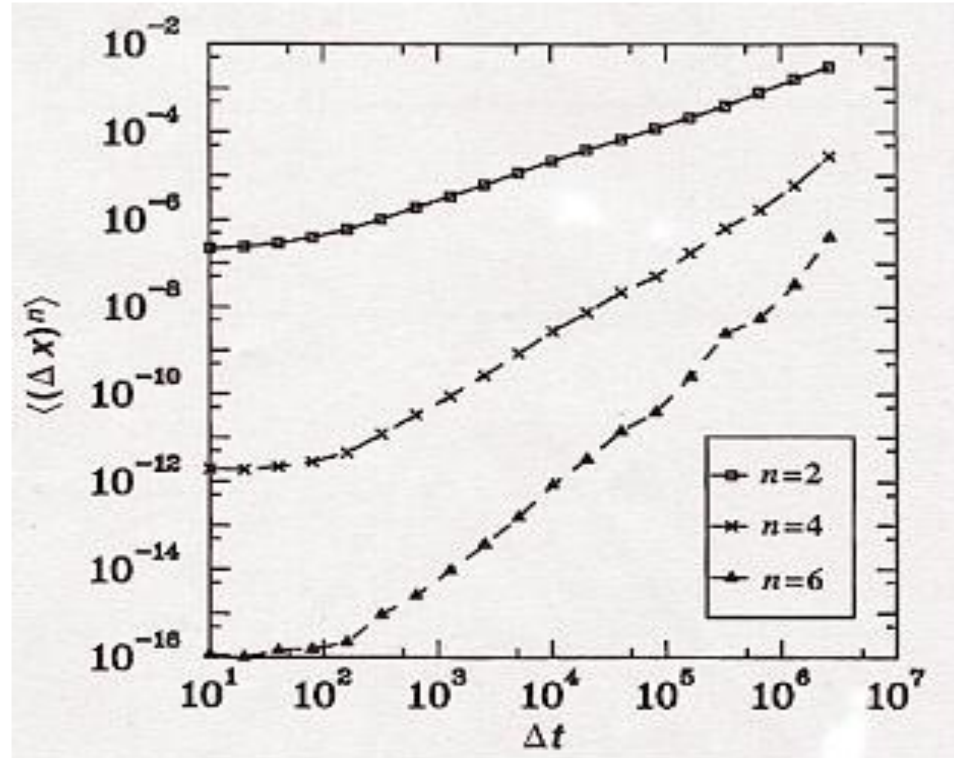
～2004年にかけて工学部材料物理工学科小園研究室を中心に行われている。

乱流と外国為替の対応表

乱流	外国為替
速度差 Δv	価格差 Δx
空間の解像度 Δr	時間の解像度 Δt
$\left\langle (\Delta v)^{n} \right\rangle \propto (\Delta r)^{\zeta_n}$	$\left\langle (\Delta x)^{n} \right\rangle \propto (\Delta t)^{\xi_n}$



本研究で求めた結果



文献[1]での結果

Ghashghaieらの論文(右)とほぼ同じ結果を再現

使用データ

外国為替データ		
データ名	詳細	データ数
FX1	1992～1993年 ドイツマルク対アメリカドル	約147万
FX2	1995～2001年 日本円対アメリカドル	約1000万
乱流データ		
TRB	2002年3月に宮崎大学の乱流風洞実験により得られたデータ	25万個

1.確率密度分布

為替、乱流のデータを用い、価格差Δx、速度差Δvの確率密度分布を求める。

ここで

- t : 時刻
- x(t) : 時刻tでの価格
- Δx=x(t)-x(t+Δt)
- r : 距離
- v(r) : 距離rでの速度
- Δv=v(r)-v(r+Δr)

とする(Δt、Δrというのはデータ上でのtickを単位とした飛ばし幅である)。

tickの説明

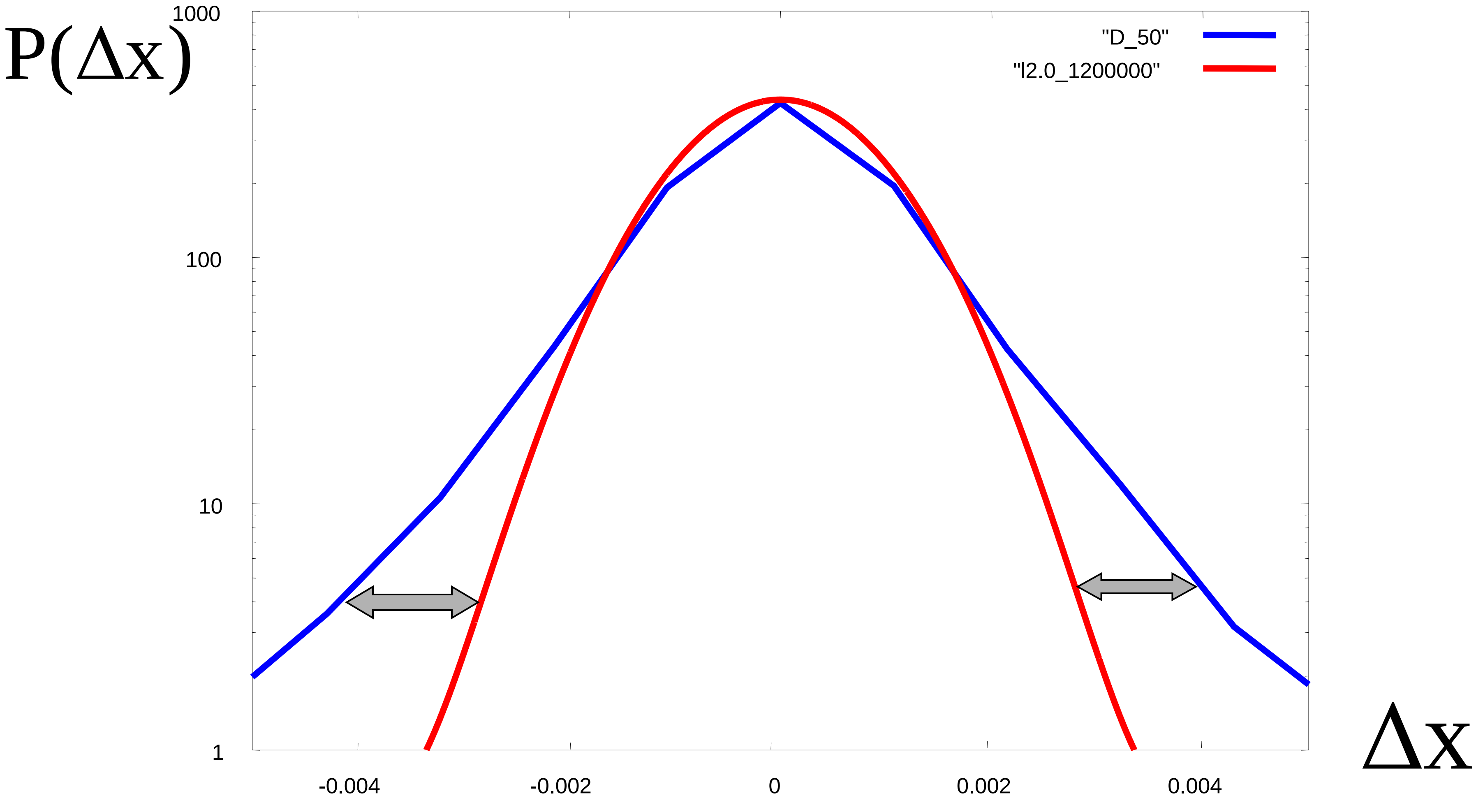
データ

1	1.41185
2	1.412
3	1.4113
4	1.41175
5	1.4107
6	1.41145
7	1.411
8	1.4118
9	1.4108
10	1.4115
11	1.41175
12	1.411

$\Delta t=5$の時
$\Delta x=0.0004$

- 為替では取引毎にデータが取られており、一回取引が行われる度に1tickデータが増えている。
- $\Delta t=5$というのは5tickを意味しており、$x(t)-x(t+\Delta t)$は左に示していることである。

FX1の確率密度分布



- FX1の確率密度分布（Δt=50）
- 正規分布

レヴィ分布

- レヴィ分布を示す(2)式において

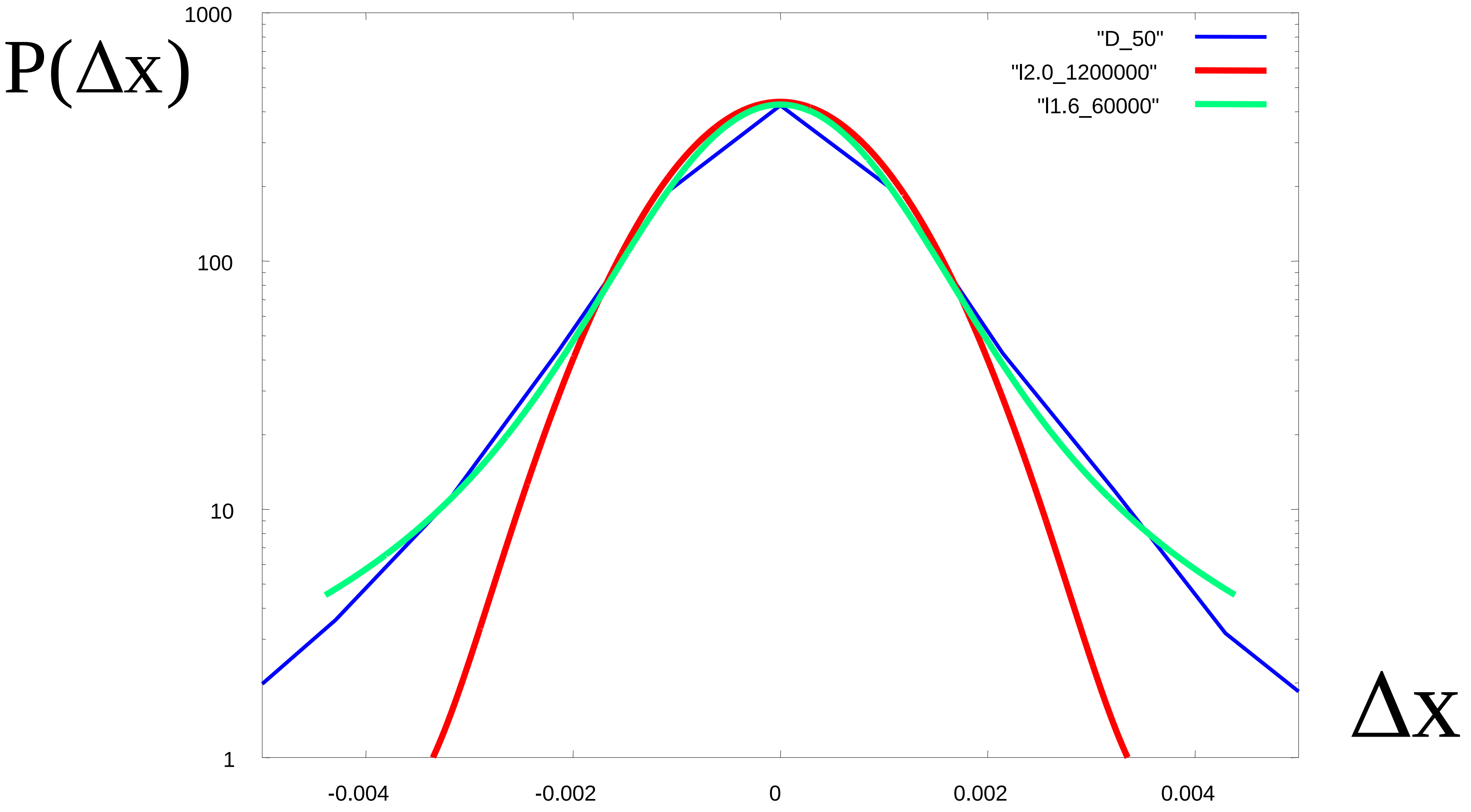
$$L_{\alpha,\beta}(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\beta|k|^{\alpha}} e^{ikx} dk \quad \cdots (2)$$

- パラメータ$\alpha=2$の場合、正規分布となる。

正規分布が頂点付近においてある程度のフィットを見せていた。

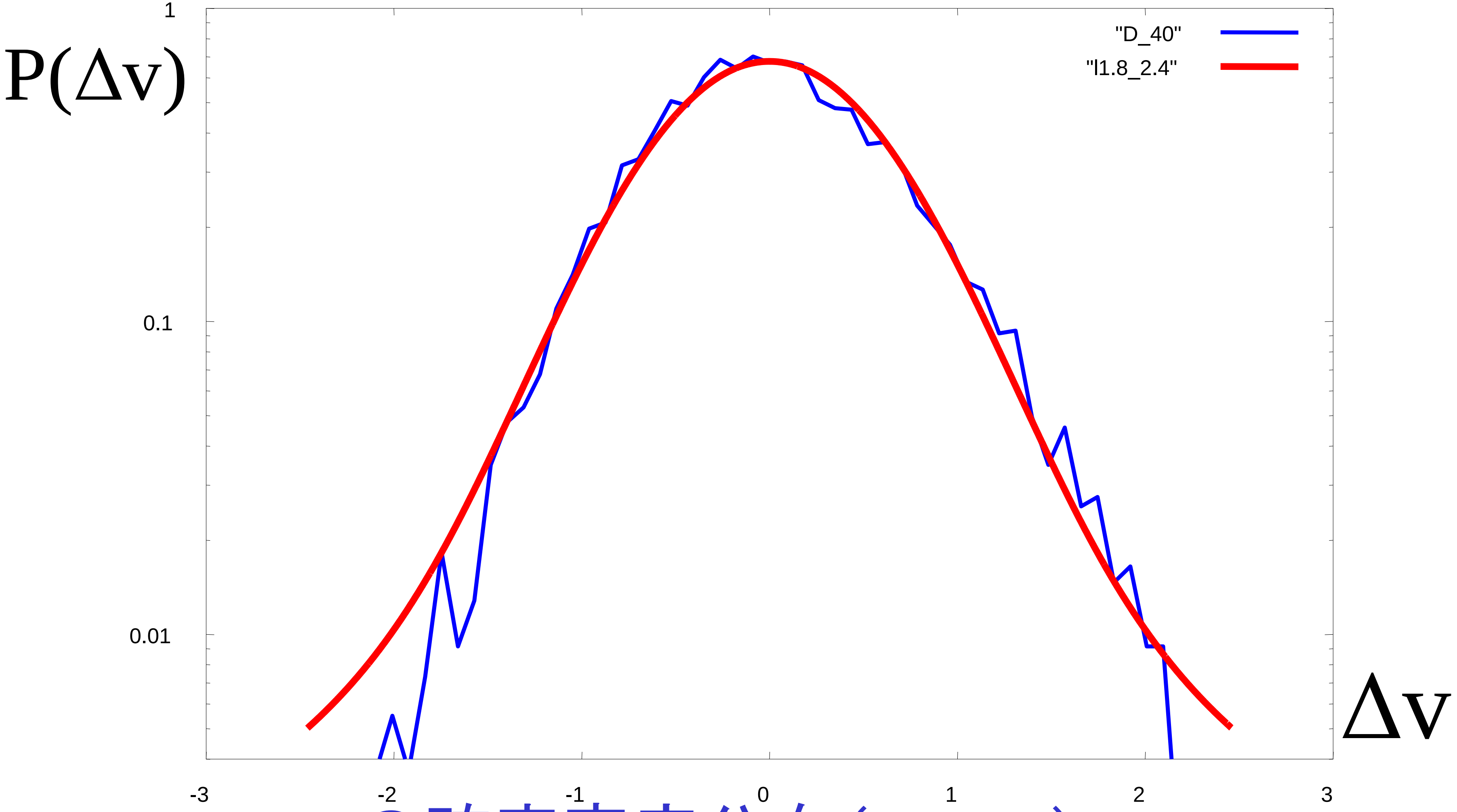
レヴィ分布によるフィット

レヴィ分布を用いたフィッティング



- α=1.6のレヴィ分布
- FX1の分布とほぼ一致している。

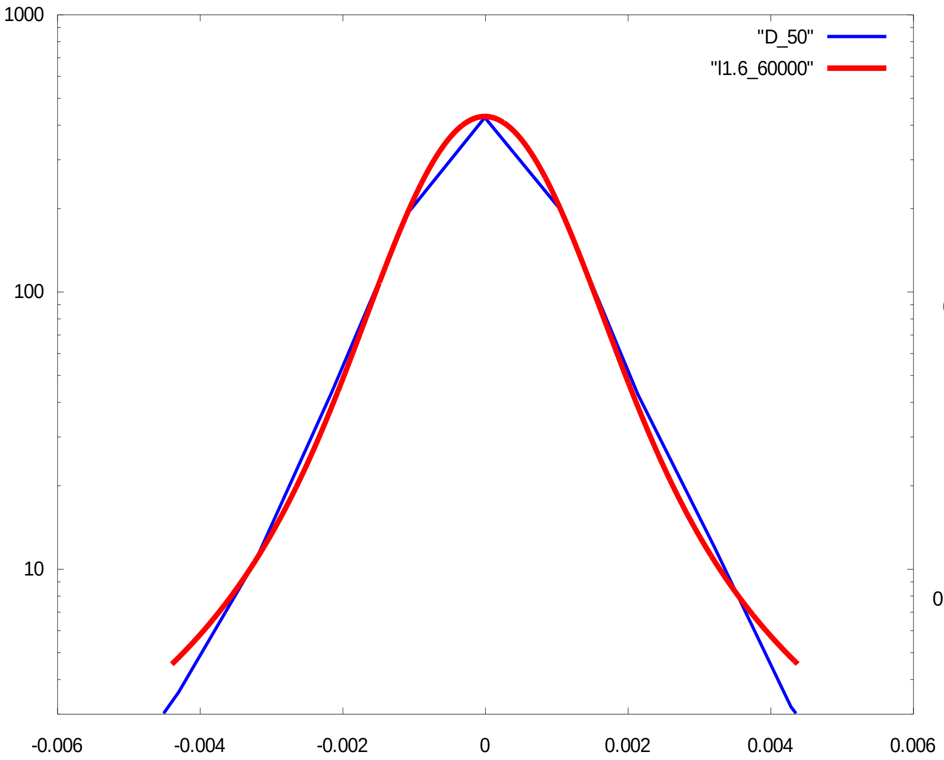
TRBの確率密度分布



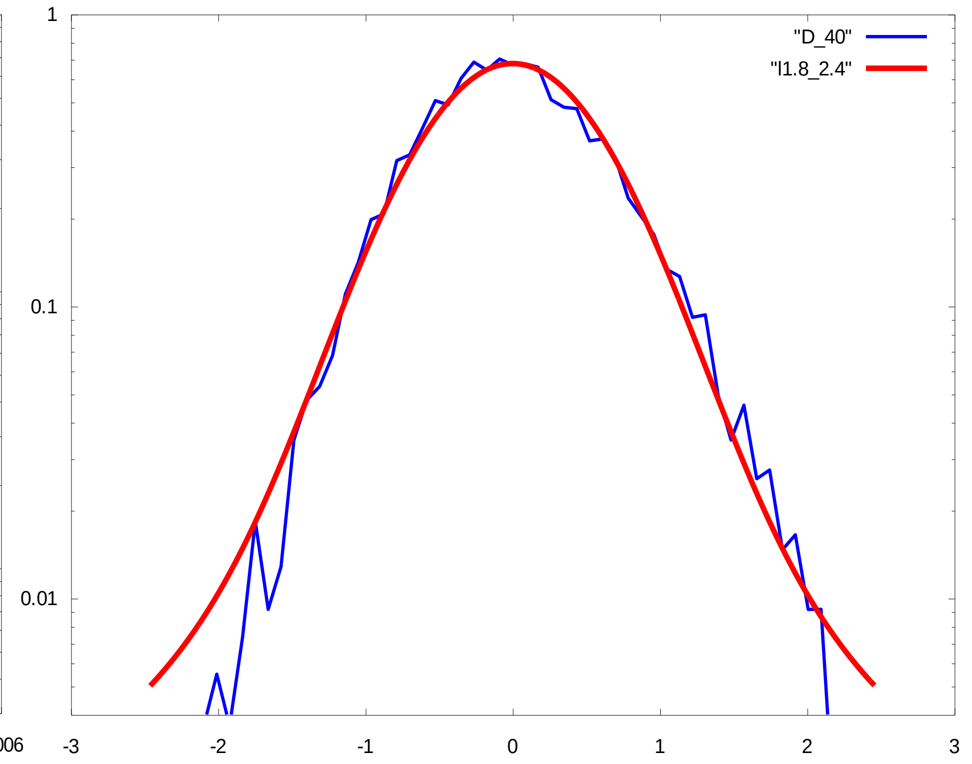
- TRBの確率密度分布（Δr=40）
- α=1.8のレヴィ分布

確率密度分布の類似

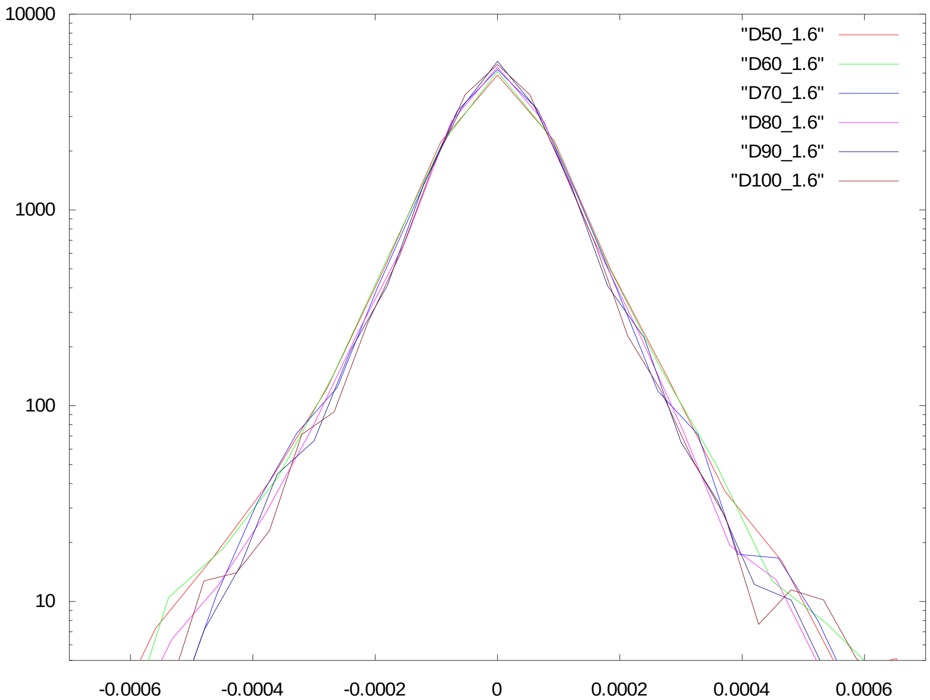
外国為替



乱流

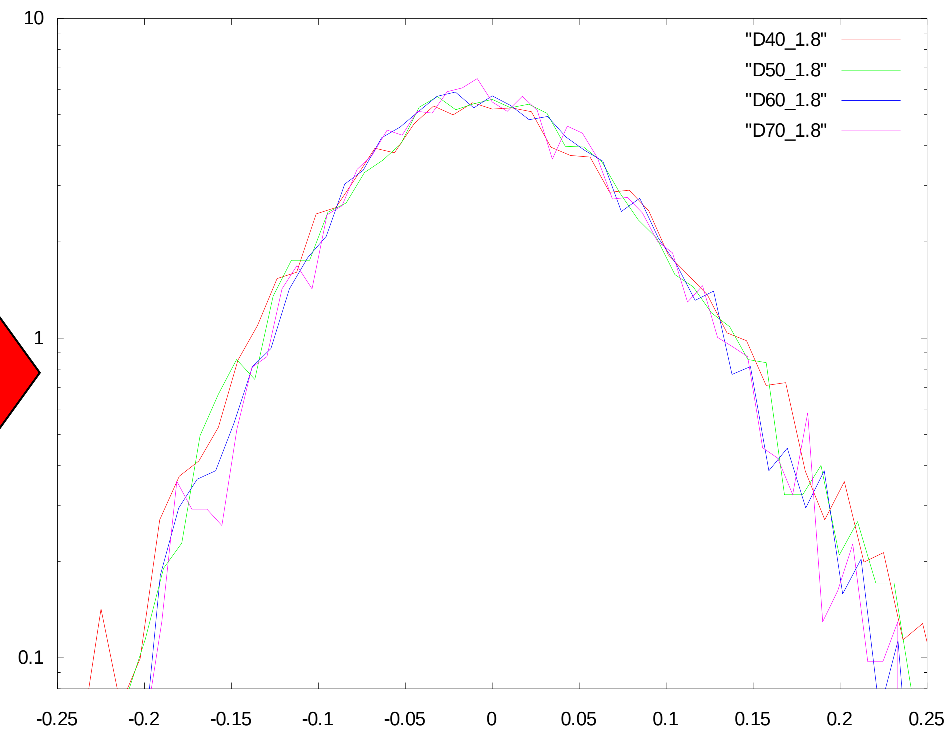


- 両者ともにレヴィ分布でフィットできており、確率密度分布は類似していると言える。



- FX1でのスケール変換のグラフ
- Δt=50~100の範囲で同一分布に重なっている。

- TRBでのスケール変換のグラフ
- Δr=40~70の範囲で同一分布に重なっている。



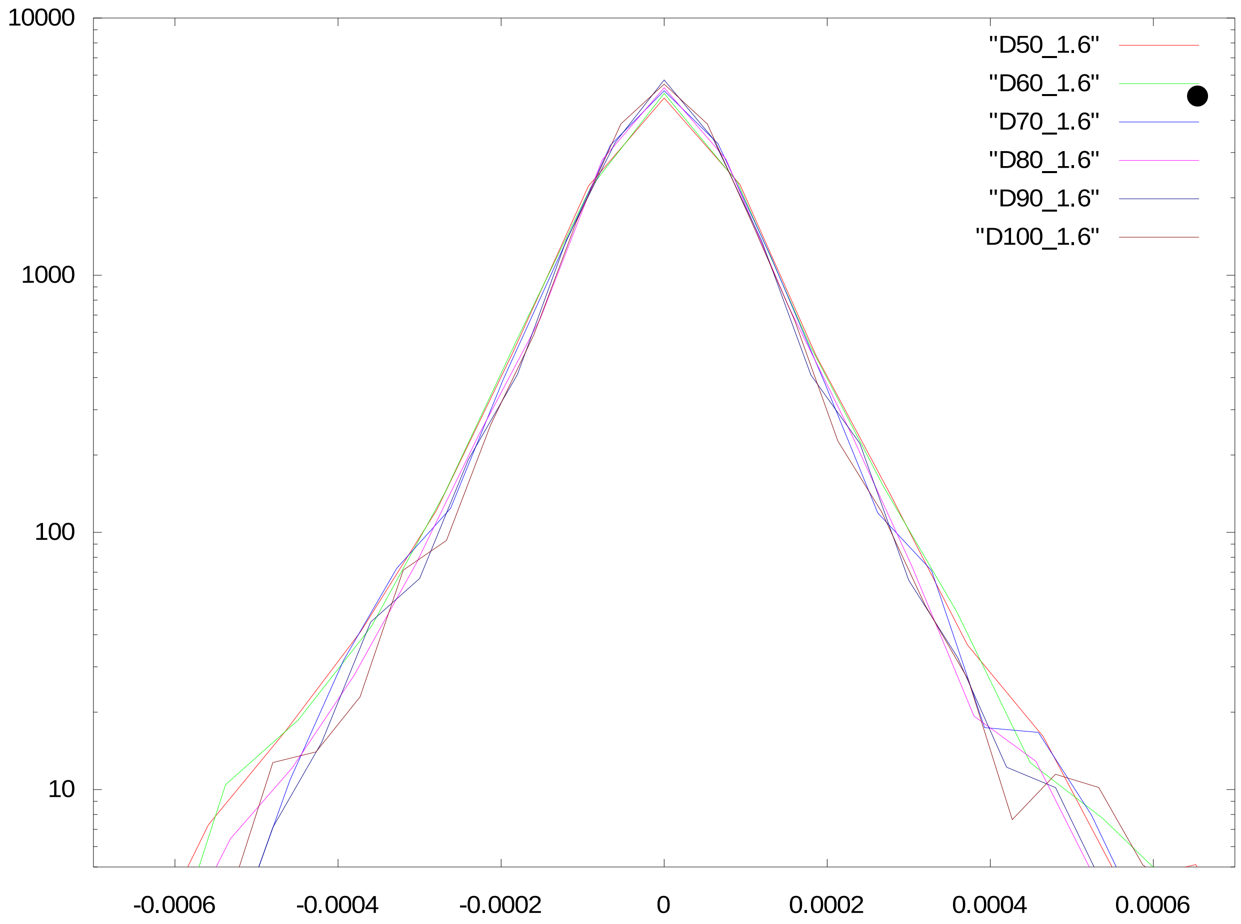
スケール変換

- レヴィ分布は次のような性質を示す。

$$P_{\alpha,\beta}(x) = \lambda^{1/\alpha} P_{\alpha,\lambda\beta}(\lambda^{1/\alpha} x) \quad \cdots (3)$$

- (3)式はスケールを合わせることで異なるβの $P_{\alpha,\beta}(x)$ が、それぞれ同一分布に重なることを意味している。
- βはΔt（Δr）に比例している。

データ長、種類、解像度で異なるα



- Δt=40-100の範囲でα=1.6

別の見方:　原点回帰率:P(0)

- スケール変換で用いた(3)式

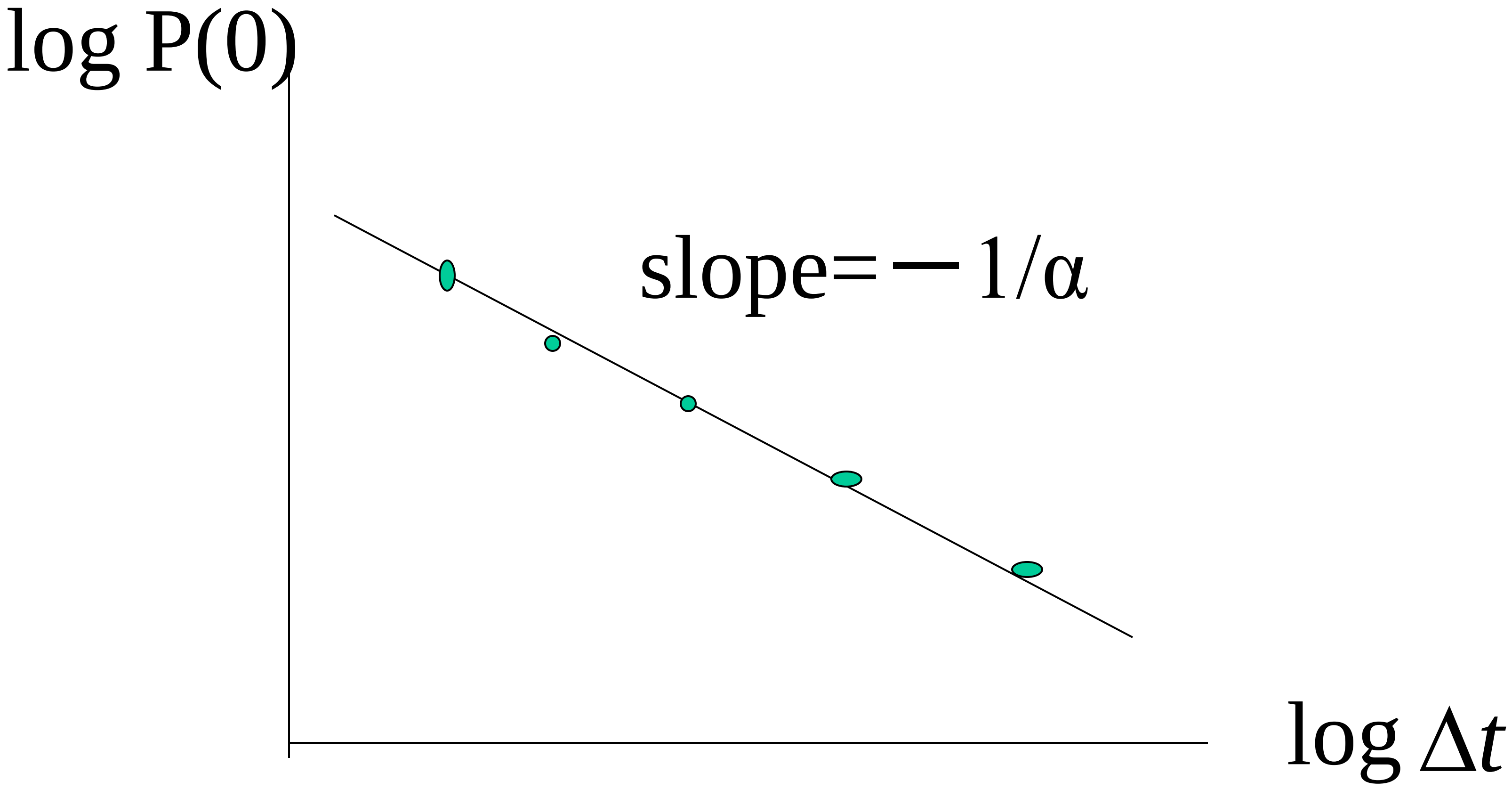
$$P_{\alpha,\beta}(x) = \lambda^{1/\alpha} P_{\alpha,\lambda\beta}(\lambda^{1/\alpha} x) \quad \cdots (3)$$

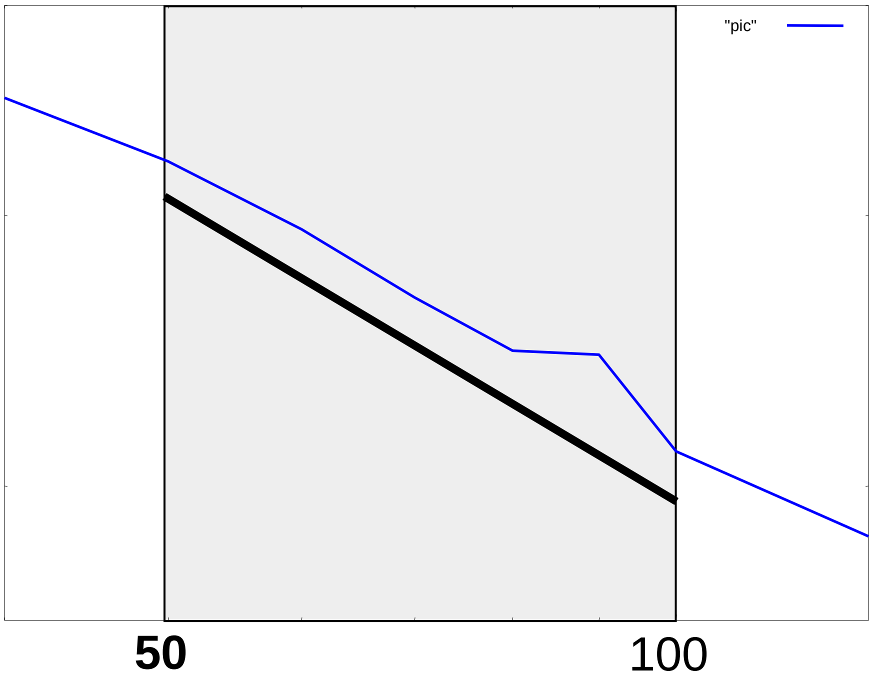
において、x=0とすると(4)式が得られる。

$$P_{\alpha,\lambda\beta}(0) = \lambda^{-1/\alpha} P_{\alpha,\beta}(0) \quad \cdots (4)$$

- (4)式を両対数グラフに表すことで、スケーリング領域において直線が現れる。

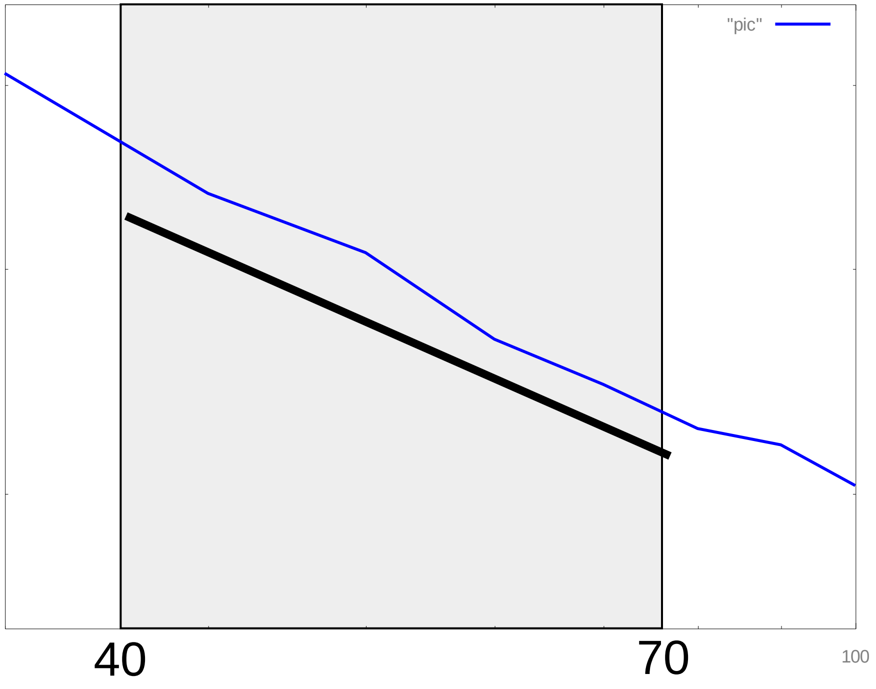
$$\log(P(0,\Delta t)) = -\frac{1}{\alpha}\cdot\log(\Delta t) + \log C$$



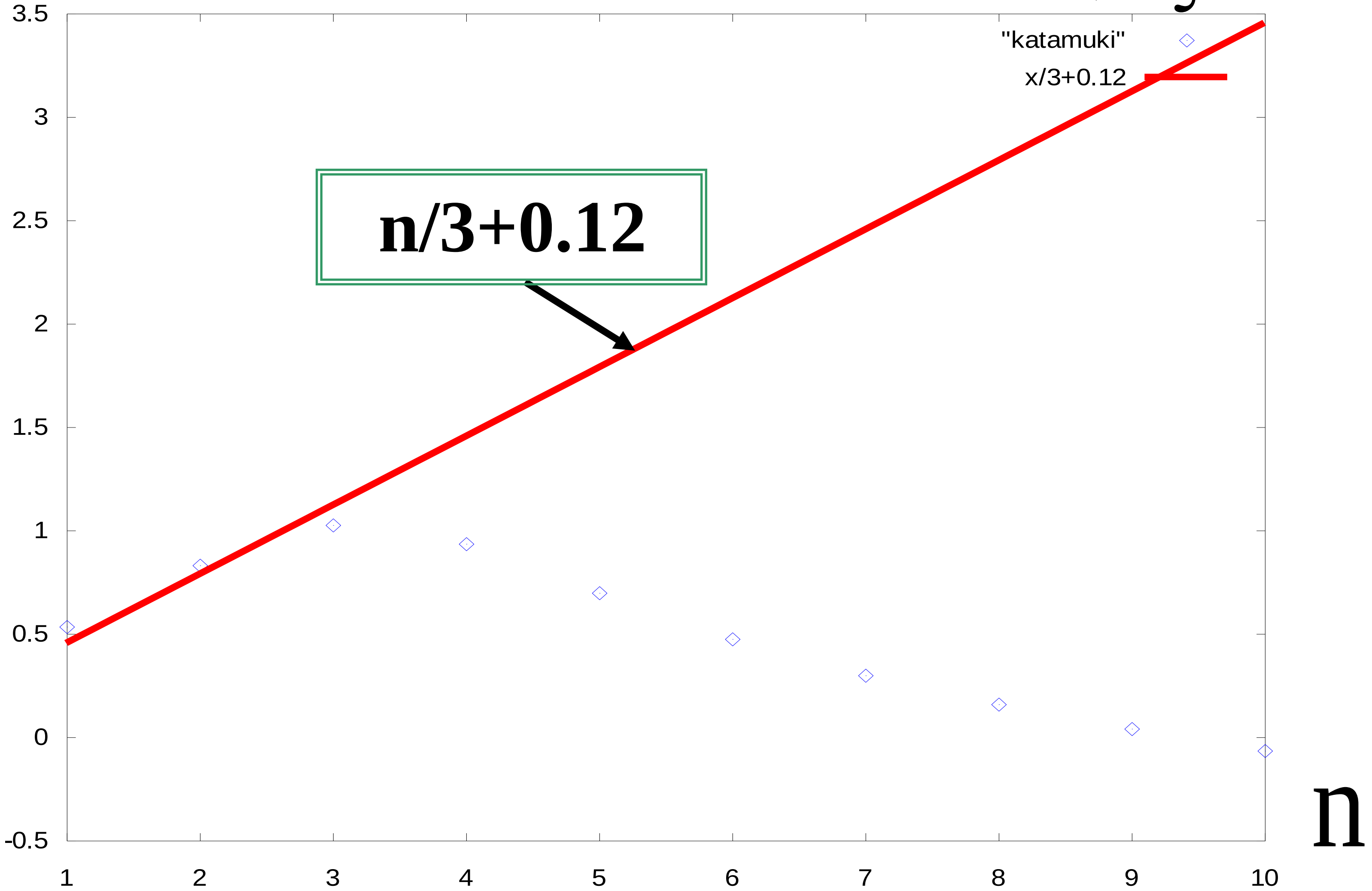


- FX1の原点回帰率
- スケール変換で求めたスケーリング領域で、直線が現れている。

- TRBの原点回帰率
- 為替と同様にスケール変換で求めたスケーリング領域で、直線が現れている。

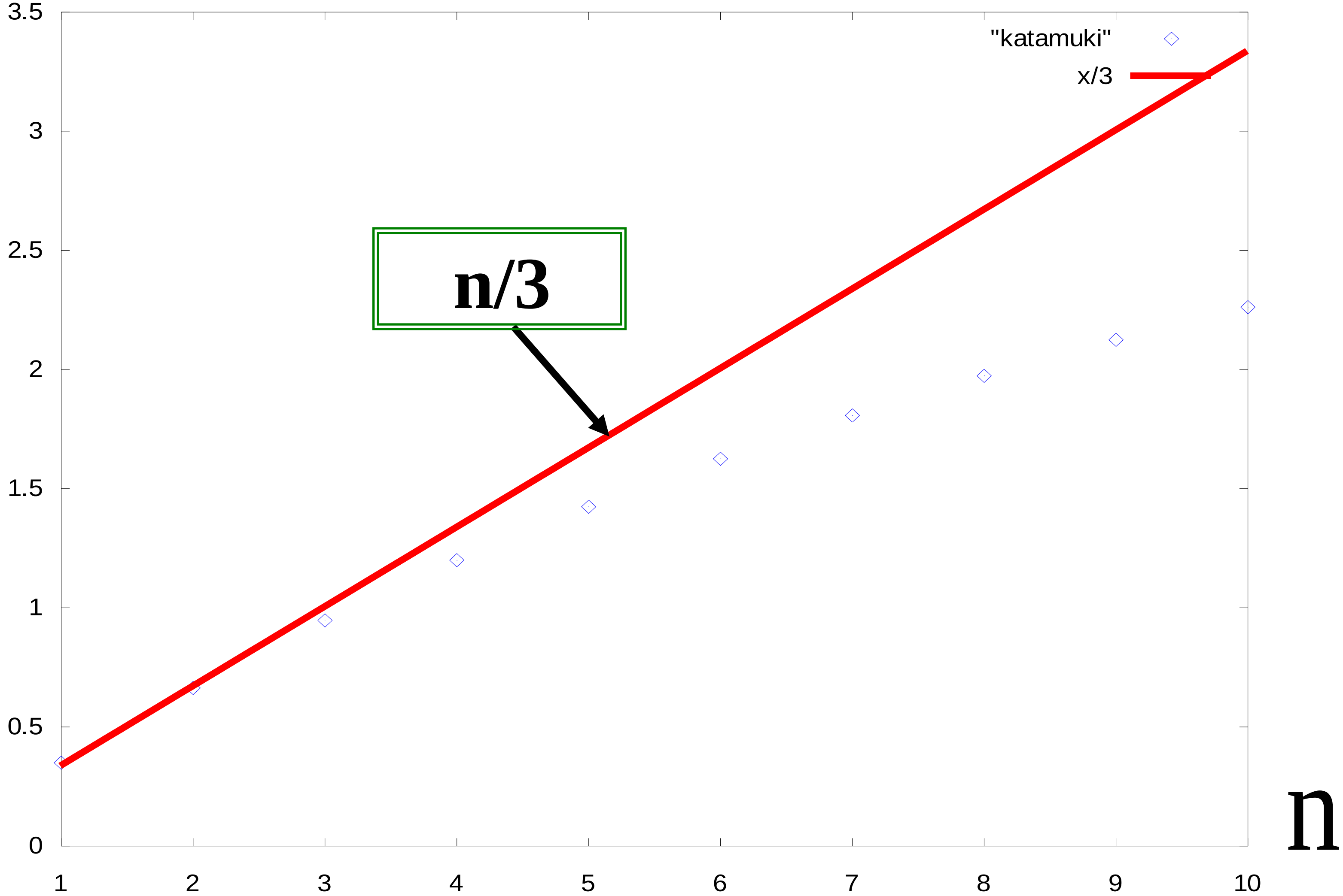


FX1のモーメントの指数ξ



- n<3で直線n/3+0.12に従っている。

TRBのモーメントの指数



- n<3で直線n/3に従っている。
- この結果はコルモゴロフの理論と一致する。

モーメントの類似

両者のモーメントの指数を調べた結果、

- 為替の指数ξ: ・FX1 $n/3+0.12$
  ・FX2 $n/3+0.2$
- 乱流の指数ζ: $n/3$

切片において違いが見られるが、両者とも傾きがn/3となり類似していると言える。

結論・考察

外国為替と乱流の類似点

1. 確率密度分布
2. 分布のモーメントが示す性質
3. 累積分布から見た分布形

- しかし、両者のヴォラティリティを調べるとデータの相関性に全く異なる性質が現れる。

問題点

- レヴィ分布のフィットをする際の数値積分
  （関数全体のフィット）

- 中心部分のフィット: P(0)

- 裾野部分のフィット: 累積分布のべき指数

三角裁定の話

Oct.13,1992 $/¥ bid ask $/M bid ask M/¥ bid ask

time	$/¥.bid	$/¥.ask	time	$/M.bid	$/M.ask	time	M/¥.ask	M/¥.ask
17:04:32	121.4	121.45	17:04:38	1.4692	1.4698	17:03:28	82.6	82.64
17:12:32	121.33	121.43	17:06:00	1.4685	1.4695	17:06:10	82.62	82.67

(A) ¥ → M → $ → ¥

82.64 yen→ 1 mark → 1/1.4698 dollar →
121.4/1.4698 = 82.60 yen

(B) ¥ → $ → M → ¥

121.45yen→1dollar→1.4692
mark→1.4692×82.6= 121.36 yen

Find time interval where either

$R_A > 1$ or $R_B > 1$

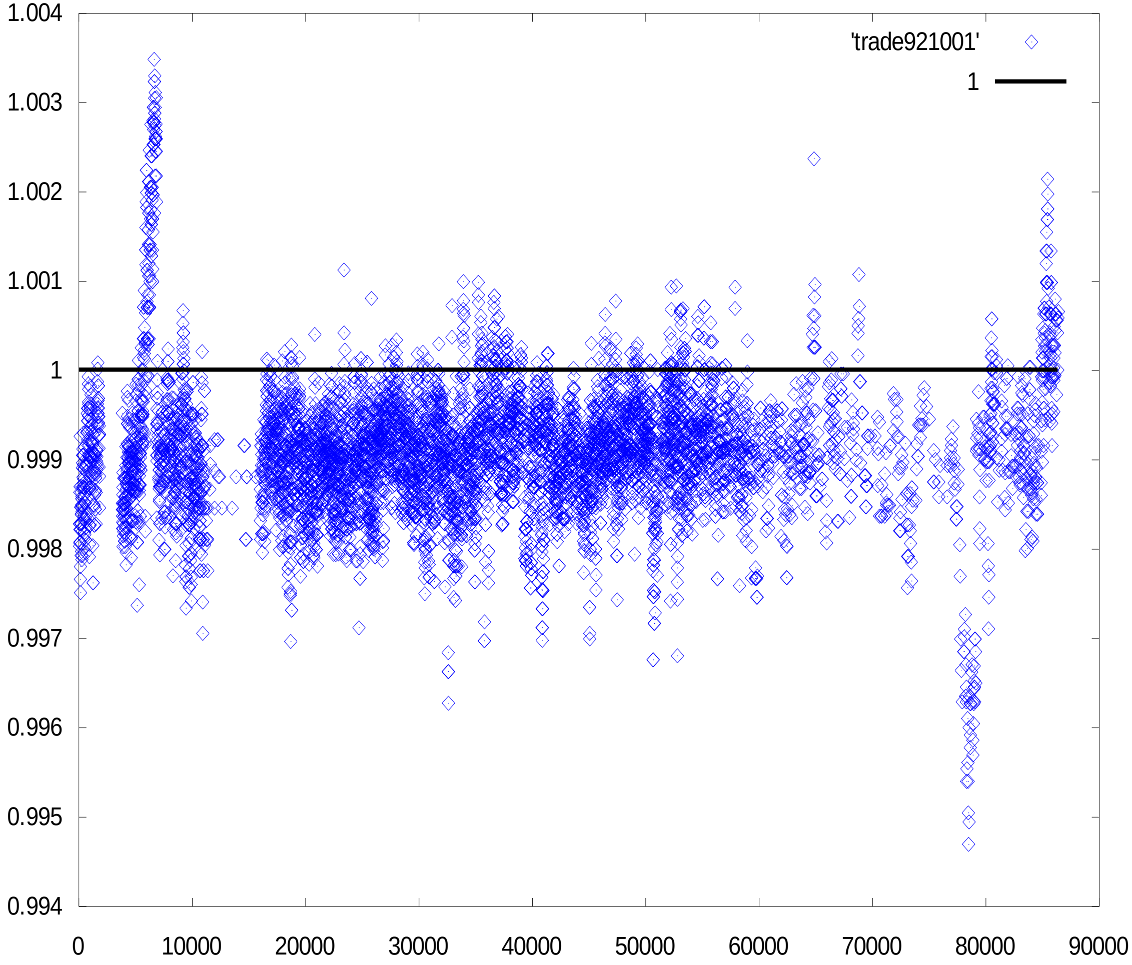
$$\mathrm{R_A} = \frac{\$/\mathrm{Y.b}}{(\$/\mathrm{M.a})\cdot(\mathrm{M}/\mathrm{Y.a})}$$

$$\mathrm{R_B} = \frac{\$/\mathrm{M.b}}{(\mathrm{Y}/\mathrm{M.a})(\$/\mathrm{M.a})}$$

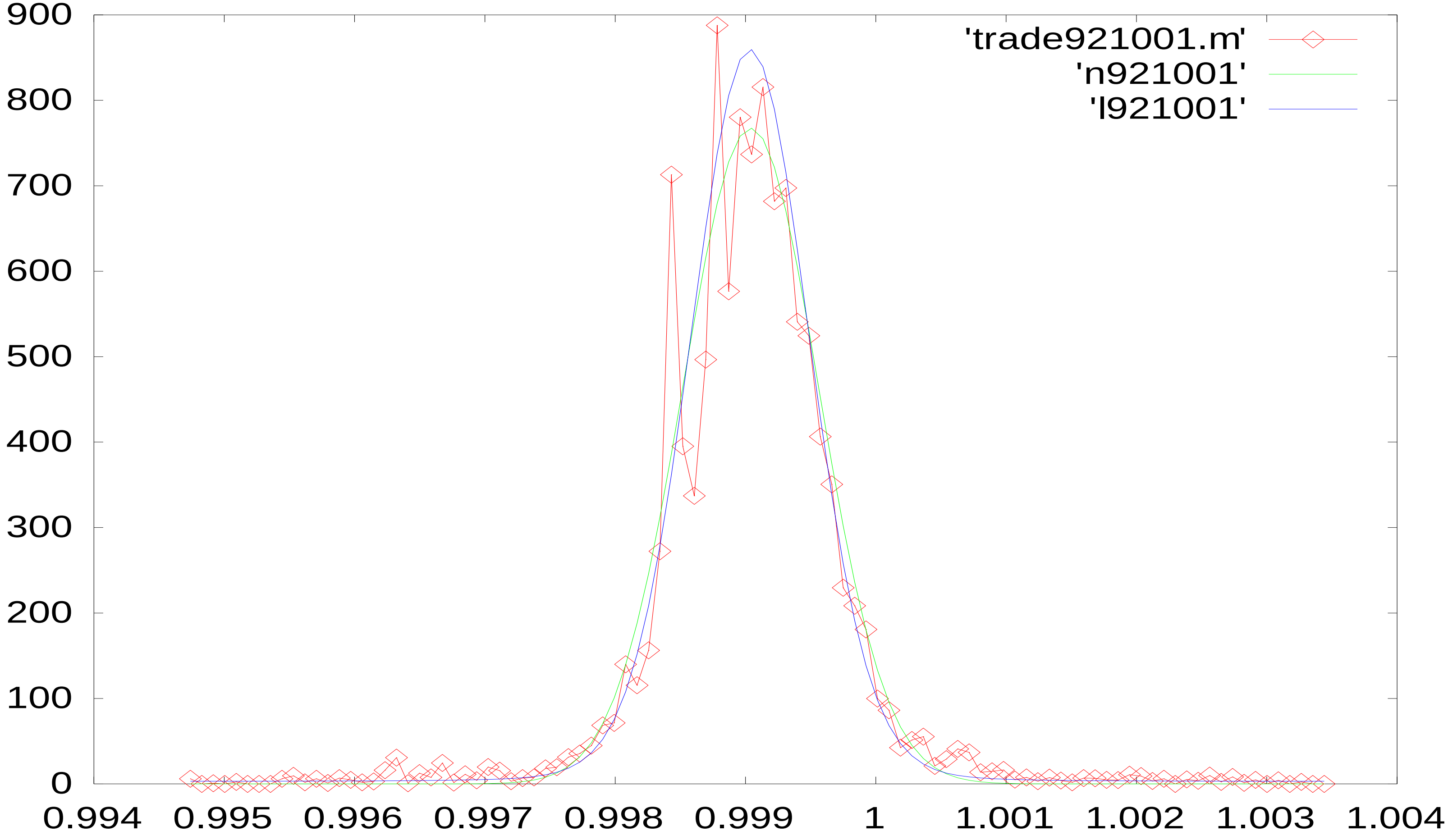
Usually both are < 1.

Occasionally one of them > 1.

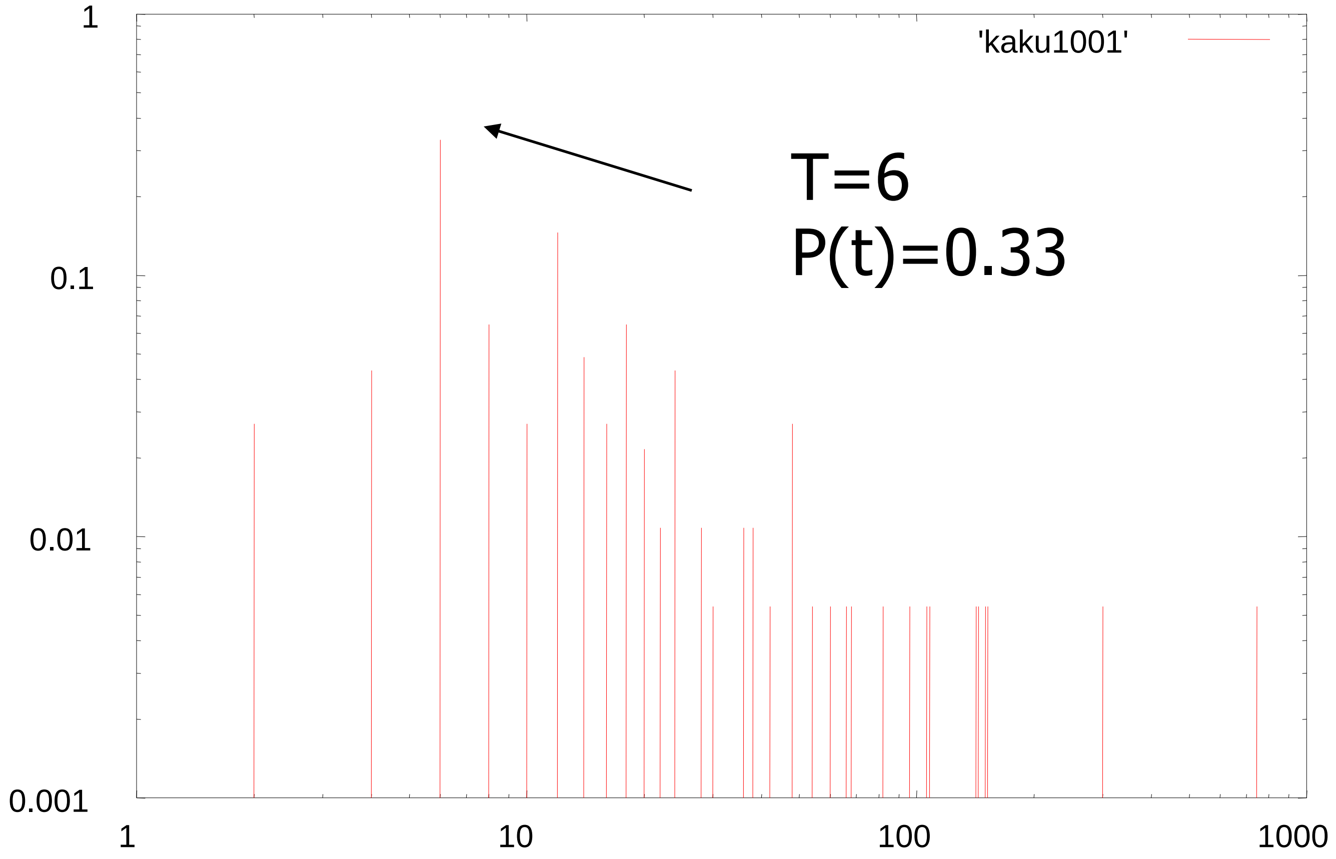
Arbitrage gain R for1day on 1992/10/1



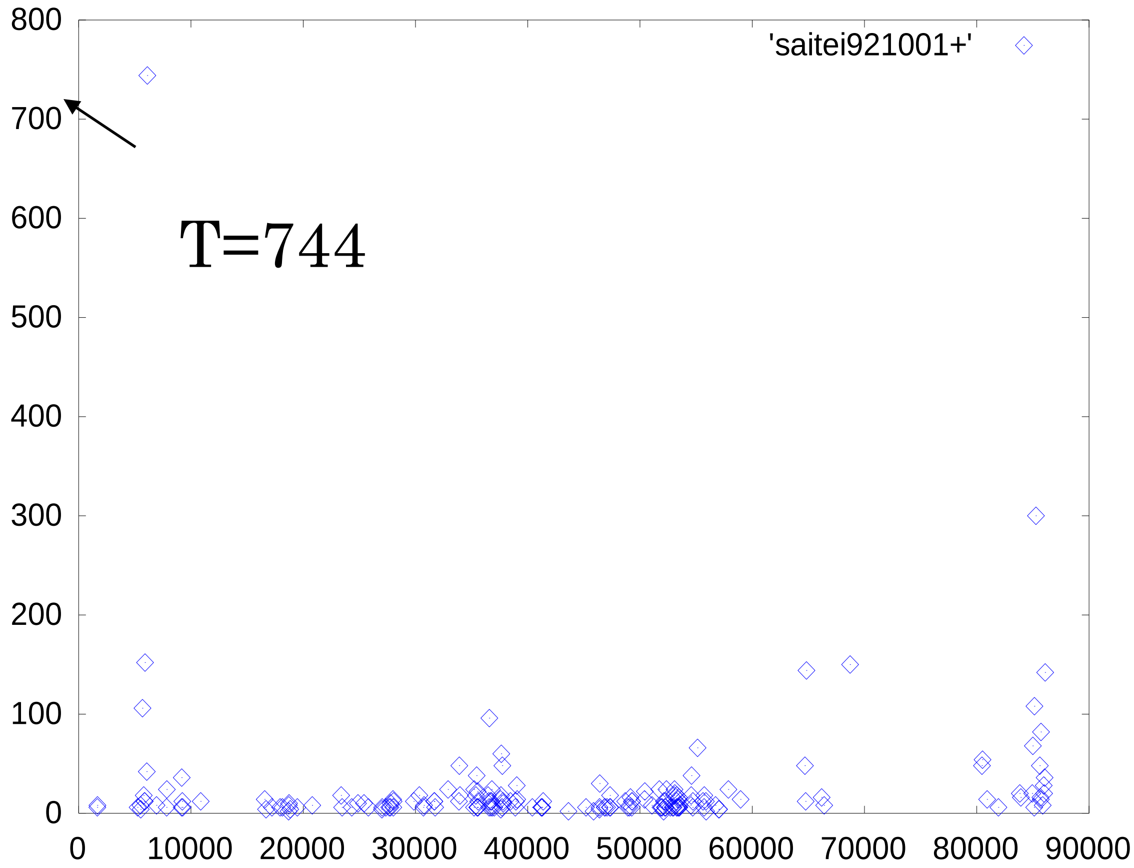
Arbitrage chance on Oct.1, 1992
P(u>1)=0.052 (75 minutes)



How long each chance lasts?



How long did it continue each time?

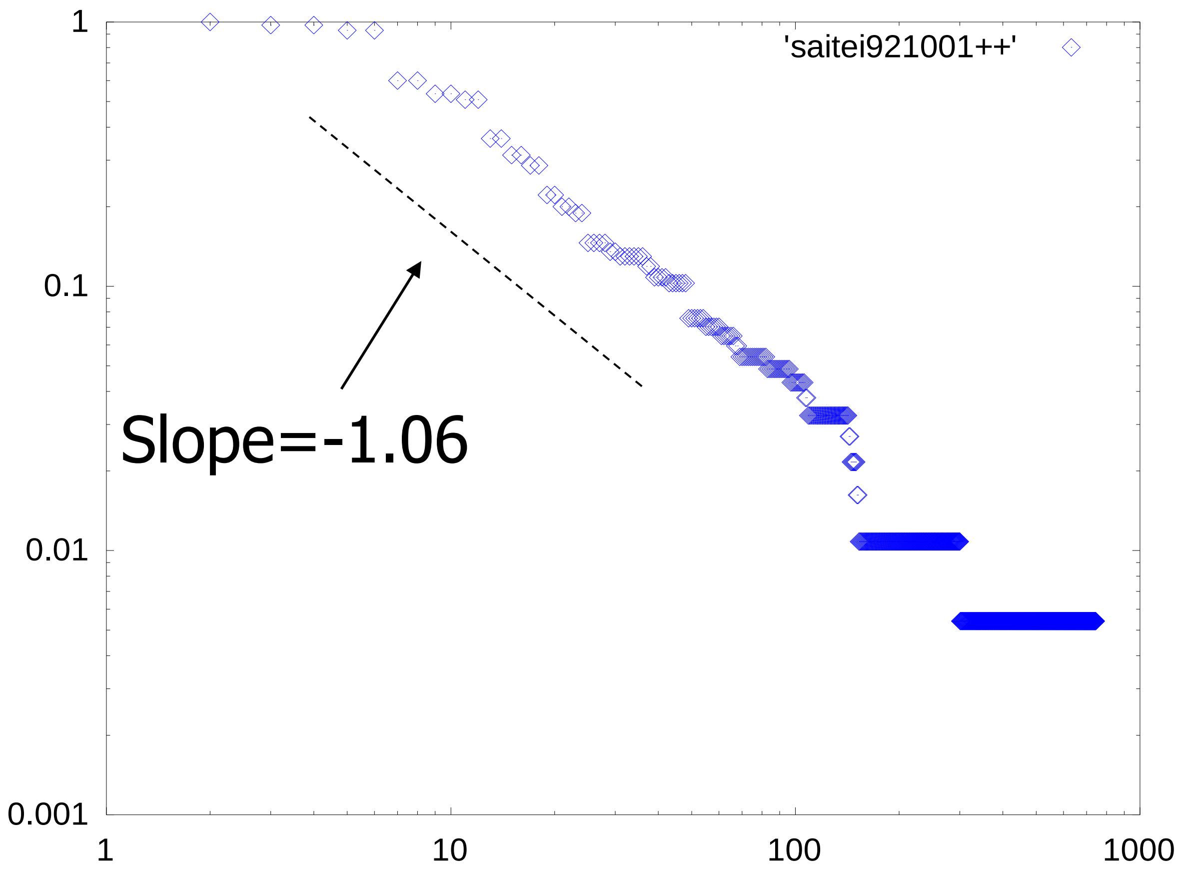


T=744

$/¥	bid	ask	$/M	bid	ask	M/¥	bid	ask	¥_$_M_¥	¥_M_$_¥
			0:00:14	1.4116	1.4121					
0:00:32	119.9	120		1.4116	1.4121					
0:00:42	119.93	119.98		1.4116	1.4121					
0:00:48	119.93	120		1.4116	1.4121					
0:00:54	119.93	120	0:00:54	1.4108	1.4118					
0:01:06	119.93	120.03	0:01:00	1.411	1.412					
	119.93	120.03		1.411	1.412	0:01:02	84.9	84.95	0.998033	0.999838
0:01:12	119.9	120		1.411	1.412		84.9	84.95	0.998283	0.999588
	119.9	120	0:01:18	1.4115	1.412		84.9	84.95	0.998636	0.999588
	119.9	120	0:01:24	1.4107	1.4117		84.9	84.95	0.99807	0.999801
0:01:30	119.95	120.05	0:01:30	1.4115	1.4125		84.9	84.95	0.99822	0.999651
0:01:42	119.9	119.95	0:01:36	1.4113	1.4123		84.9	84.95	0.998911	0.999376
0:02:06	120	120.05	0:01:42	1.411	1.412		84.9	84.95	0.997867	1.000422
0:02:12	120	120.1	0:01:54	1.4118	1.4128		84.9	84.95	0.998017	0.999855
0:02:18	120	120.1	0:02:06	1.4113	1.4123		84.9	84.95	0.997663	1.000209
0:02:24	120	120.05	0:02:18	1.4115	1.413		84.9	84.95	0.99822	0.999714
0:02:30	119.97	120.02	0:02:24	1.4105	1.4115		84.9	84.95	0.997762	1.000526
0:02:42	119.95	120.05	0:02:30	1.4115	1.4125		84.9	84.95	0.99822	0.999651
0:03:00	119.95	120	0:02:48	1.4115	1.4125		84.9	84.95	0.998636	0.999651
0:03:14	119.95	120	0:02:54	1.4116	1.4121		84.9	84.95	0.998707	0.999934
0:03:26	119.9	120	0:03:06	1.411	1.412		84.9	84.95	0.998283	0.999588
0:03:32	119.98	120.03	0:03:20	1.4105	1.4115		84.9	84.95	0.997679	1.000609
0:03:38	119.92	120.02	0:03:26	1.4115	1.4125		84.9	84.95	0.99847	0.999401
0:04:04	119.95	120	0:03:52	1.4108	1.4118		84.9	84.95	0.998141	1.000147
0:04:10	119.9	120	0:03:58	1.4114	1.4121		84.9	84.95	0.998566	0.999517
0:04:28	119.91	119.98	0:04:04	1.4113	1.4118		84.9	84.95	0.998661	0.999813
0:04:42	119.95	120	0:04:10	1.4114	1.4121		84.9	84.95	0.998566	0.999934
0:05:04	119.9	119.95	0:04:28	1.4102	1.4112		84.9	84.95	0.998132	1.000155
0:05:20	119.91	119.96	0:04:34	1.4112	1.4117		84.9	84.95	0.998757	0.999884
0:05:34	119.88	119.95	0:04:46	1.4105	1.4115		84.9	84.95	0.998345	0.999775
0:05:54	119.85	119.95	0:05:00	1.4113	1.4123		84.9	84.95	0.998911	0.998959
0:05:58	119.9	119.95	0:05:14	1.4103	1.4113		84.9	84.95	0.998203	1.000084
0:06:04	119.85	119.92		1.4103	1.4113	0:05:18	84.93	84.96	0.998806	0.999549
	119.85	119.92	0:05:26	1.411	1.412		84.93	84.96	0.999301	0.999054
	119.85	119.92	0:05:40	1.411	1.4125		84.93	84.96	0.999301	0.9987
	119.85	119.92	0:05:48	1.4105	1.4115		84.93	84.96	0.998947	0.999408
	119.85	119.92	0:05:54	1.4105	1.4115		84.93	84.96	0.998947	0.999408
	119.85	119.92	0:05:58	1.4111	1.4116		84.93	84.96	0.999372	0.999337
	119.85	119.92	0:06:04	1.4103	1.4113		84.93	84.96	0.998806	0.999549
0:06:10	119.85	119.92		1.4103	1.4113		84.93	84.96	0.998806	0.999549
0:06:22	119.85	119.95	0:06:22	1.411	1.412		84.93	84.96	0.999052	0.999054
0:06:42	119.85	119.92	0:06:36	1.411	1.4115		84.93	84.96	0.999301	0.999408
0:06:52	119.85	119.95	0:06:46	1.4105	1.4115		84.93	84.96	0.998697	0.999408
0:07:00	119.87	119.92	0:07:00	1.4102	1.4112		84.93	84.96	0.998735	0.999787
0:07:06	119.85	119.9		1.4102	1.4112		84.93	84.96	0.998901	0.99962
0:07:16	119.85	119.95		1.4102	1.4112	0:07:10	84.92	84.97	0.998368	0.999502

0:07:24	119.88	119.98		1.4102	1.4112	0:11:32	84.95	85	0.99847	0.9994
0:08:04	119.85	119.95	0:07:30	1.411	1.412	0:16:34	84.96	85.01	0.999404	0.998466
0:08:12	119.9	119.95	0:07:52	1.4105	1.4115	0:16:50	84.98	85.03	0.999285	0.999001
0:08:18	119.88	119.98	0:07:58	1.4113	1.4118	0:17:26	85	85.05	0.999837	0.998388
0:08:24	119.9	120	0:08:18	1.4103	1.4113	0:18:16	84.95	85	0.998375	0.999496
0:08:46	119.88	119.98	0:08:30	1.411	1.412	0:20:00	84.96	85.01	0.999155	0.998716
0:09:08	119.9	120	0:08:40	1.4105	1.4115	0:20:18	84.95	85	0.998516	0.999354
0:09:14	119.9	119.95	0:08:52	1.411	1.412	0:21:06	84.95	85	0.999287	0.999
0:09:20	119.92	120.02	0:09:50	1.4115	1.4125	0:23:00	85	85.05	0.999646	0.998226
0:09:30	119.9	120	0:10:02	1.4112	1.4117	0:23:18	85.02	85.07	0.999835	0.99839
0:09:42	119.95	120.05	0:10:28	1.4115	1.412	0:23:36	85	85.05	0.999396	0.998829
0:09:50	119.93	120.03	0:10:34	1.4113	1.4118	0:23:56	84.96	85.01	0.998951	0.999274
0:10:16	119.98	120.05	0:10:40	1.4108	1.4118	0:25:44	84.97	85.02	0.998548	0.999573
0:10:28	119.95	120.02	0:10:58	1.4116	1.4123	0:29:38	84.94	84.99	0.999011	0.999322
0:10:34	119.95	120.05	0:11:04	1.4115	1.412	0:29:56	84.95	85	0.998808	0.999417
0:10:52	119.95	120	0:11:58	1.4107	1.4117	0:30:22	84.91	84.96	0.998188	1.0001
0:11:24	119.96	119.99	0:12:20	1.4118	1.4125	1:11:14	84.96	85.01	0.999638	0.999029
0:12:46	119.95	120	0:12:34	1.4115	1.4125	1:11:58	84.94	84.99	0.999107	0.999181
0:13:18	119.95	120.05	0:12:46	1.4118	1.4123	1:17:16	85	85.05	0.999608	0.998617
0:13:24	119.9	120	0:12:54	1.412	1.413	1:22:38	84.98	85.03	0.999931	0.997941
0:13:56	119.96	120.01	0:13:00	1.4115	1.4125	1:23:12	84.95	85	0.999141	0.999146
0:14:22	119.95	120	0:13:06	1.412	1.413	1:24:20	84.9	84.95	0.99899	0.999297
0:14:34	119.92	120.02	0:13:18	1.412	1.413	1:54:16	84.6	84.7	0.995294	1.001996
0:14:48	120.02	120.07	0:13:30	1.4118	1.4123	2:00:08	84.6	84.7	0.994739	1.003329
0:15:04	119.95	120.05	0:13:42	1.412	1.4125	2:03:14	84.6	84.7	0.995045	1.002602
0:15:12	120.05	120.15	0:13:50	1.4125	1.414	2:07:40	84.65	84.7	0.995157	1.002373
0:15:22	120.05	120.15	0:14:04	1.412	1.4135	2:11:26	84.58	84.63	0.993982	1.003557
0:15:28	120.1	120.2	0:14:10	1.412	1.413	2:13:26	84.57	84.62	0.993451	1.004449
0:15:40	120.05	120.15	0:14:16	1.4118	1.4128	2:14:32	84.56	84.61	0.993606	1.004291
0:15:46	120.05	120.15	0:14:22	1.4121	1.4126	2:17:18	84.51	84.56	0.99323	1.005028
0:15:52	120.1	120.2	0:14:28	1.4123	1.4133	2:18:48	84.5	84.6	0.99284	1.004473
0:16:10	120.1	120.2	0:14:40	1.4115	1.4125	2:24:04	84.55	84.6	0.992865	1.005042
0:16:24	120.03	120.13	0:14:48	1.4121	1.4126	2:28:56	84.47	84.52	0.992925	1.005336
0:16:30	120.05	120.15	0:14:58	1.412	1.413	2:36:46	84.41	84.45	0.991984	1.006052
0:16:42	120.05	120.1	0:15:12	1.4125	1.4135	2:37:16	84.42	84.46	0.992866	1.005577
0:16:48	120.09	120.14	0:15:18	1.413	1.4145	2:37:58	84.43	84.47	0.993005	1.005082
0:16:54	120.1	120.2	0:15:22	1.412	1.413	2:39:28	84.41	84.45	0.991572	1.006471
0:17:06	120.05	120.15	0:15:34	1.4122	1.4132	2:43:36	84.45	84.49	0.992595	1.005433
0:17:12	120.1	120.2	0:15:40	1.4125	1.4135	2:44:06	84.43	84.47	0.992158	1.005877
0:17:24	120.1	120.2	0:15:58	1.4125	1.4135	2:45:34	84.48	84.52	0.992745	1.005282
0:17:30	120.15	120.17	0:16:18	1.4125	1.414	2:47:56	84.43	84.48	0.992406	1.00582
0:17:36	120.1	120.2	0:16:42	1.413	1.4145	2:50:02	84.51	84.54	0.99345	1.004333
0:17:48	120.08	120.18	0:16:48	1.4135	1.415	2:50:24	84.46	84.51	0.993378	1.004167
0:18:04	120.05	120.15	0:16:54	1.413	1.4135	2:51:26	84.44	84.49	0.99304	1.00522
0:18:22	120.05	120.15	0:17:06	1.4125	1.4135	2:56:54	84.45	84.55	0.992806	1.004506
0:18:32	120	120.1	0:17:12	1.413	1.4145	3:04:08	84.47	84.52	0.993806	1.003734
0:18:38	120.03	120.1	0:17:36	1.413	1.414	3:04:56	84.45	84.5	0.993571	1.004578

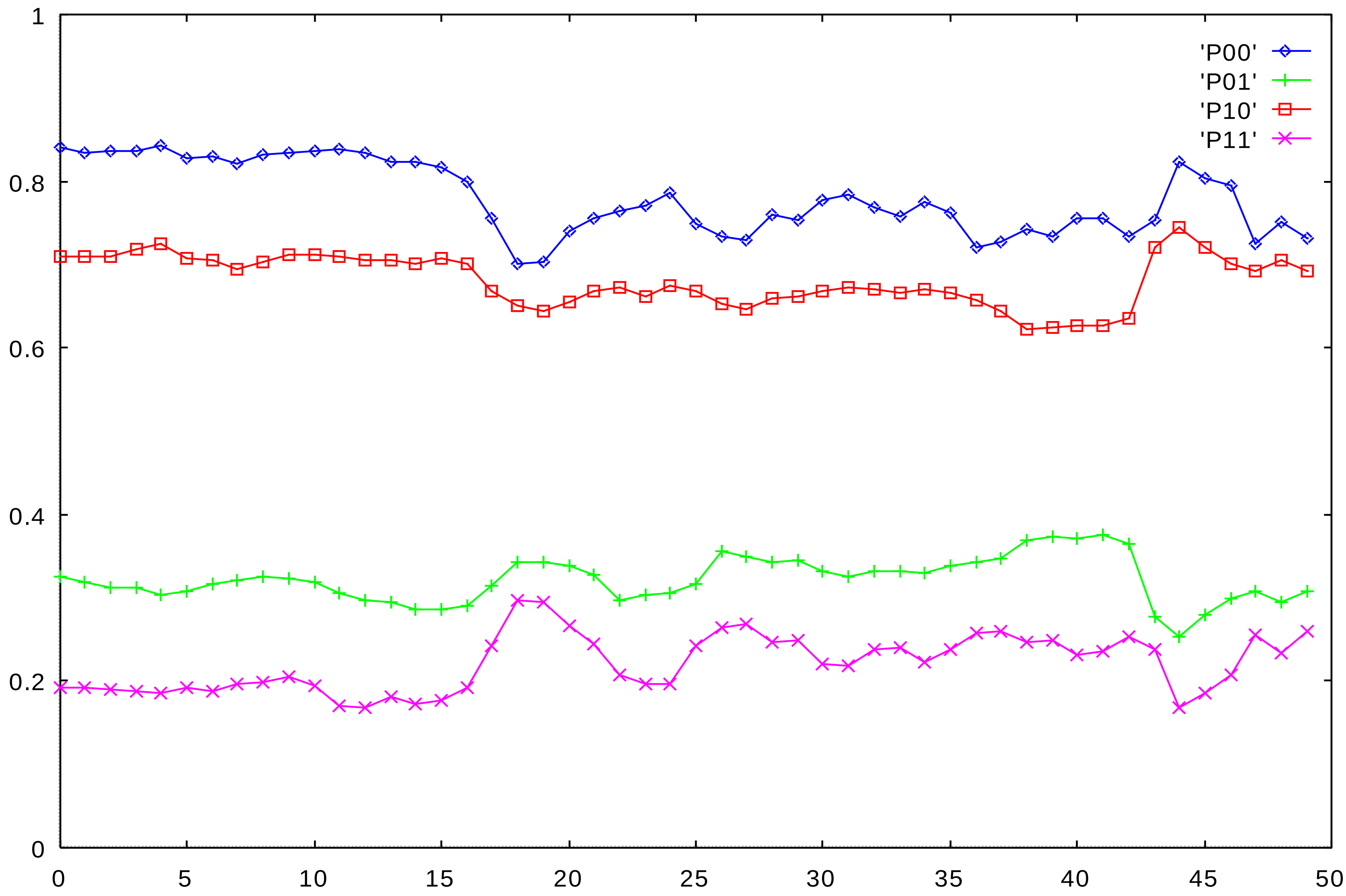
Accumulated distribution P(T) for 1 day



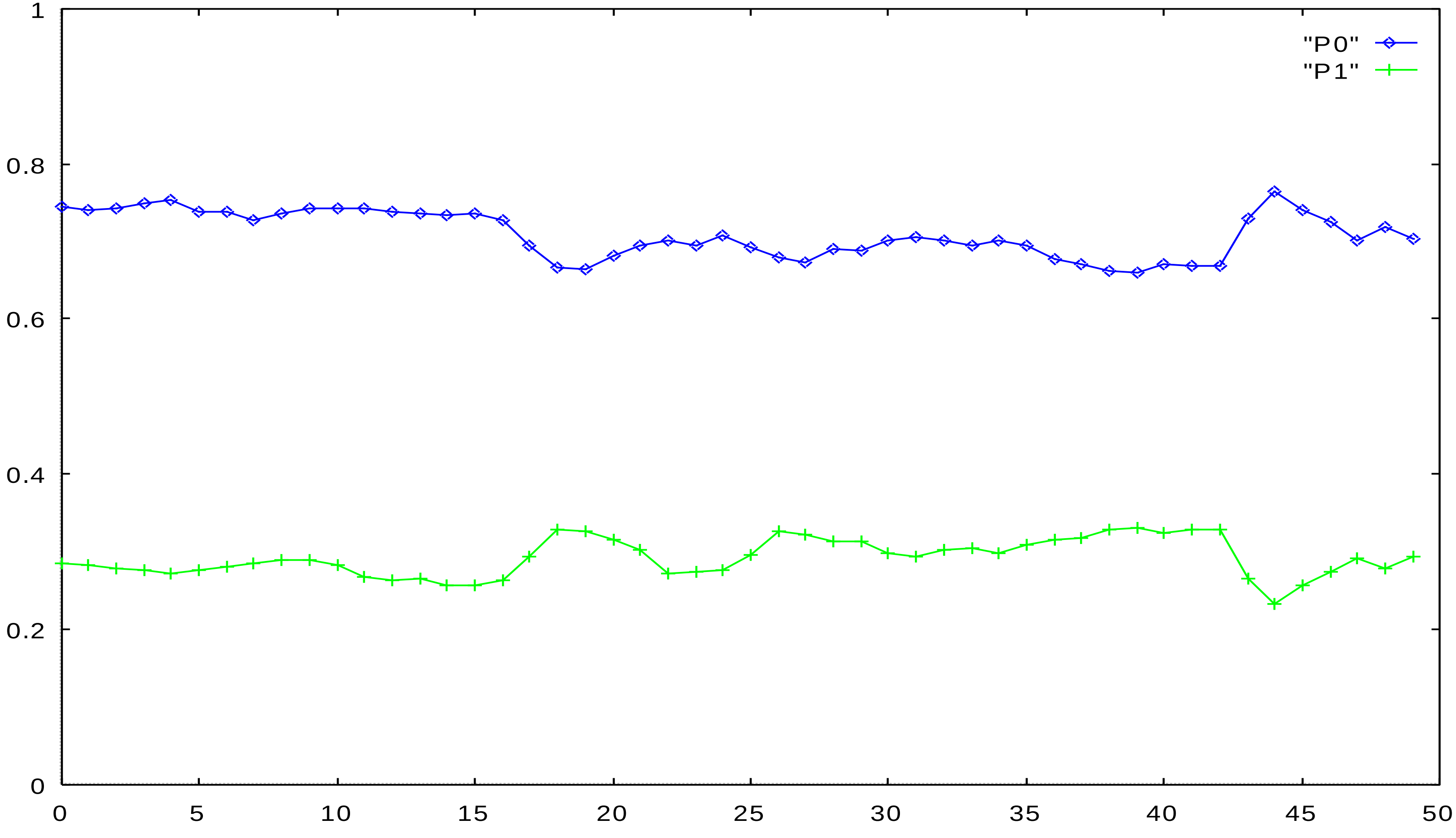
Zip rule

記憶長の話

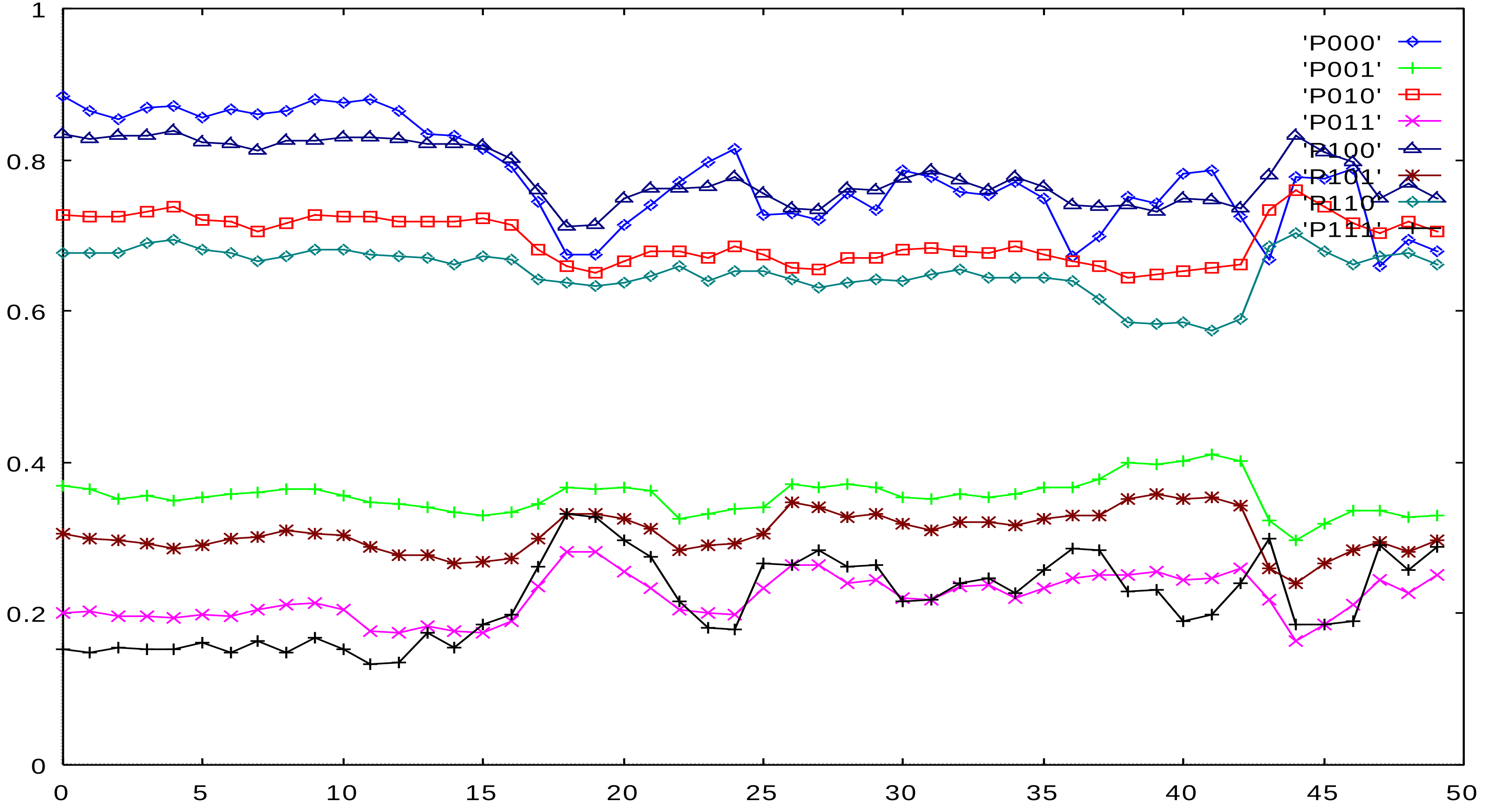
Depth=2　P(1|00), P(1|01), P(1|10), P(1|11)



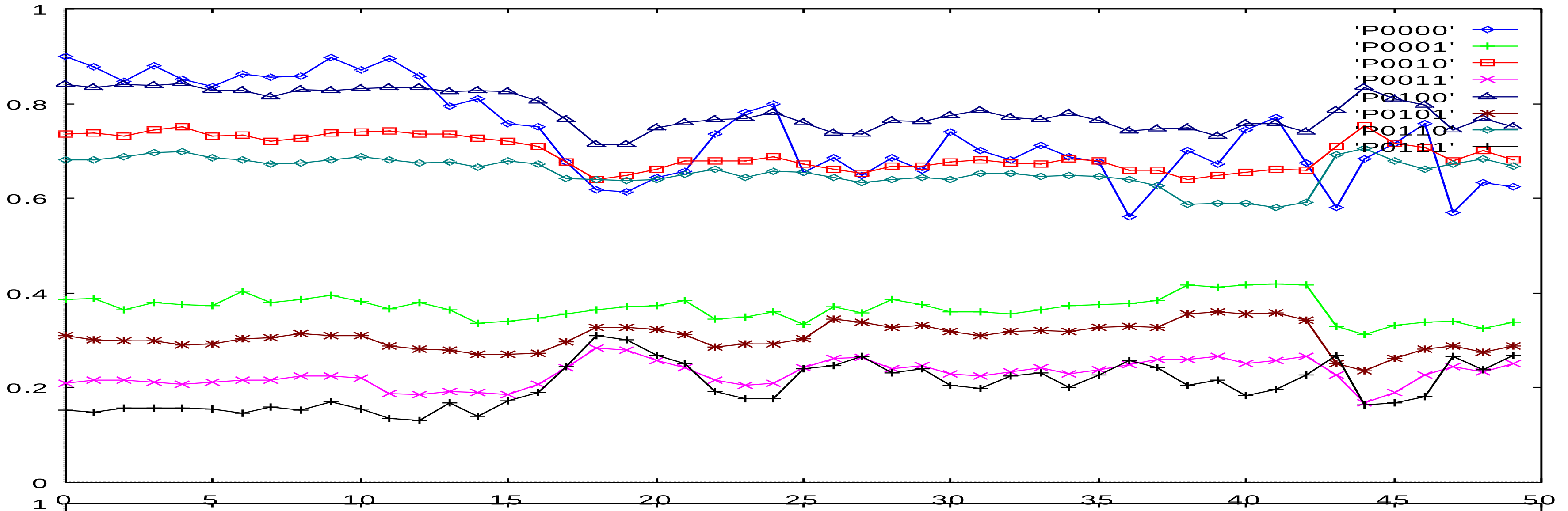
USDJPY1995.1.2－2001.4.12(200000*50)
Depth＝1　　　　P(1|0) and P(1|1)

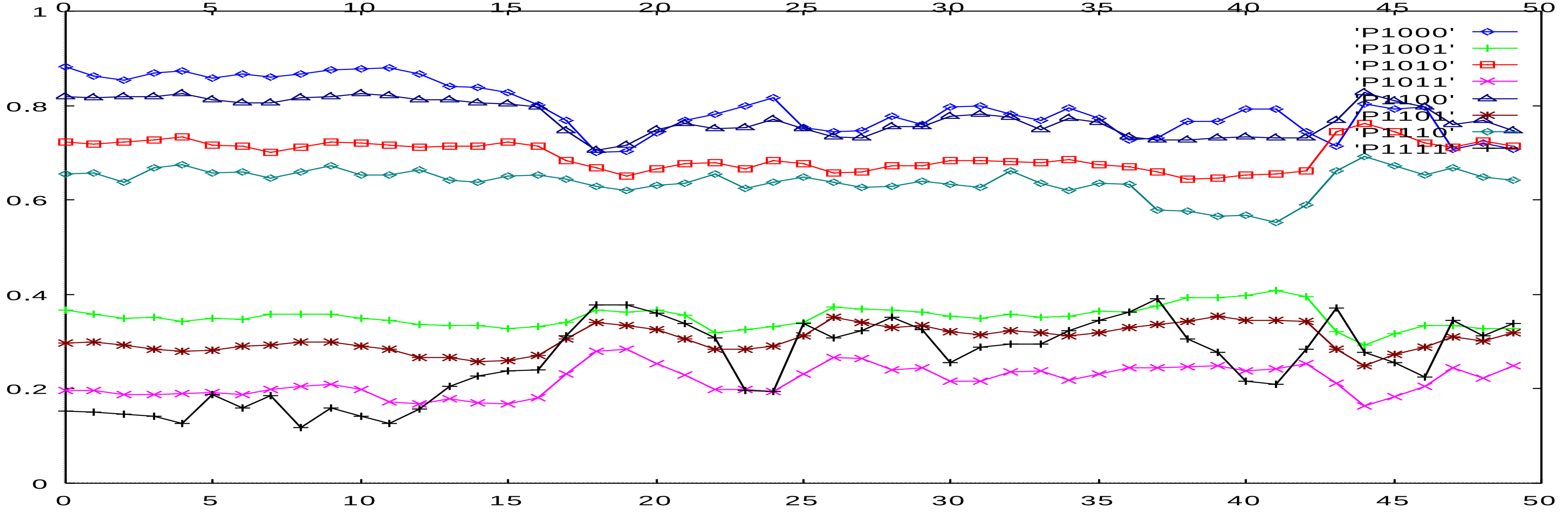


Depth=3
P(1|000),P(1|001),P(1|010), P(1|011), P(1|100),P(1|101),P(1|110), P(1|111)

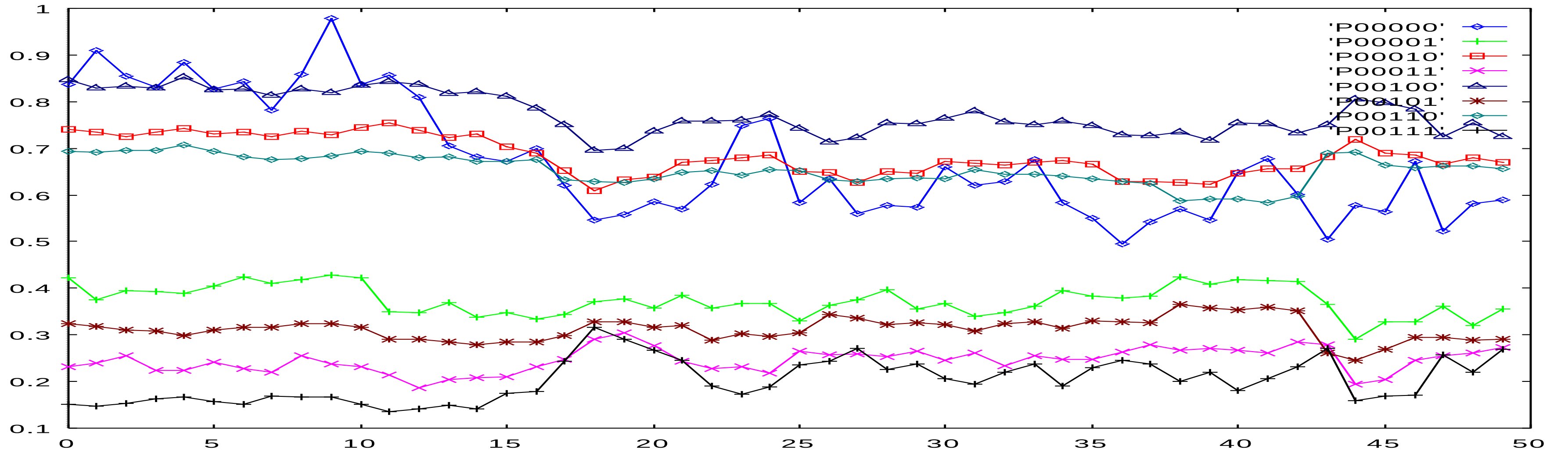


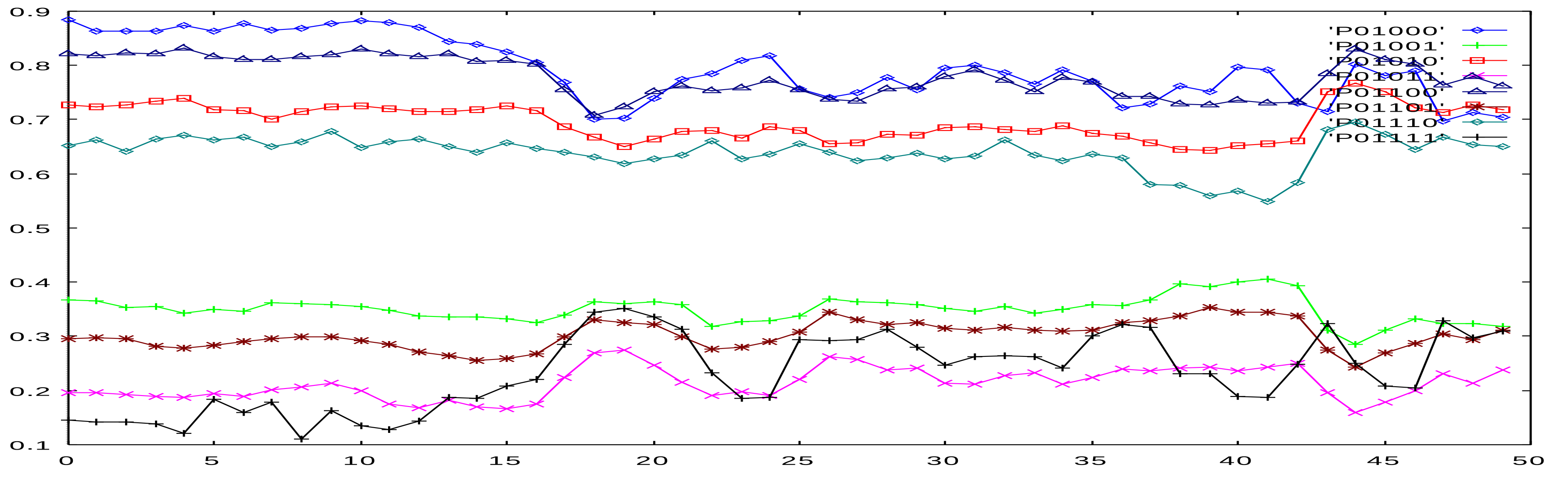
Depth=4 P(1|0000), P(1|0001), etc.



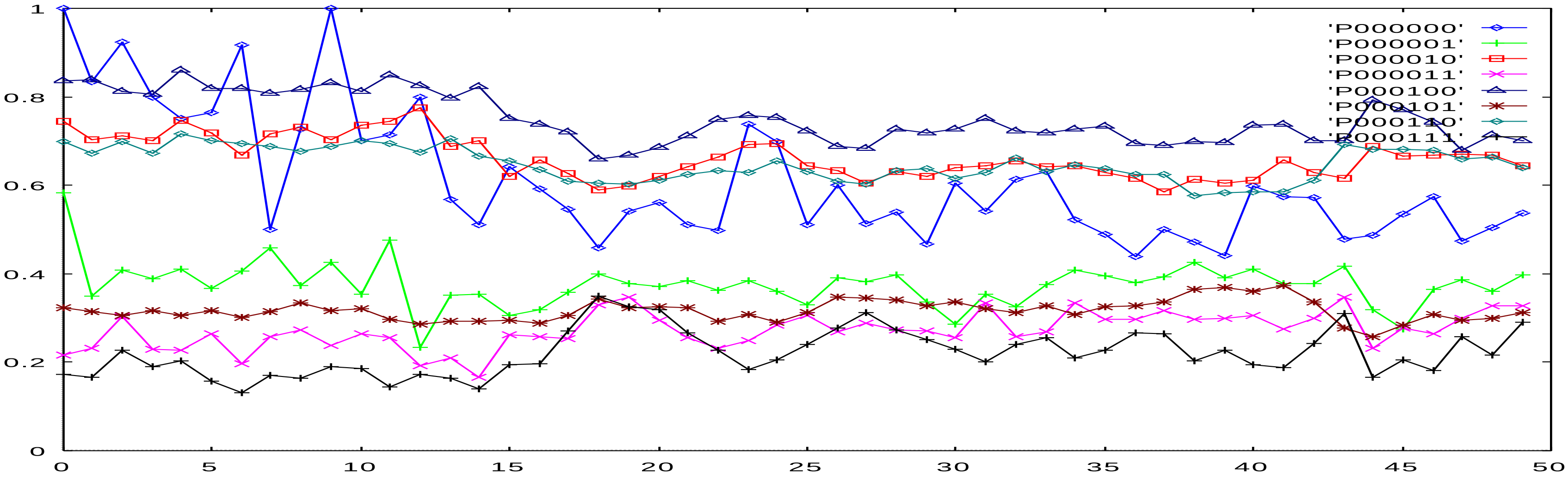


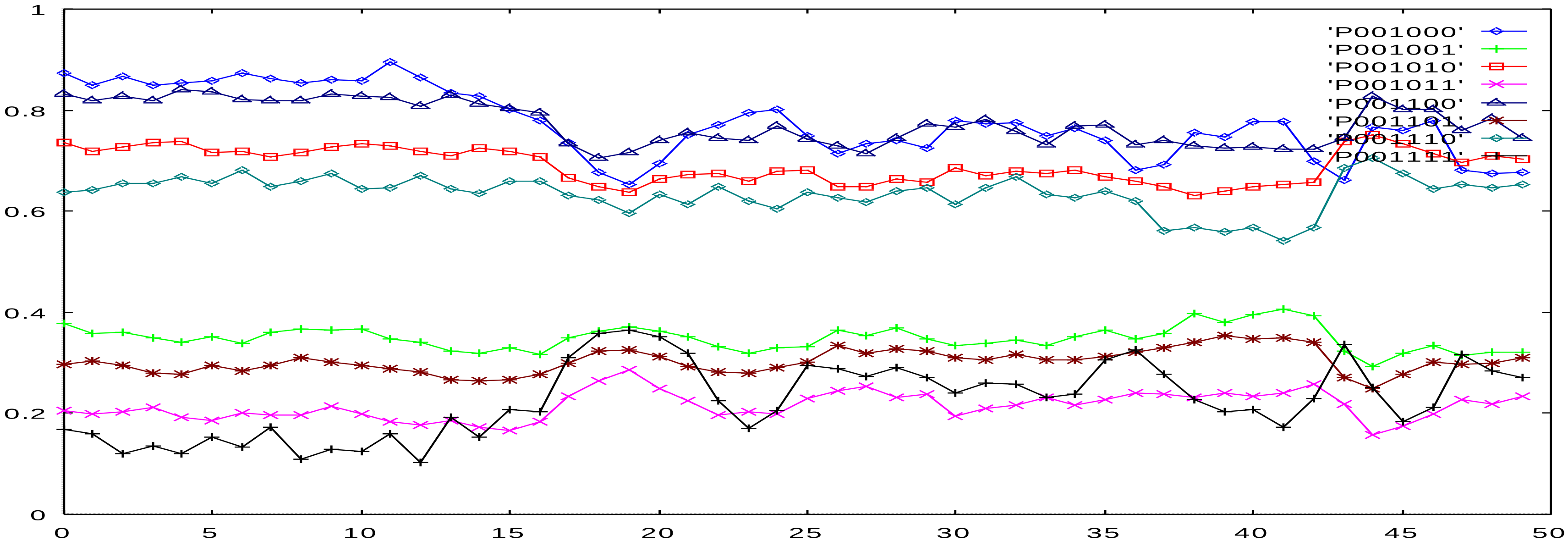
Depth=5　P(1|00000), P(1|00001), etc.



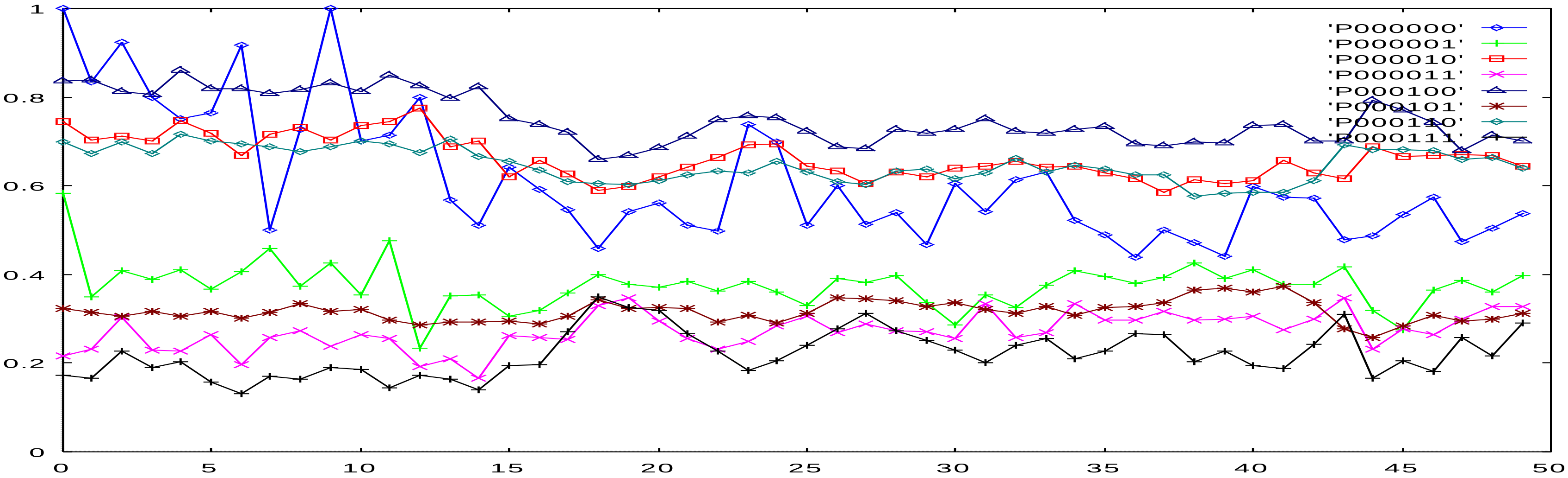


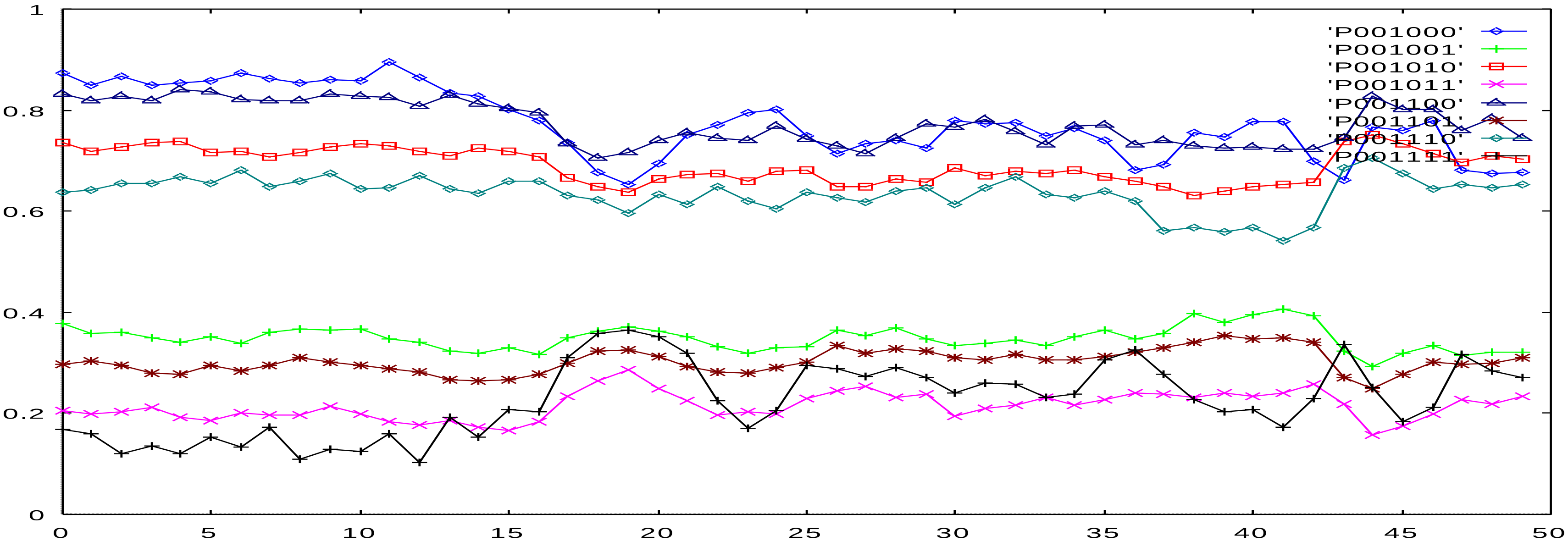
Depth=6 P(1|000000),P(1|000001), etc.



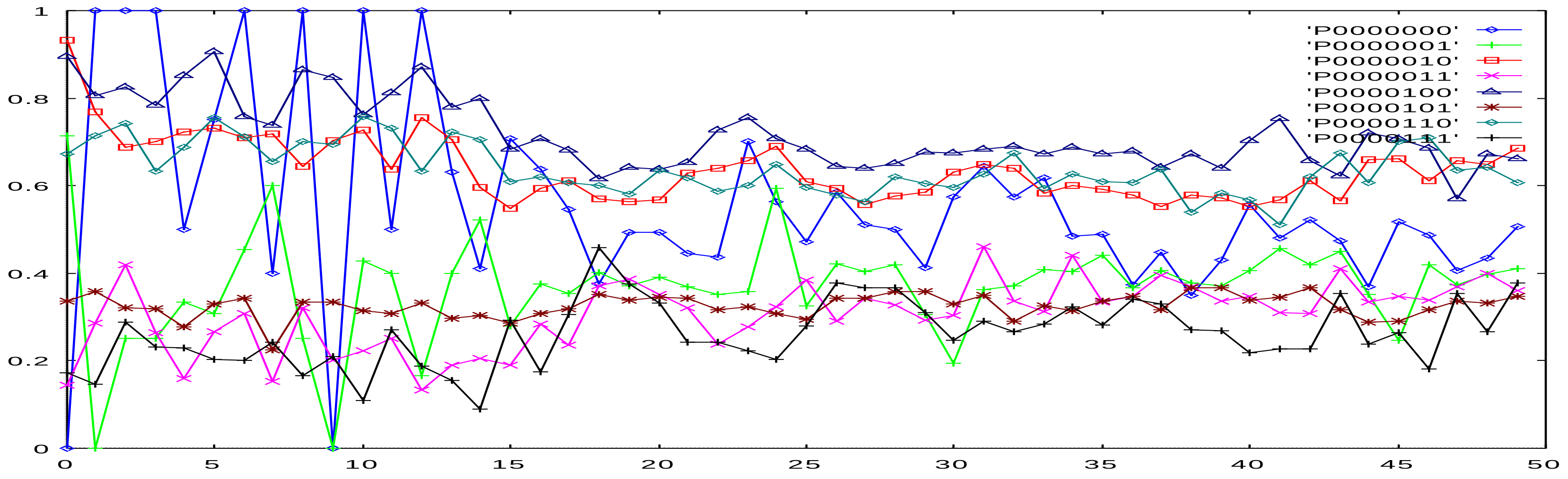


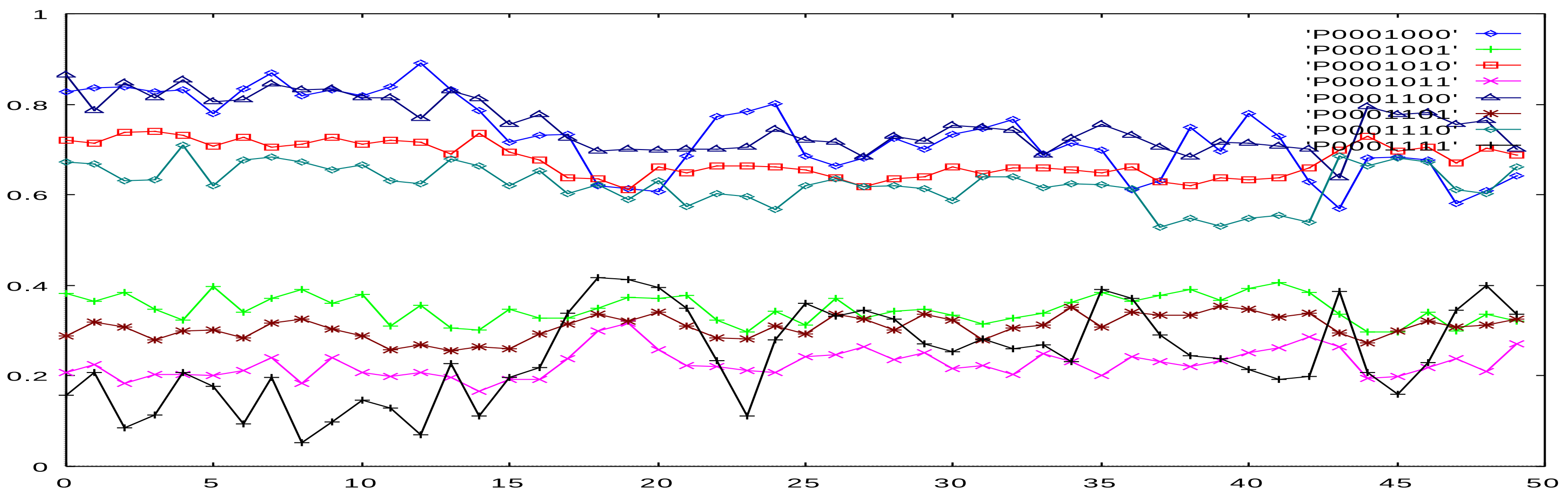
Depth=6 P(1|000000),P(1|000001), etc.



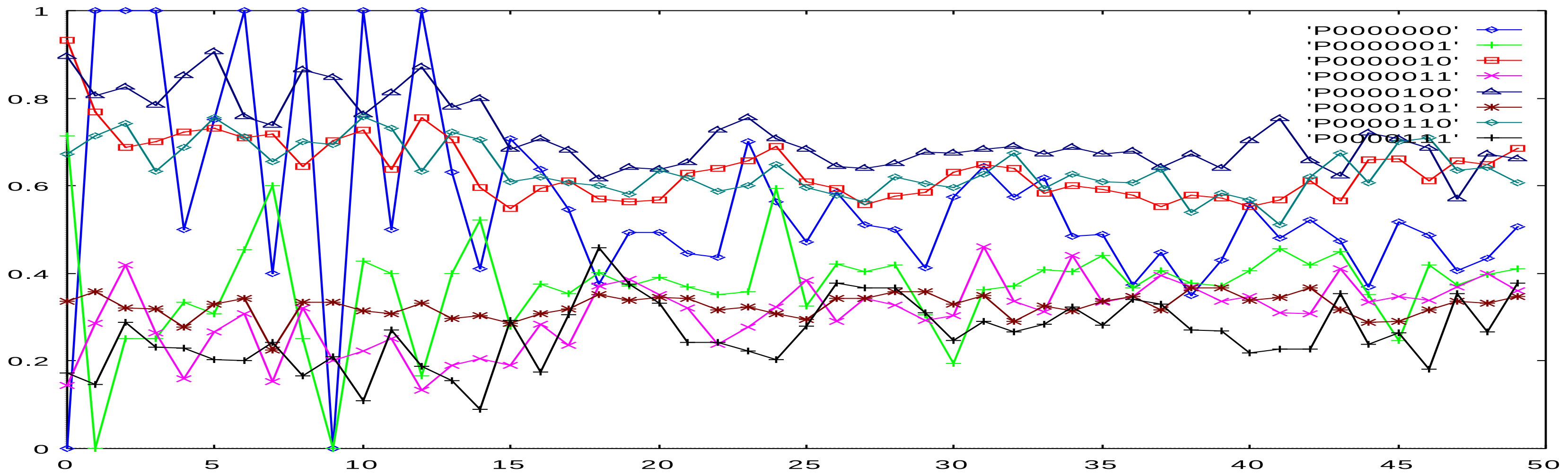


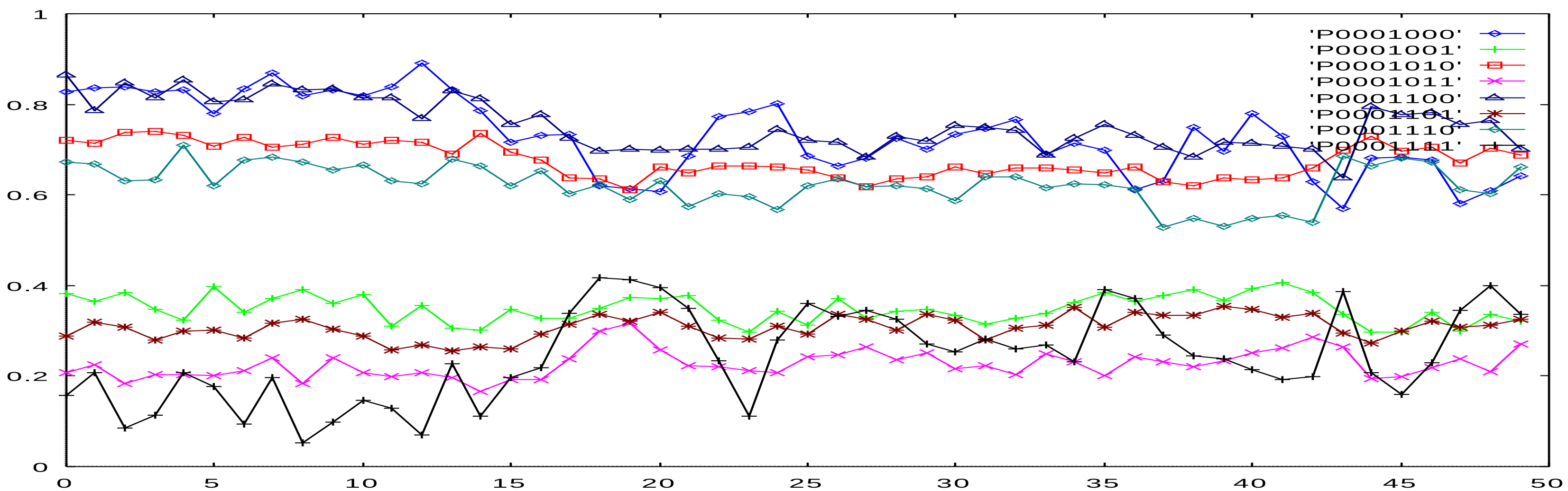
Depth=7　P(1|0000000), etc.





Depth=7 P(1|0000000), etc.

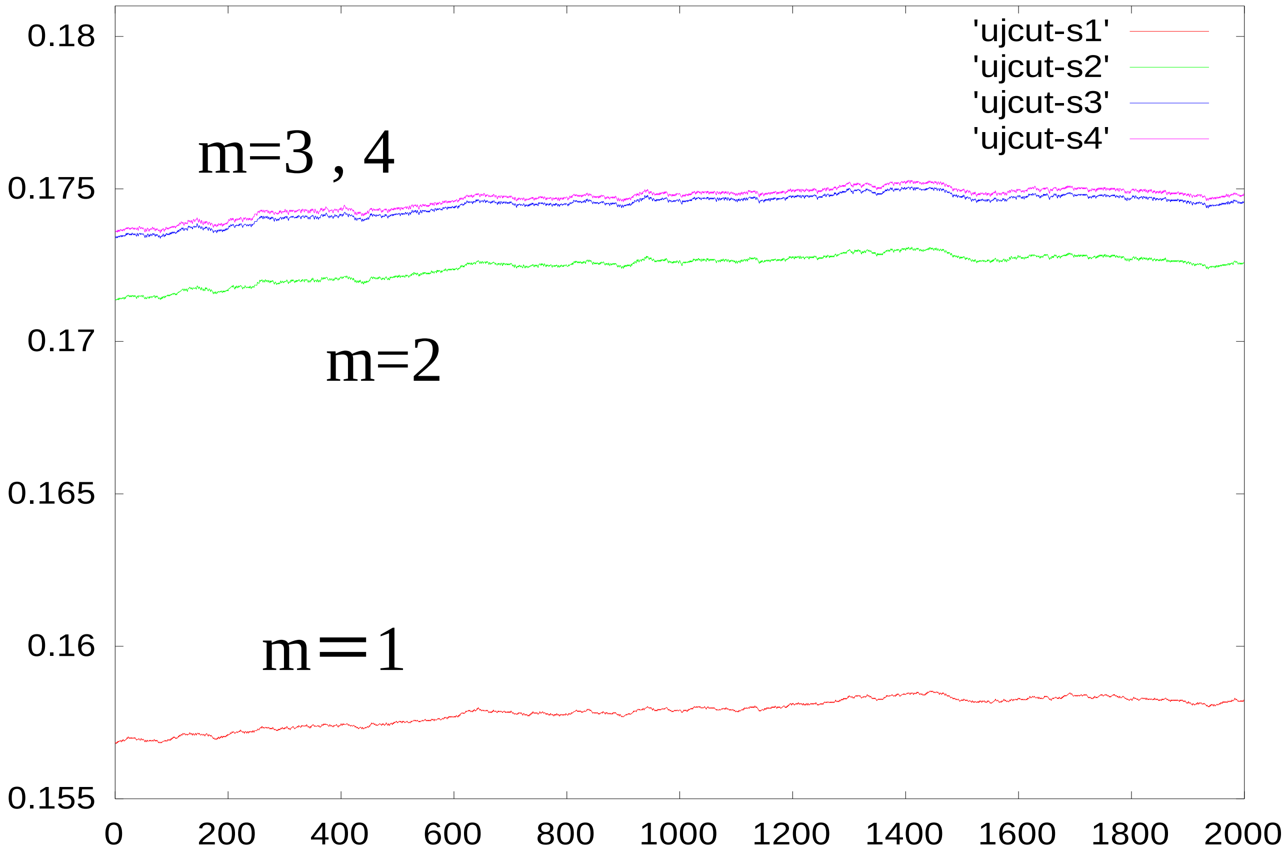




Memory depth from M(x:y)

$$M(x,y) = H(x) - H(x|y)$$

$$= -\sum_x P(x)\log_2 P(x) + \sum_x \sum_y P(x|y)\log_2 P(x|y)$$



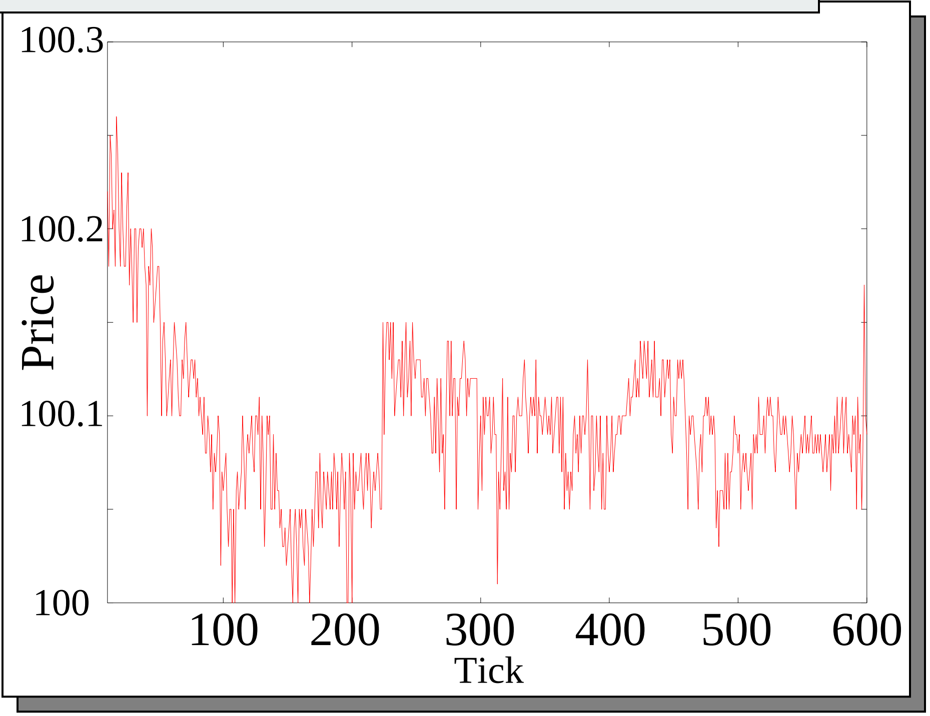
m=3 and 4 overlap

No more information expected from 4 steps before

進化的手法による
為替TICKデータの予測

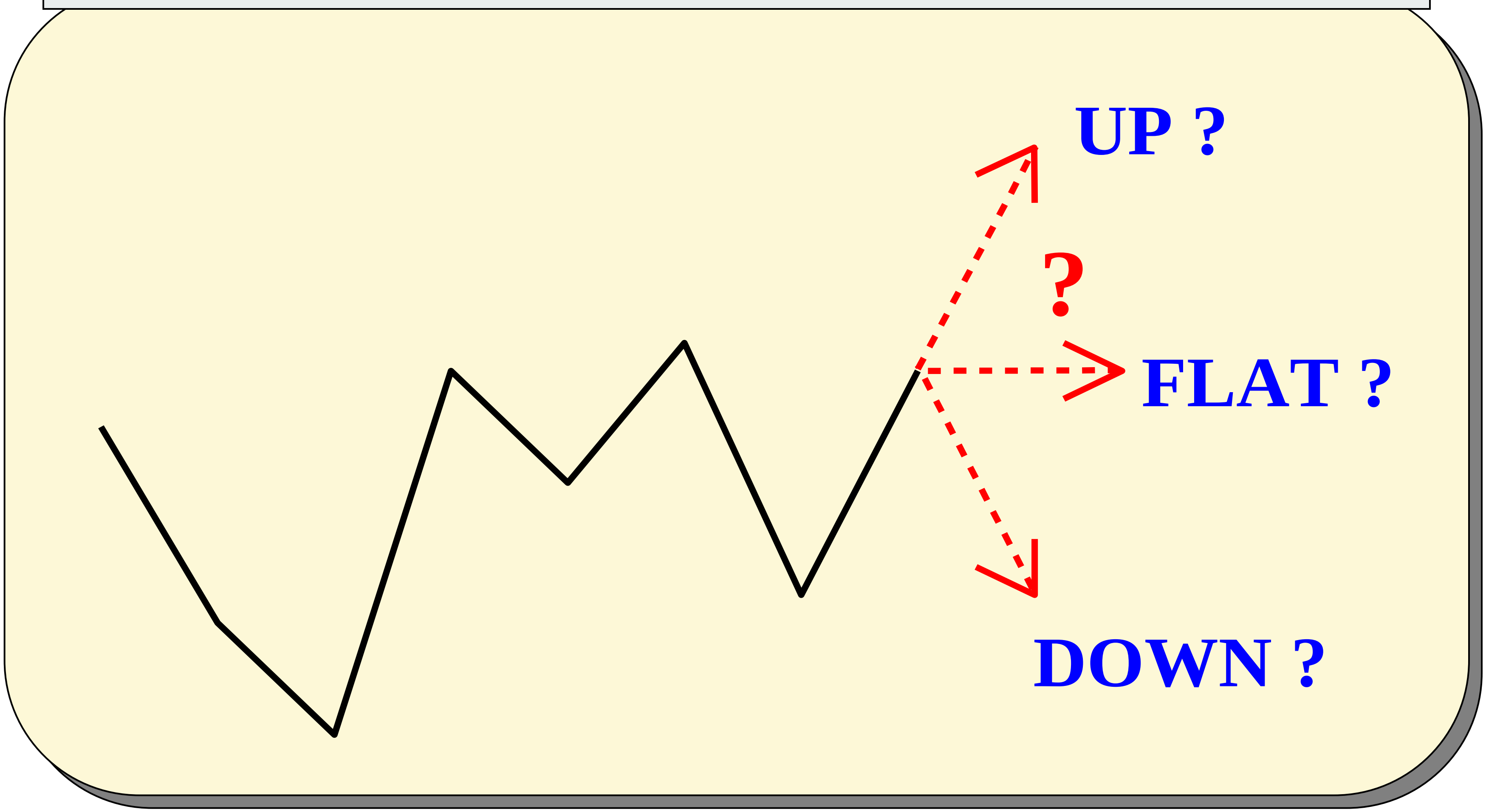
金融Tickデータとは

日付	時間	価格
1995/01/02	14:53	100.22
1995/01/02	14:53	100.18
1995/01/02	14:53	100.23
1995/01/02	14:55	100.20
1995/01/02	14:55	100.25
1995/01/02	14:55	100.16
1995/01/02	14:55	100.21
1995/01/02	14:56	100.20
:	:	:



<u>極短期価格変動の記録</u>

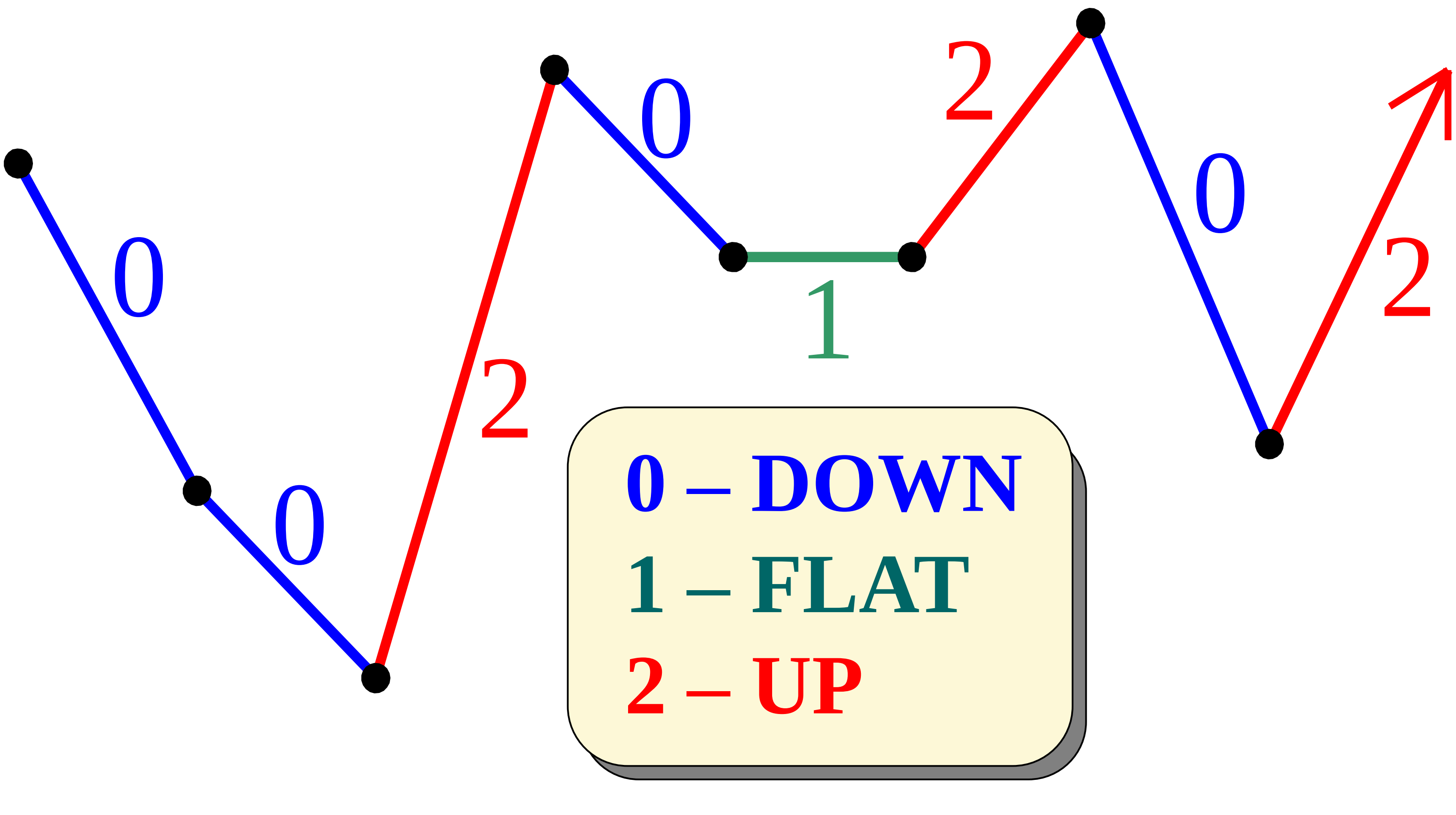
価格変動は予測可能か?



目的

- 進化的計算法を用いて，過去のパターンから予測を行う
- 過去の値動きがどの程度影響するのか？

価格変化の符号化



履歴（H）

・ 履歴の長さ = N

$$H = h_1\ h_2\ h_3 \cdots h_N$$

$$h_i \in \{0, 1, 2\}$$

0 – DOWN
1 – FLAT
2 – UP

予測ゲノム (P)

遺伝子長 = 3^N

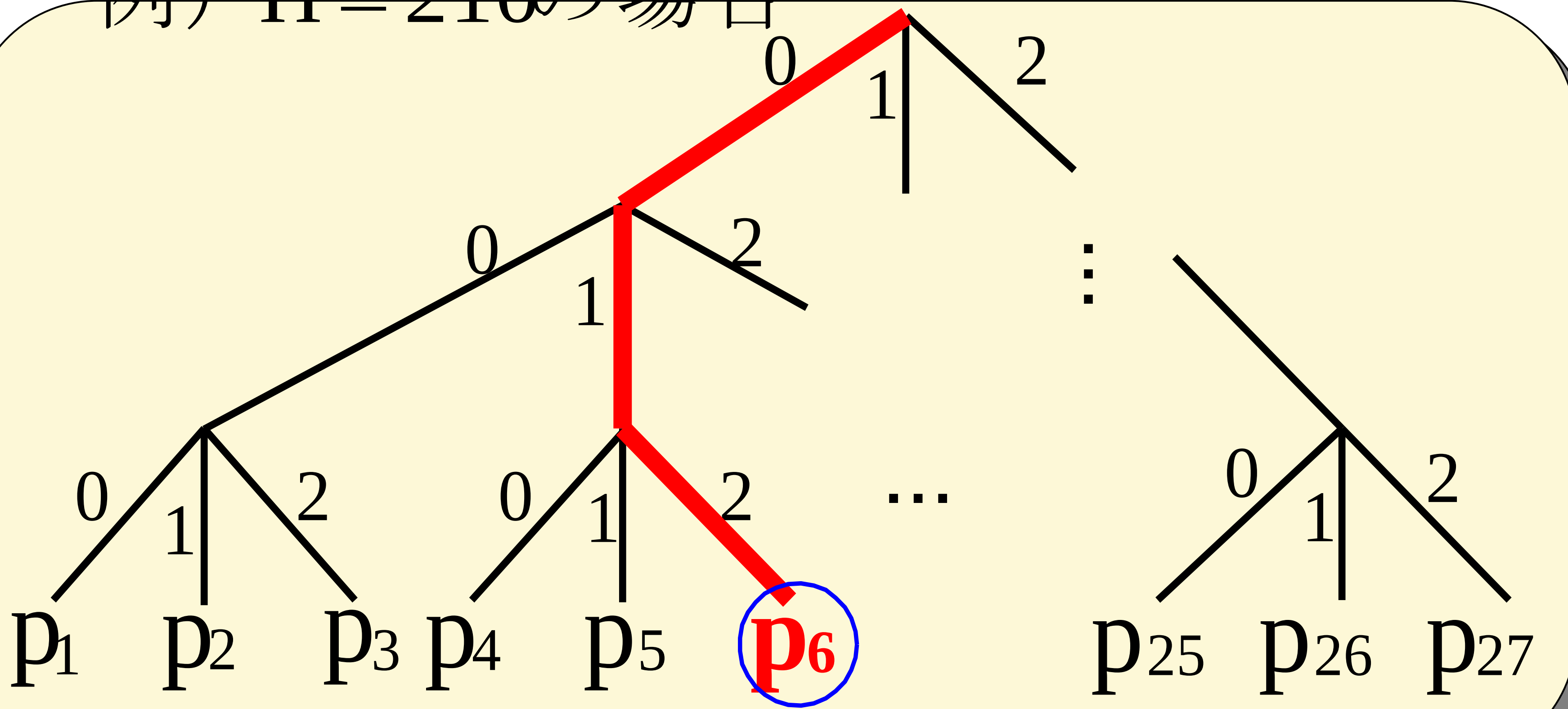
$$P = p_1\, p_2\, p_3 \cdots p_{3^N}$$

$$p_j \in \{0, 1, 2\}$$

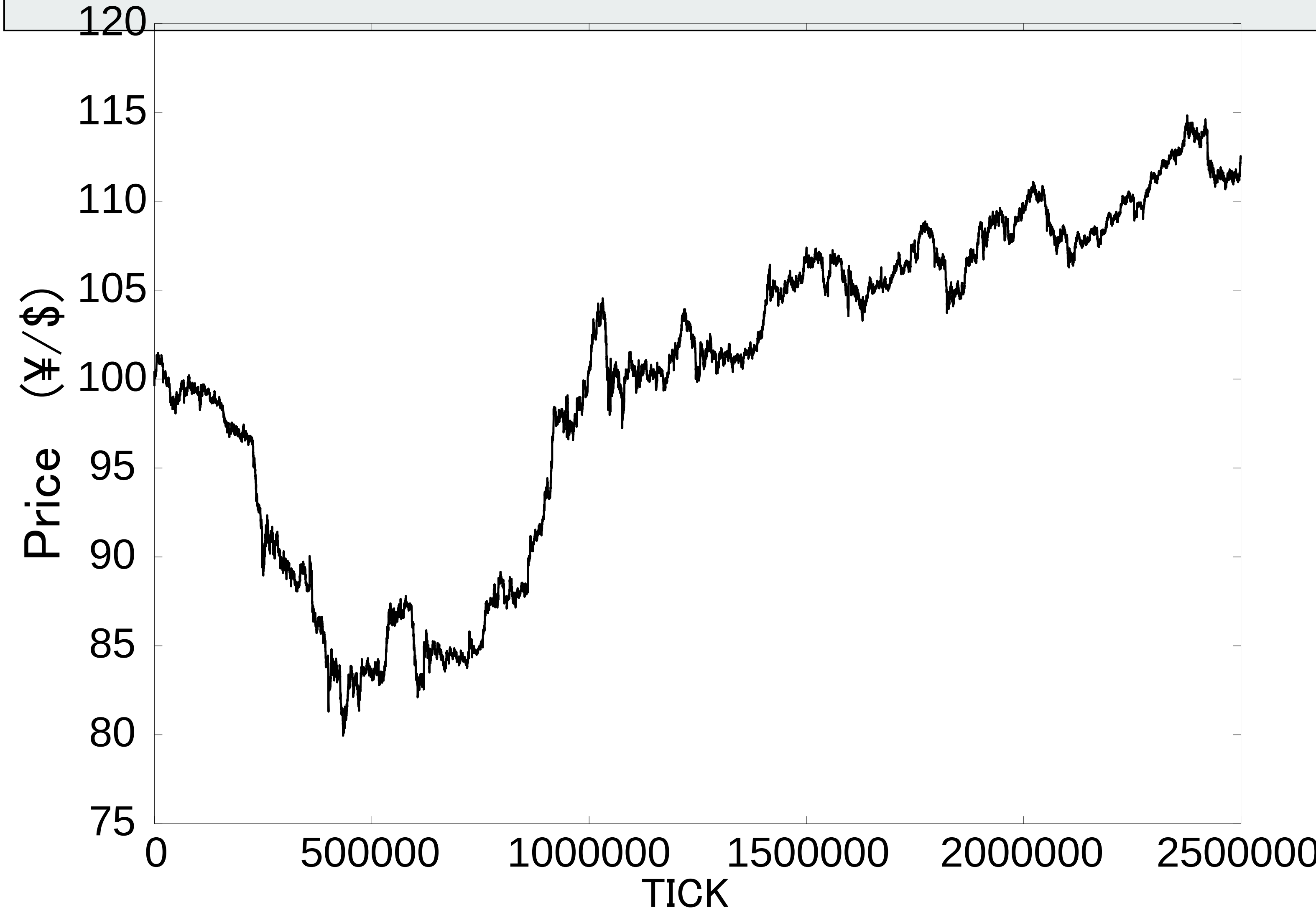
p_1	p_2	p_3	$\cdots$	p_{3^N-1}	p_{3^N}

予測の決定方法

例）H＝210の場合

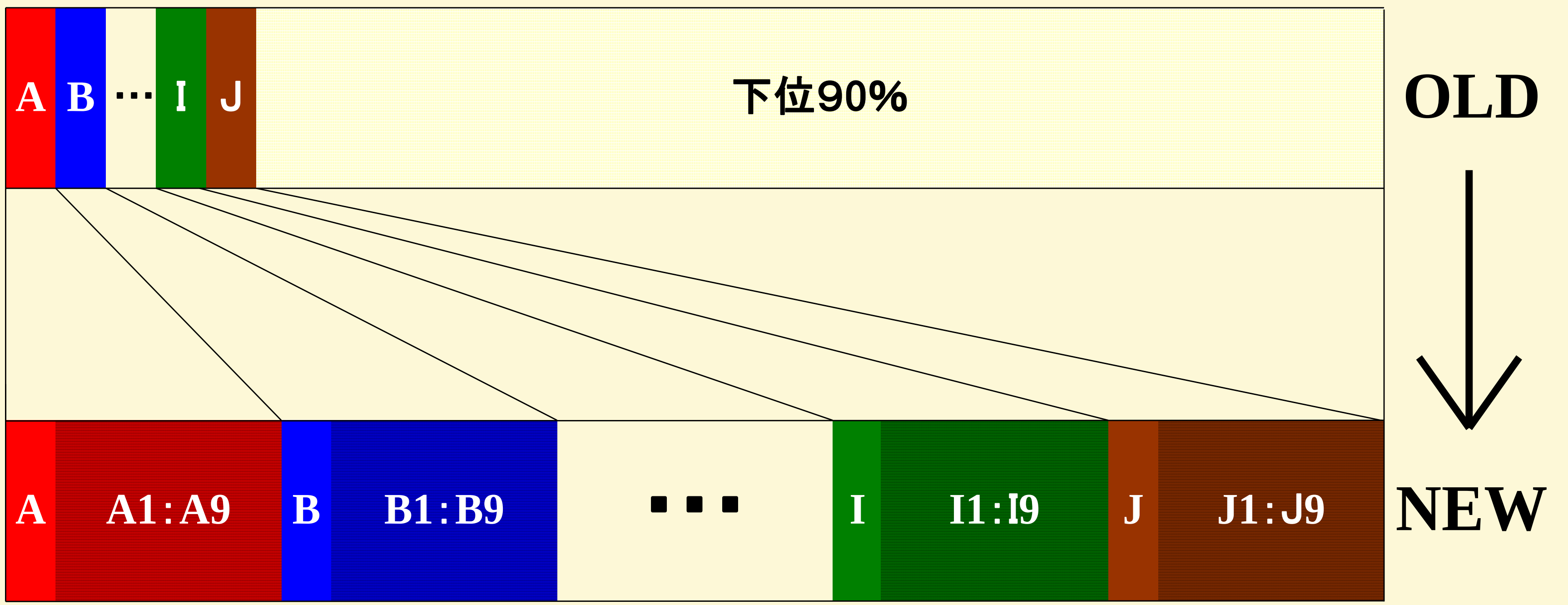


実験データ



120
115
110
105
100
95
90
85
80
75
Price (¥/$)
0
500000
1000000
1500000
2000000
2500000
TICK

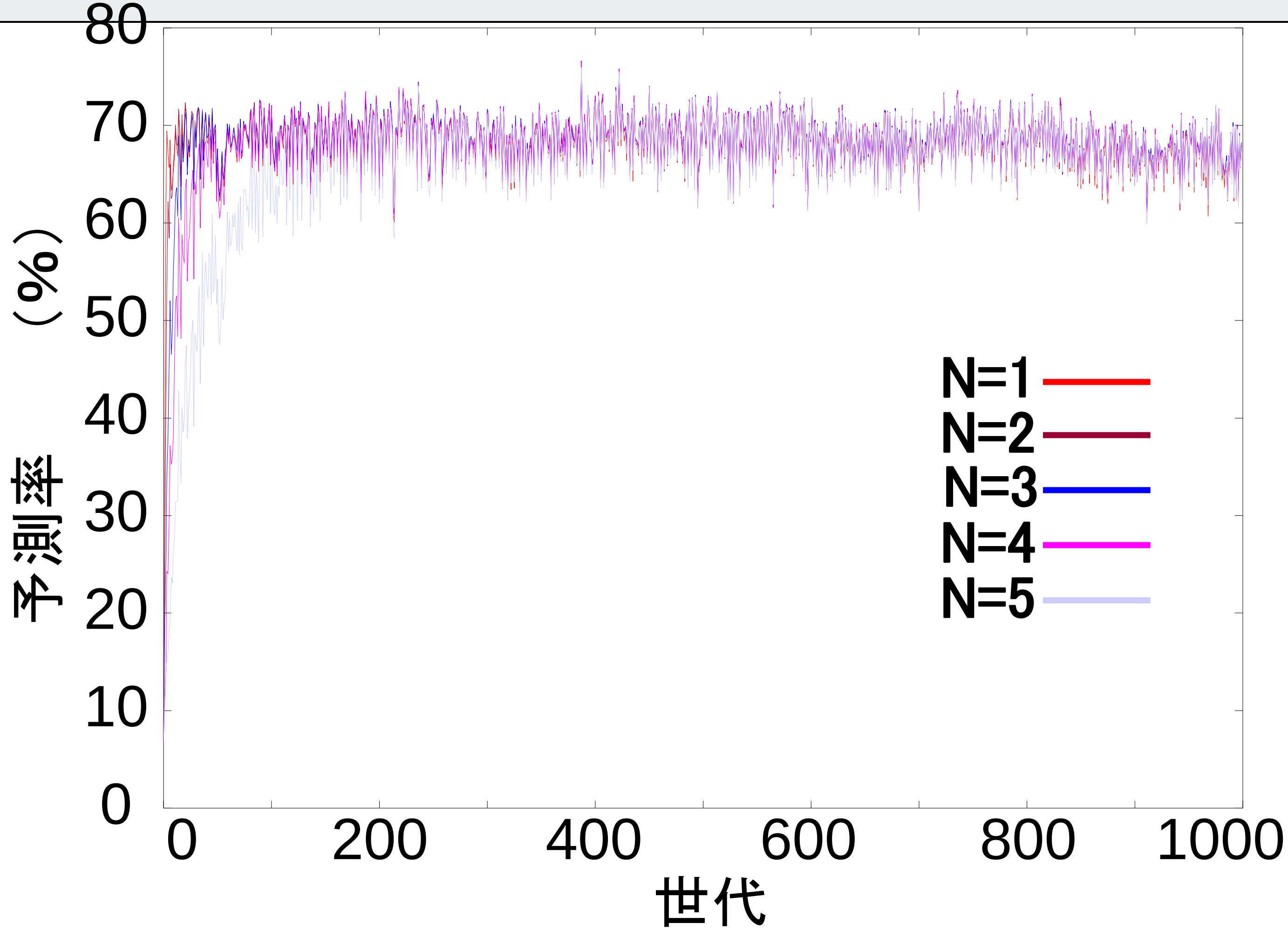
世代交代



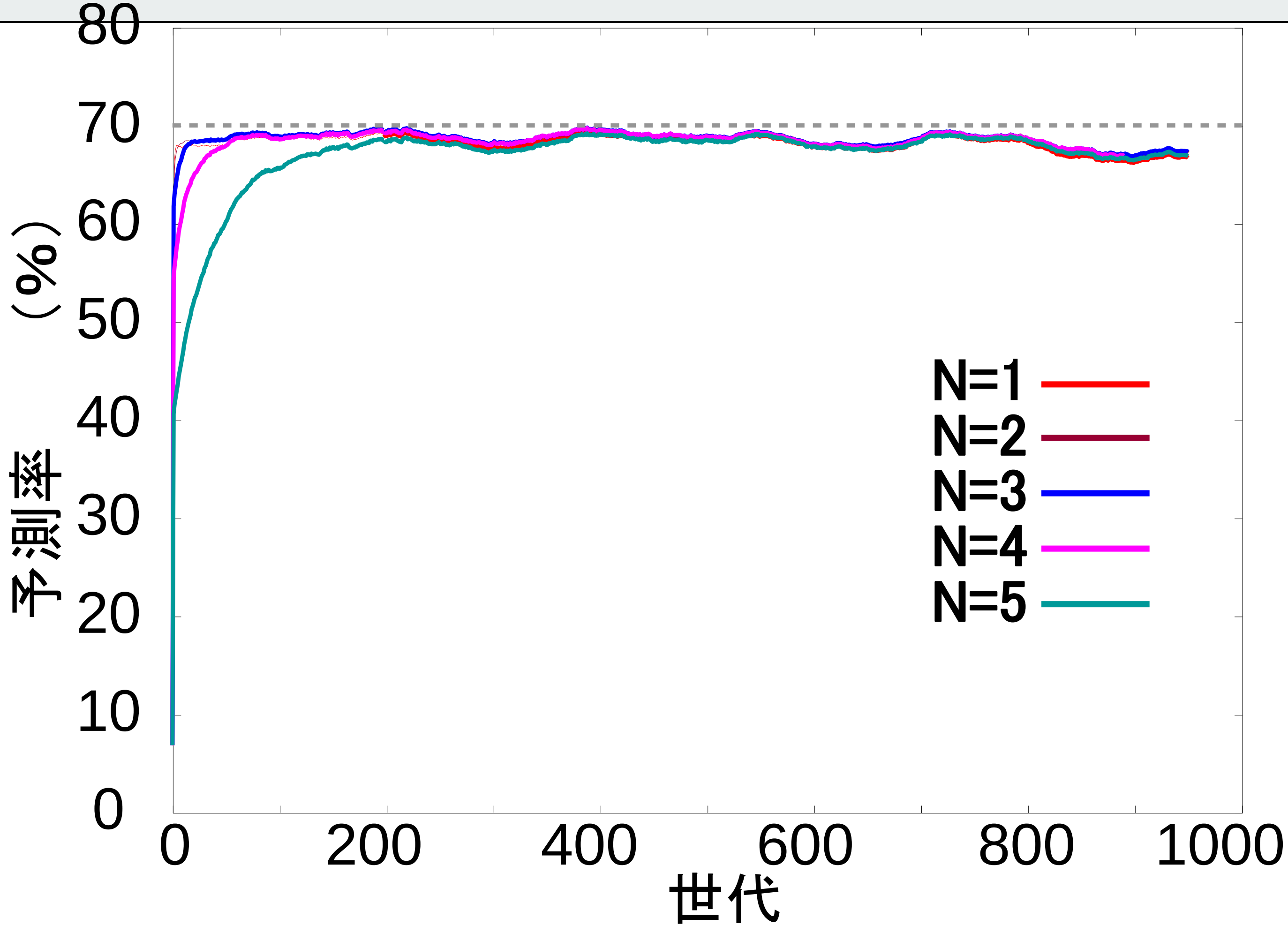
シミュレーション

・ エージェント数	100
・ 予測ゲノム初期値	全て不変と予測
・ 1世代の予測期間	2,500 ticks
・ 最大世代数	1,000 世代

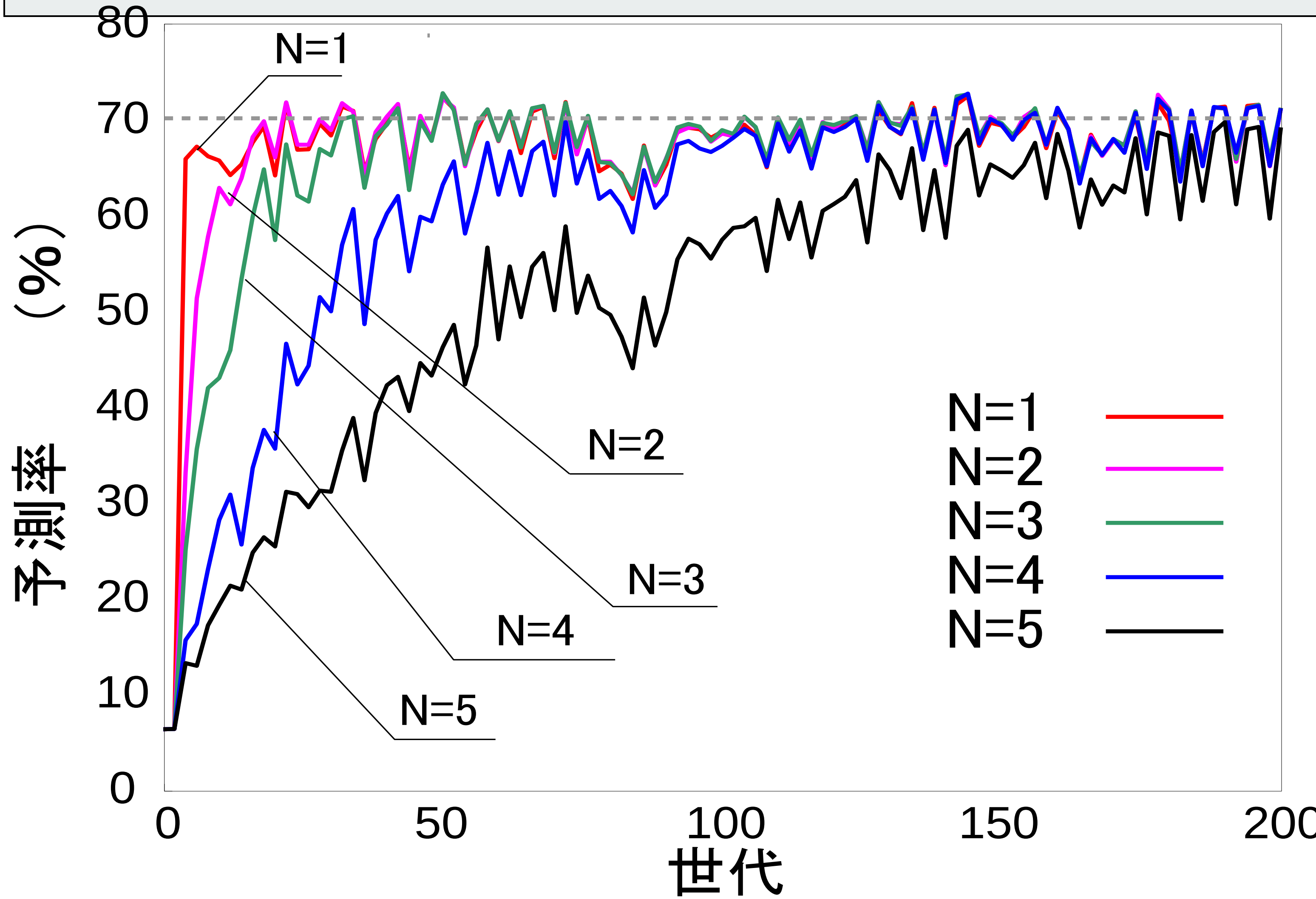
実験結果



40世代ごとの平均予測率

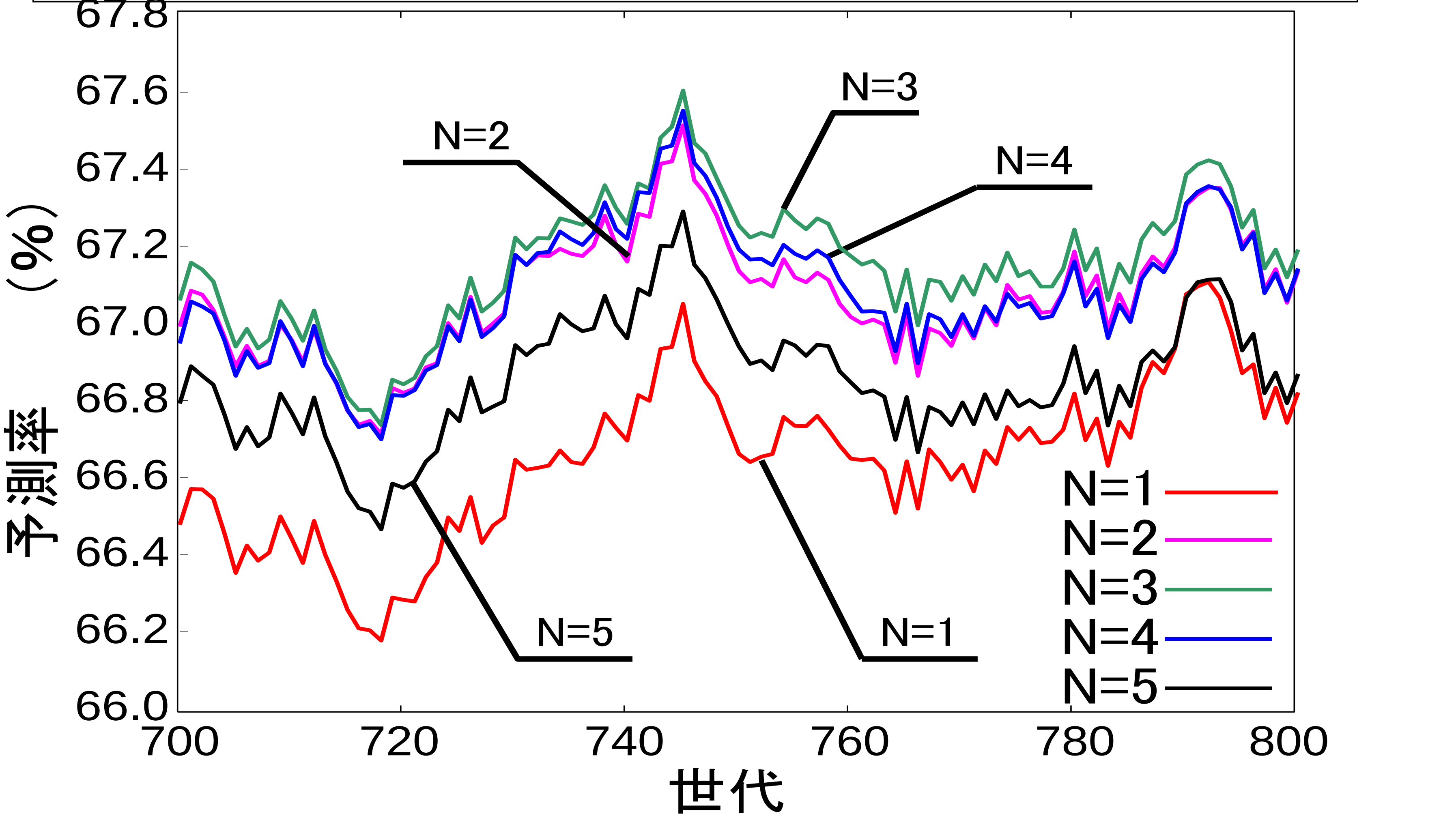


200世代までの様子



予測率 （%）
80
70
60
50
40
30
20
10
0
0
50
100
150
200
世代
N=1
N=2
N=3
N=4
N=5
N=1
N=2
N=3
N=4
N=5

700から800世代までの様子



ここまでのまとめ

- 予測率は67～ 68%
- 履歴長Nが3以上で頭打ち

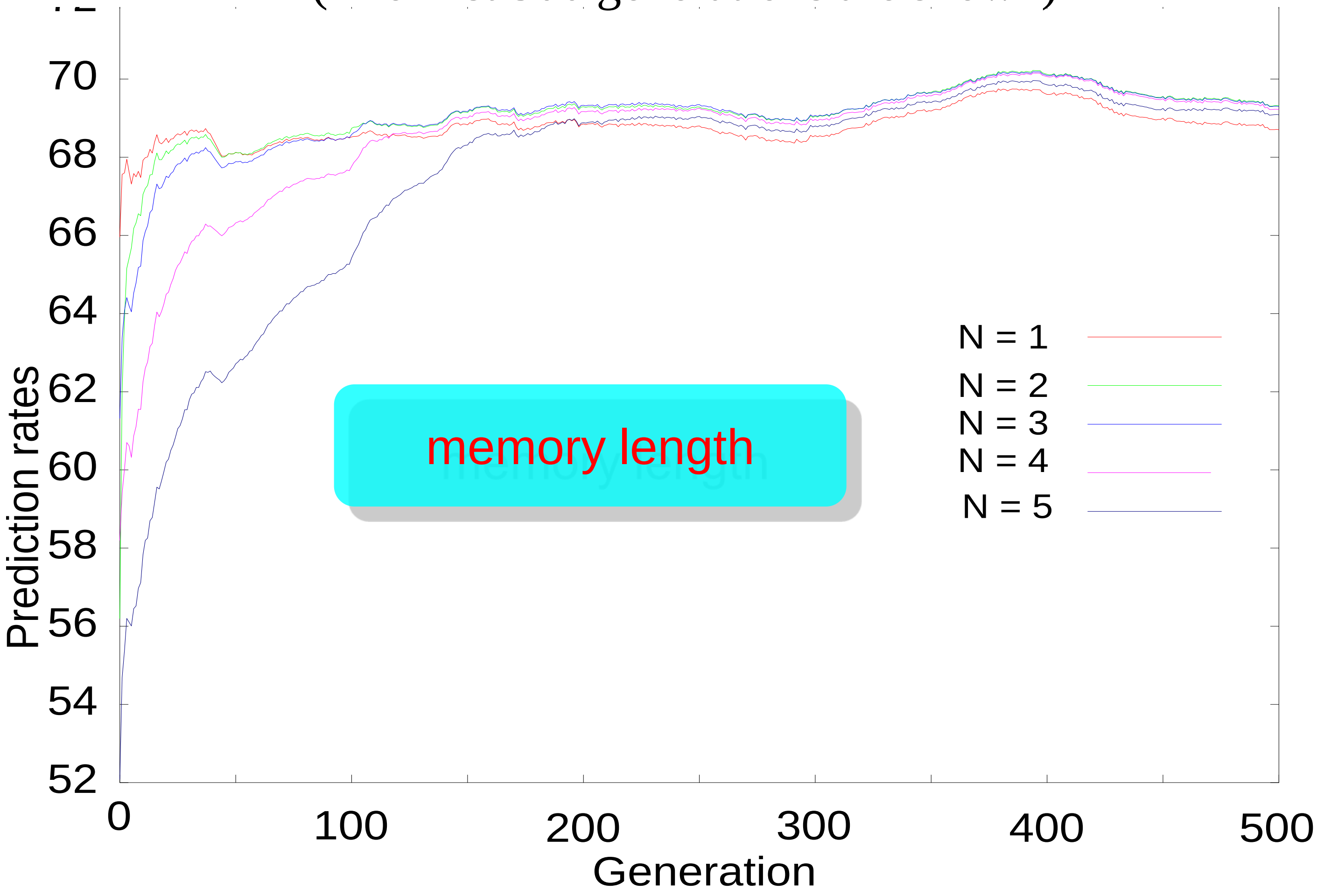
さらに予測率を上げるには？

- もう少し広い判断材料が必要

- 新たな性質が見つかった！　改良できそう

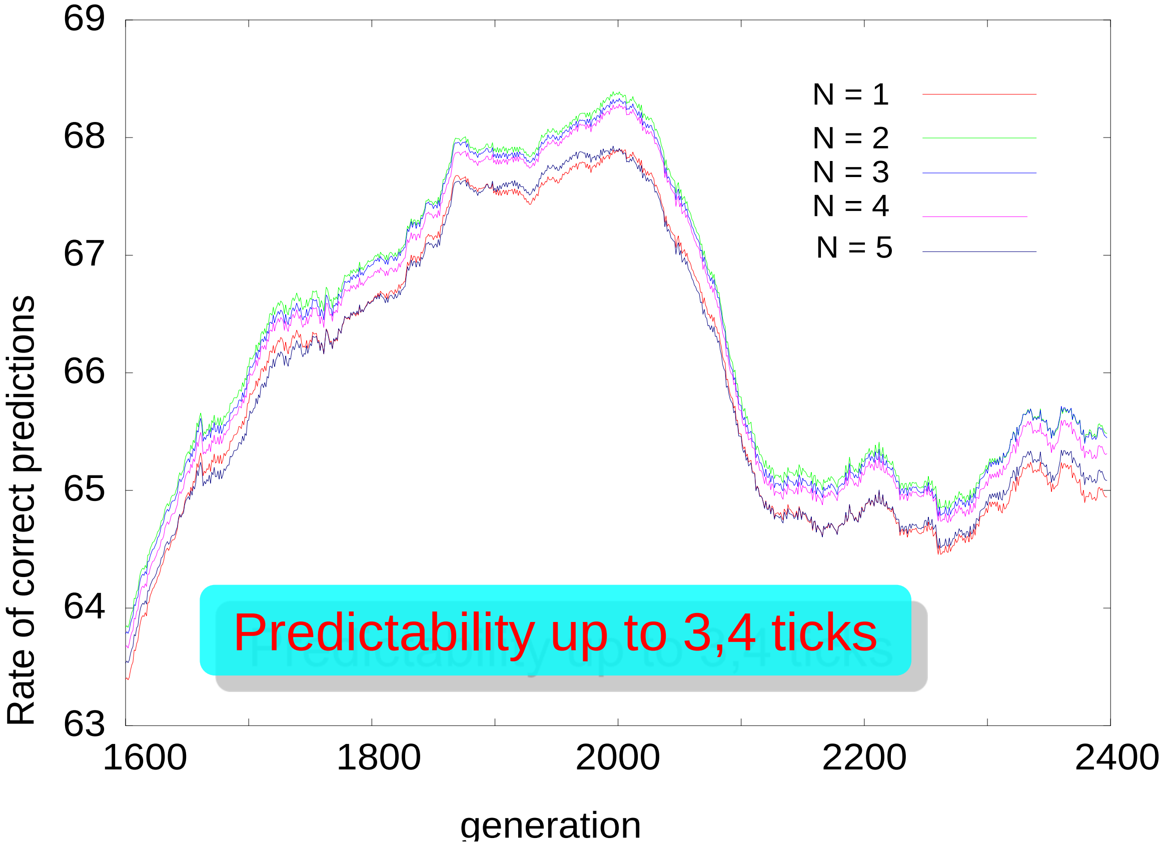
Result: Predictability based on GA leaning

(The first 500 generations are shown)



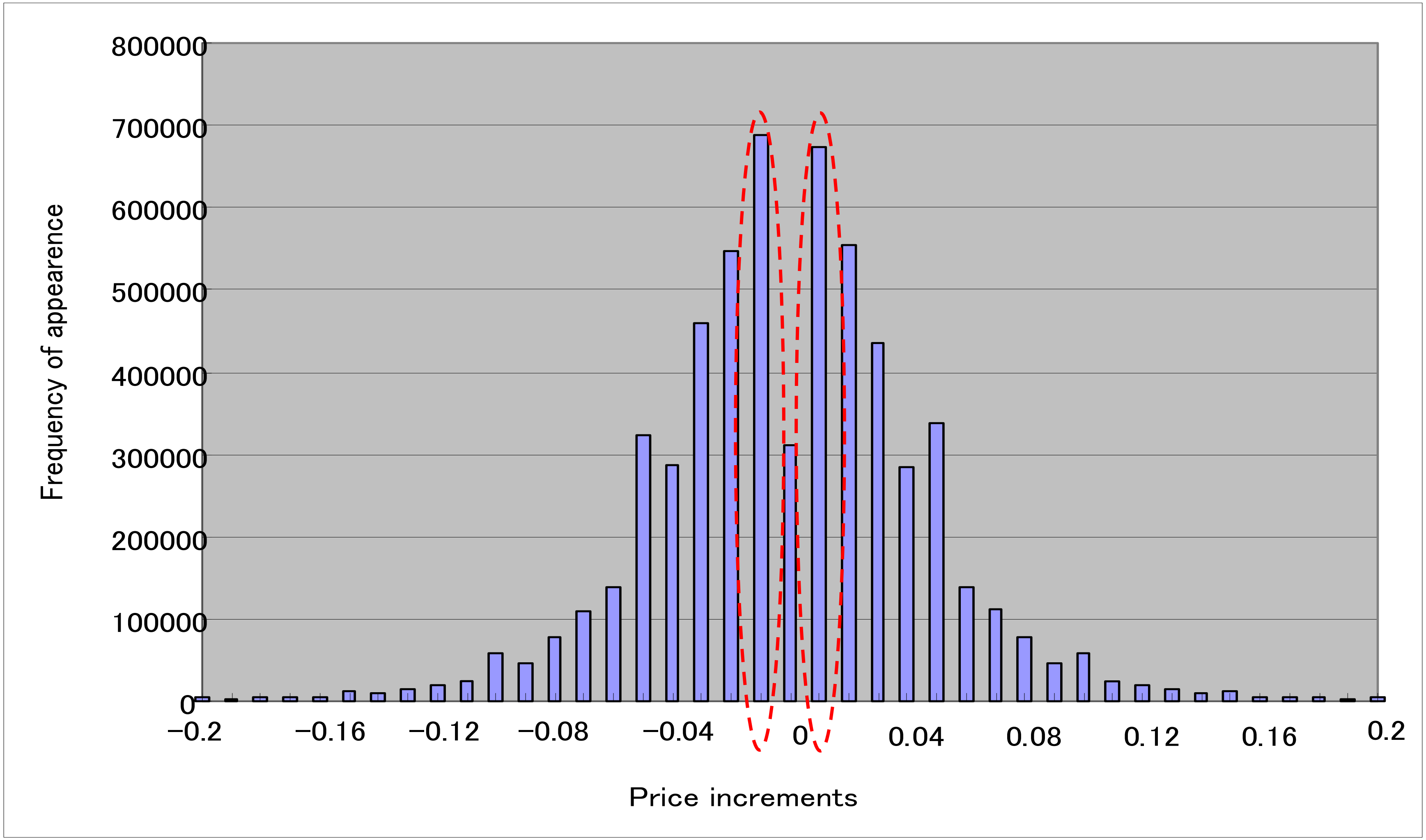
Predictability after sufficient learning period (1600-2400)

From the top to bottom N=2,3,4,5,1



二つの条件付き確率
P(UP|m) と P(DOWN|m)
の奇妙な振る舞い

Histogram of price increments

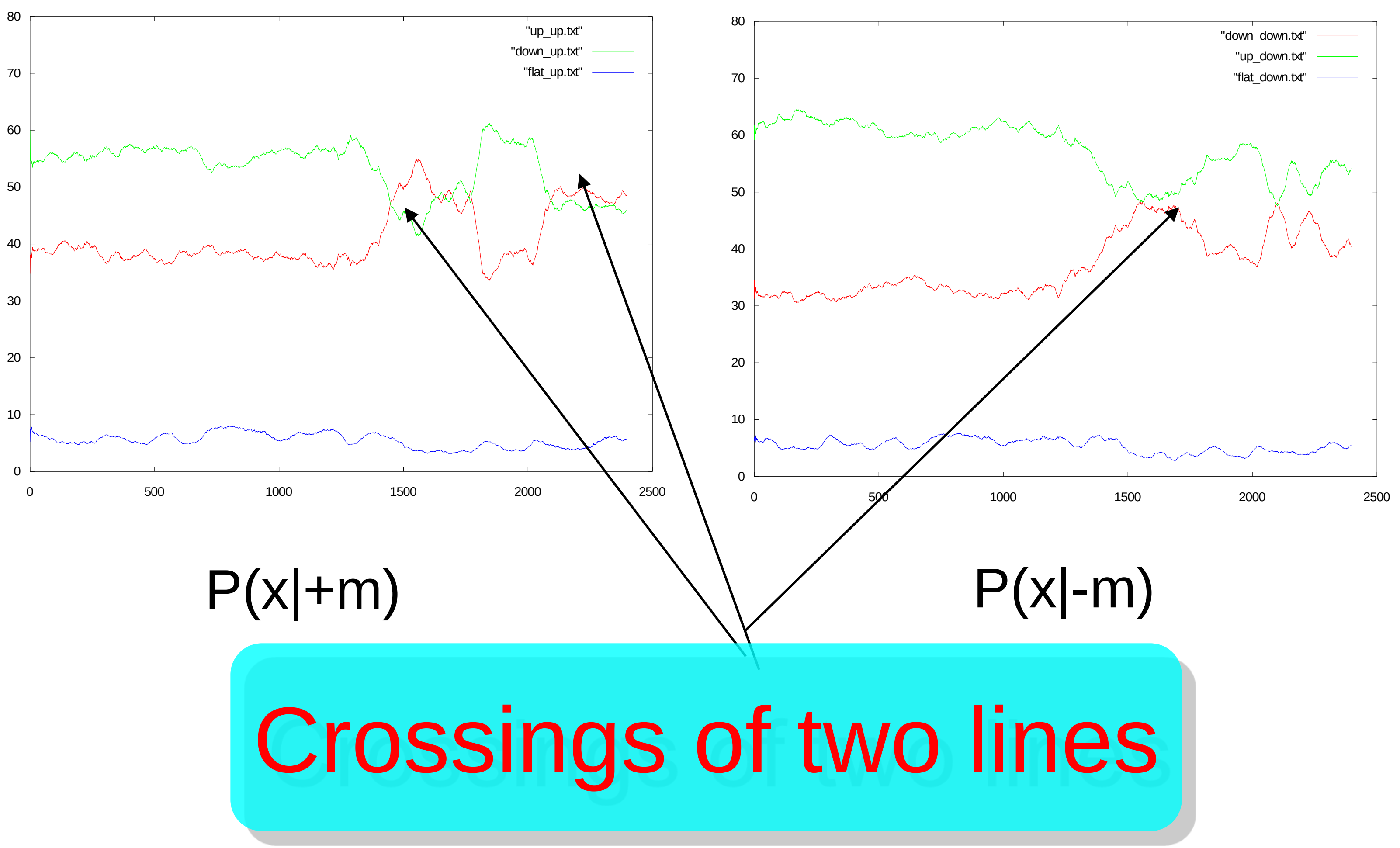


Conditional probability after ±0.01

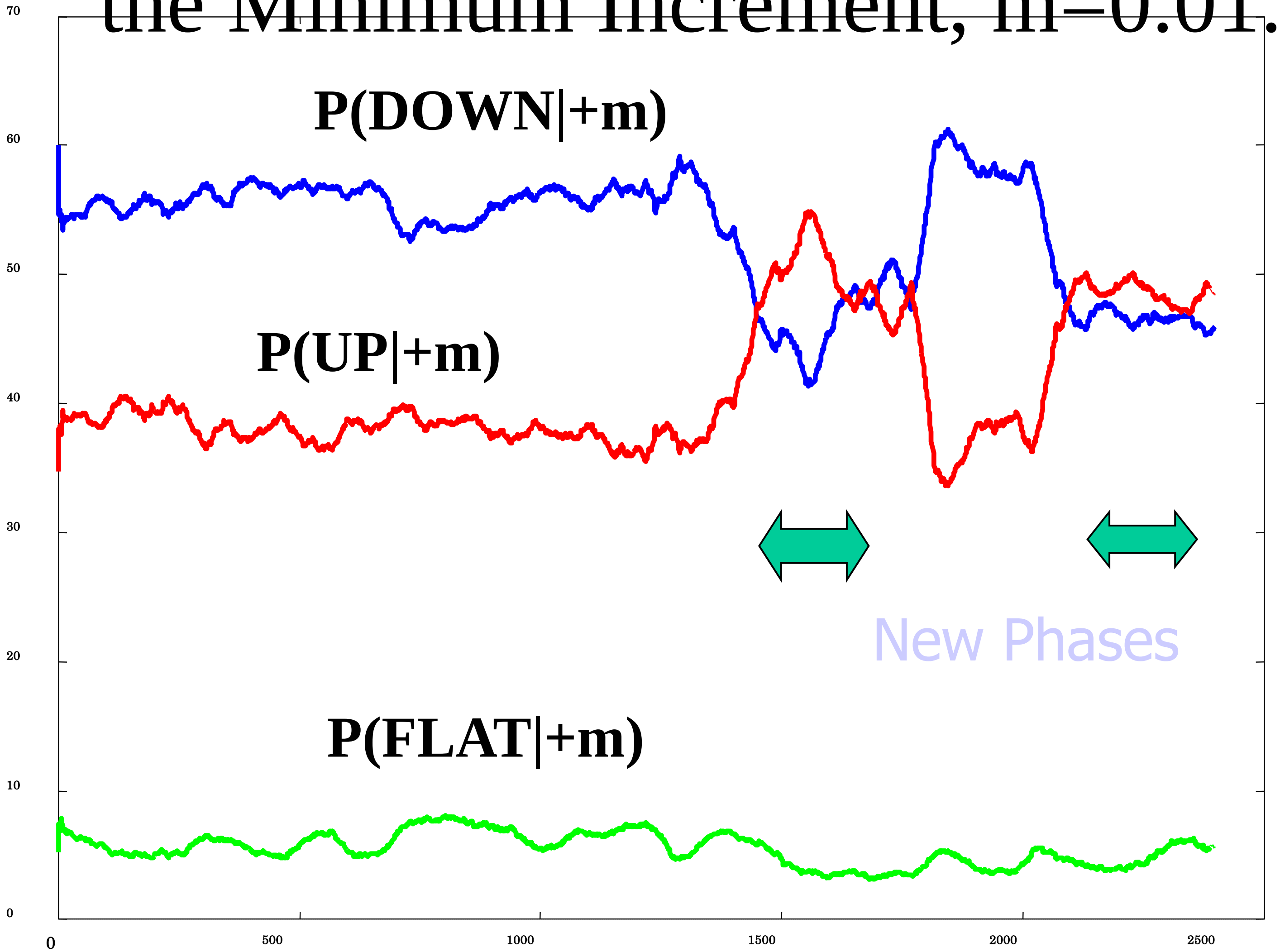
±0.01 occurs most frequently

After the move of ±m

One point represents condtional probability over 2500ticks



Conditional Probabilities After a Rise by the Minimum Increment, m=0.01.



More states

down: 0
flat : 1
up : 2

Down more than 0.01 : 0
Down by 0.01 : 1
flat : 2
Up by 0.01 : 3
up more than 0.01 : 4

$$H = h_1 h_2 h_3 \cdots h_N$$

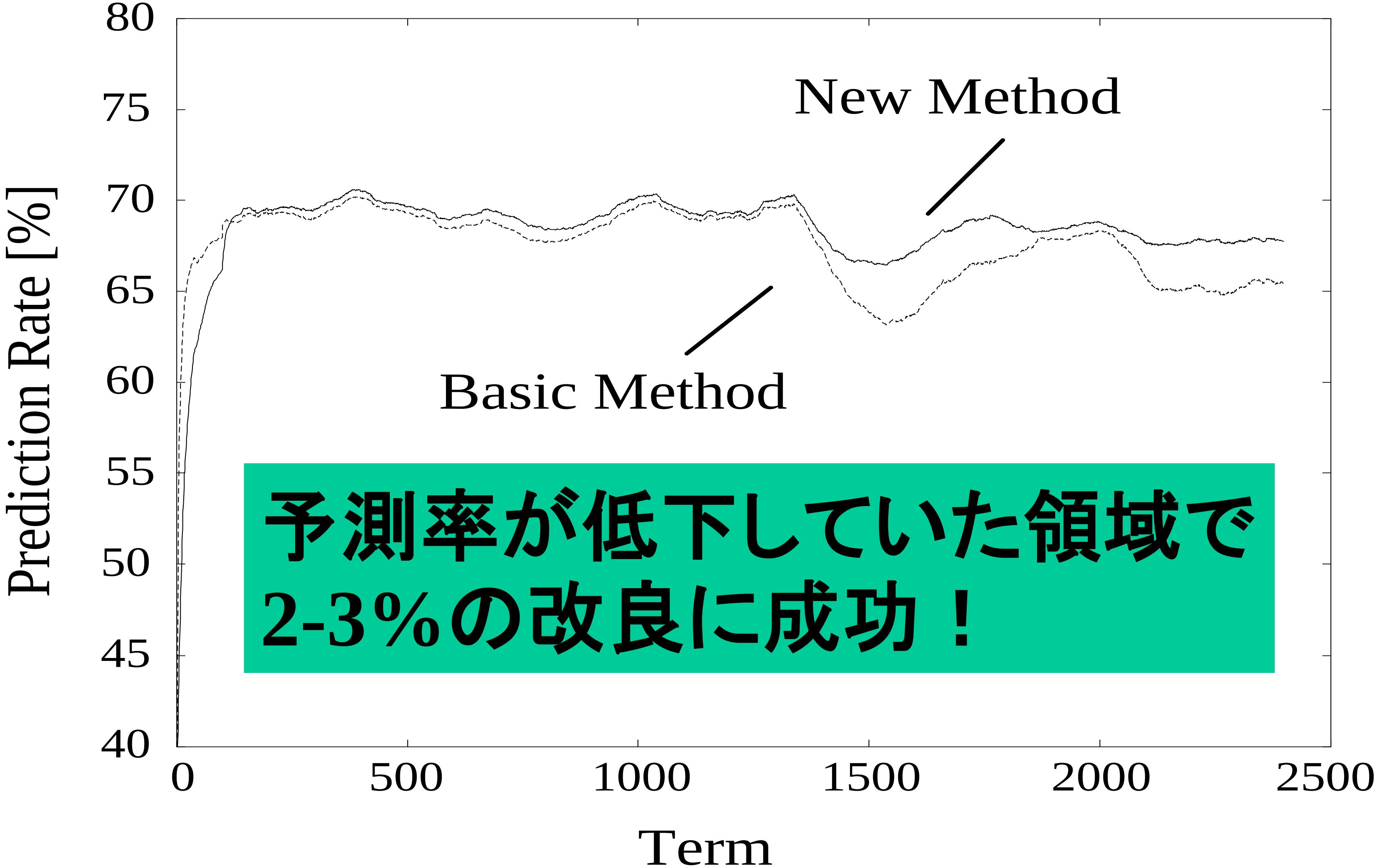
$$h_i \in \{0,1,2,3,4\}$$

History types $= 5^N$

New Method with Size Effect and Discovery of Two Phases

- Not only the direction of move but also the size of price moves must be taken into account in this new phase.

Comparison：Basic Method (dashed) vs. New Method (solid) for N=3.



予測率が低下していた領域で2-3%の改良に成功！

テクニカル指標自動選択の試み

Labeling Different Phases by Means of Technical Indicators

Constructing a Prediction Generator

Technical Indicators

- **Trend Indicators**
  - **MA (Moving Average)**
  - **MAD (Deviation MA)**
  - **EMA (Exponential MA)**
  - **MO (Momentum)**
  - **MACD**

    (Moving Average Convergence and Divergence)

  ➡ **price direction**

- **Oscillator Indicators**
  - **RSI (Relative Strength Index)**
  - **RCI (Rank Correlation Index)**
  - **PL (Psychological Line)**
  - **Etc…..**

  ➡ **overbought and oversold signals**

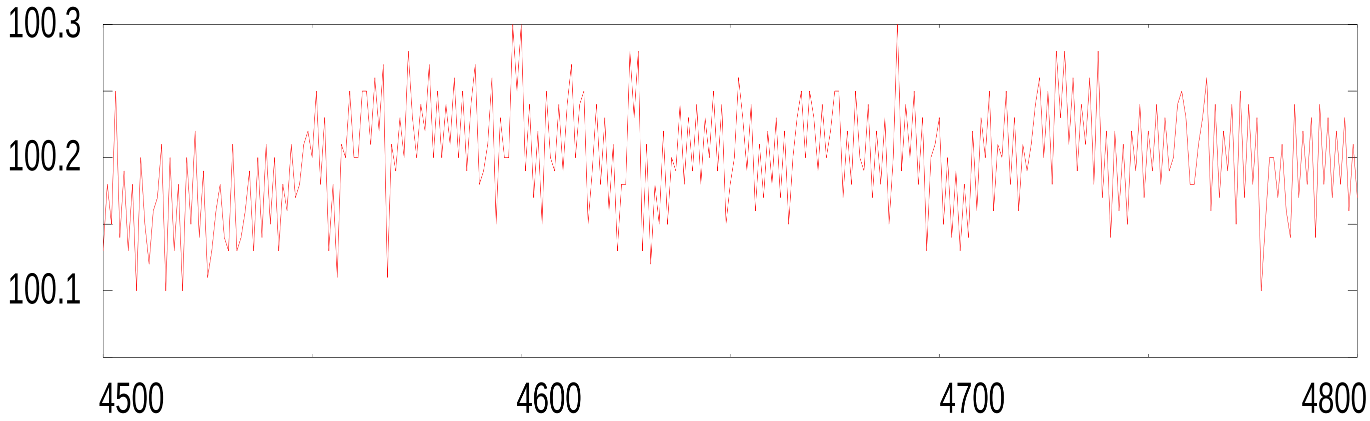
Data: Tick-wise Stock Price

NYSE, 8 symbols, 1993/1/1～1993/12/31

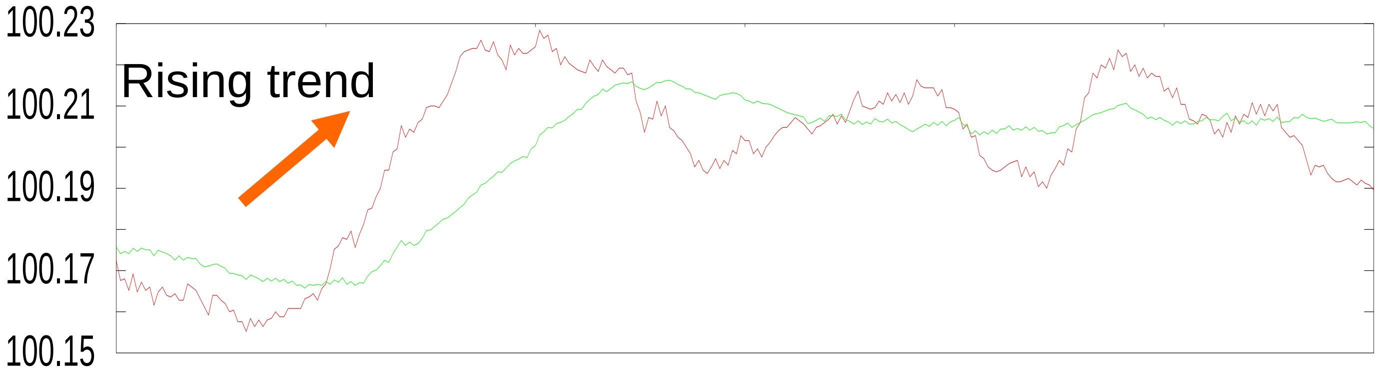
Stock symbol	Business type	Data size (ticks)	Tick interval (s/ticks)	10-ticks (minute)
BBY	retail	54821	109	18
SMRT	retail	12525	473	78
APC	oil	23685	253	42
BP	oil	73562	83	14
CA	computer	65051	92	15
IBM	computer	455233	14	2
F	automobile	194561	32	5
GM	automobile	277241	23	4

Example of trend indicator : MA

Tick price



MA
(Trend Indicators)



Genetic operations & parameters

Genetic operations

- Selection ： Elite selection(10%) +
  Roulette-wheel selection (90%)
- Crossover ： Two-point crossover

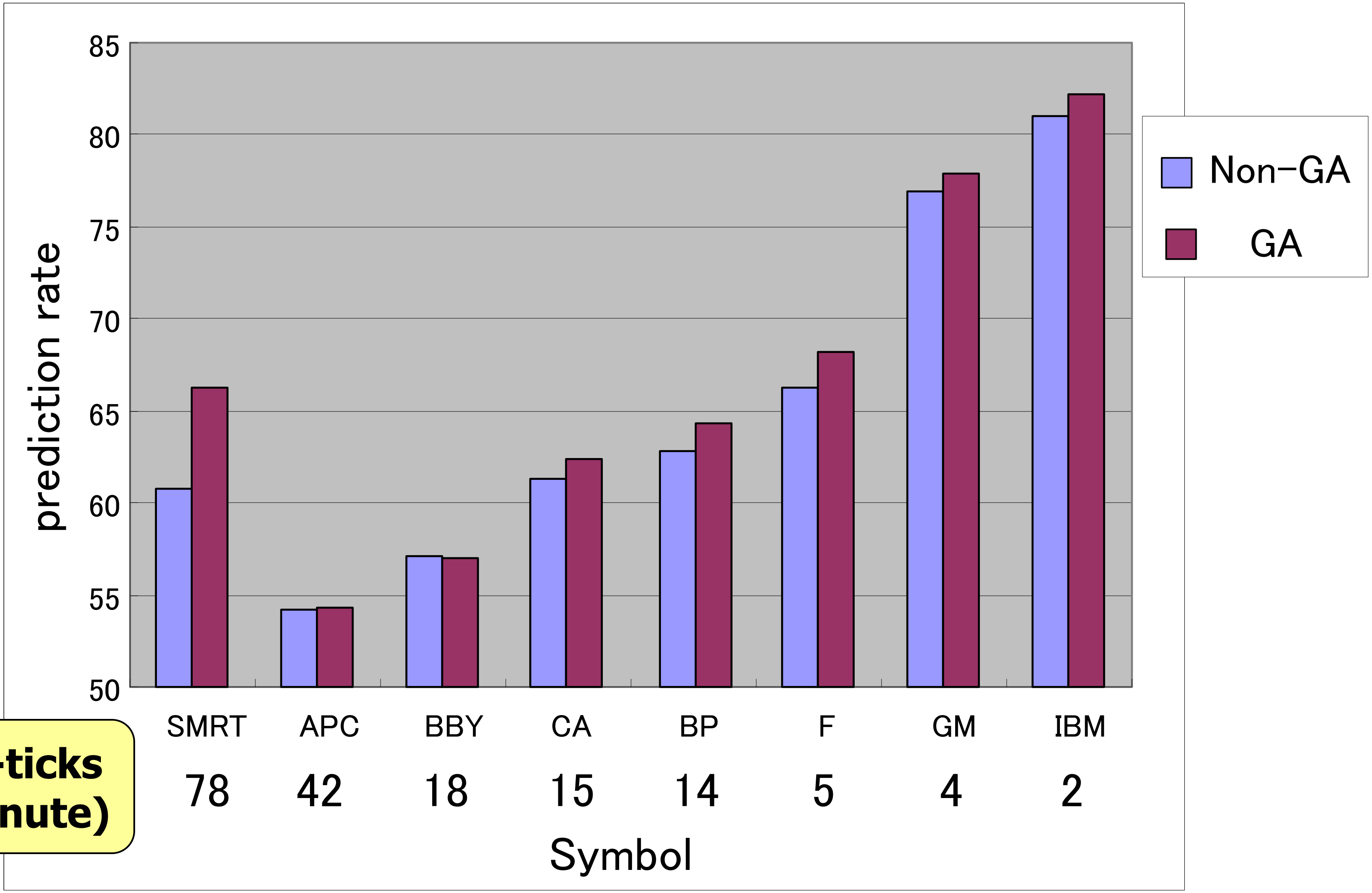
- Parameters
  - Population：100
  - Generation：500
  - Crossover rate：90%
  - Mutation rate：1%

Combination of indicators selected by GA

	APC	BBY	BP	CA
1st	(MO1,MA1,MACD)	(MO2,RSI)	(MO2,MA2)	(MO1,MA2)
2nd	(MO1,MA1,MACD,RCI)	(MO1,RSI)	(MO2,EMA)	(MO1,MA2,MAD)
3rd	(MO1,MA1,RSI)	(RSI,RCI)	(MO2,MA1)	(MO1,MA1)

	F	GM	IBM	SMRT
1st	(MO2,MA2,EMA)	(MO2,MAD,EMA)	(MO2,MAD,EMA)	(MO2,MACD)
2nd	(MO2,MA2)	(MO2,MAD,PHL)	(MO2,MAD)	(MA2,MACD)
3rd	(MO2,MA1,EMA)	(MO2,MAD)	(MO2,EMA)	(MA2,EMA)

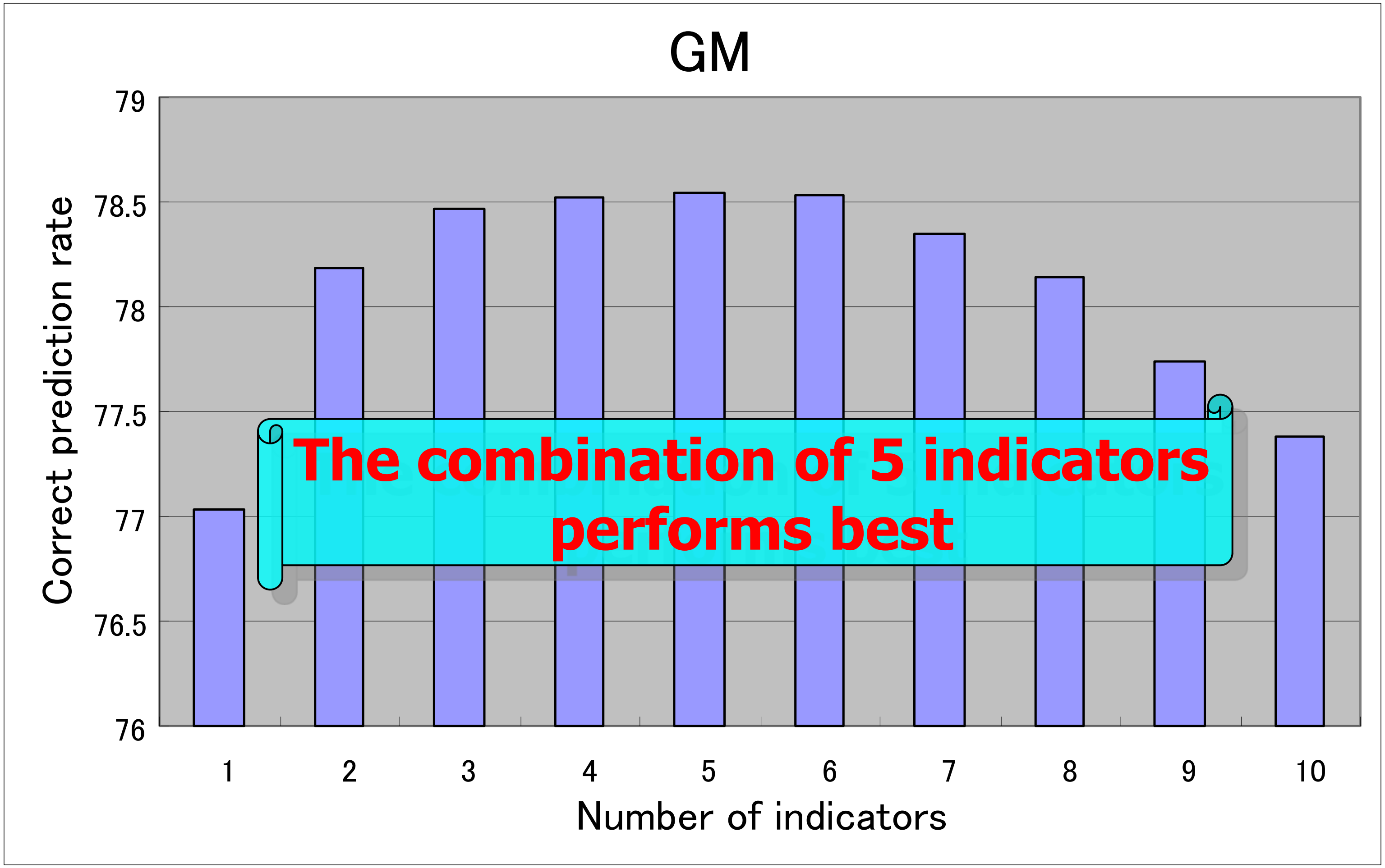
GA vs. Non-GA



10-ticks (minute)

How many indicators should be combined ?

Prediction rates vs. number of indicators



<u>The prediction rate improves when two or more indicators are used</u>

◎ : very effective
○ : effective
△ : effective only to a part of symbol
× : Non-effective

Evaluation

	Indicator	Single	Multiple
Trend type	MA1,MA2	◎	◎
	MAD	◎	◎
	EMA	◎	◎
	MACD	×	○
	MO1,MO2	○	◎
Oscillator type	RSI	△	○
	RCI	△	○
	PHL	○	○

Trend type × **Oscillator type** is effective

Technical indicators seem to work

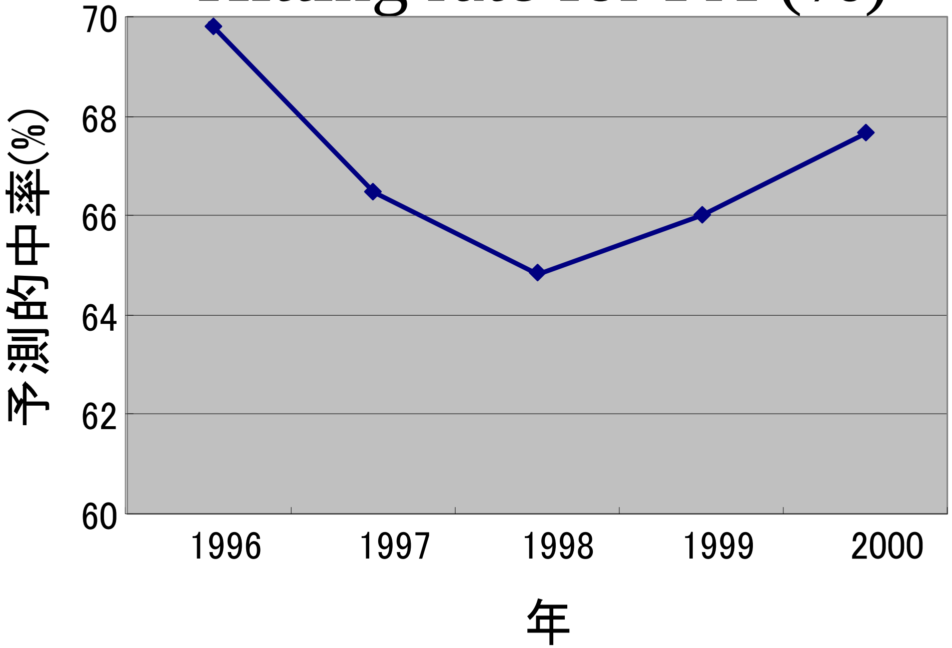
- **Predict price level 10-ticks ahead**
- **Combination of several technical indicators by GA**

↓

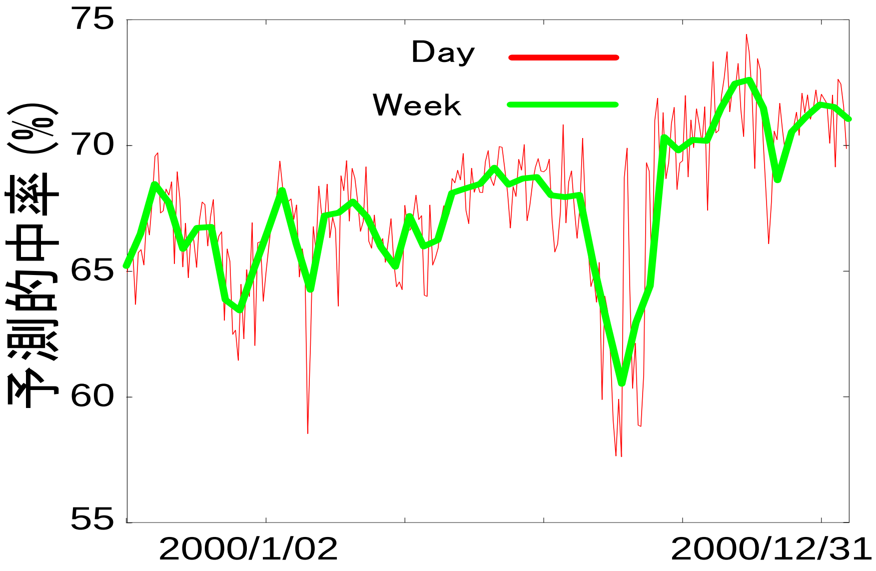
- **Correct prediction in 80% for busy stocks (IBM, GM)**
- **65% for average of eight stocks**
- **Best result given by combination of 3-5 indicators**
- **Poor performance on slow stocks can be explained by the large tick intervals (approx.1-hour)**
- **Our method should perform well for other popular stocks or foreign exchange rates**

Result: Forecast of 10 ticks ahead

- Hitting rate for FX (%)



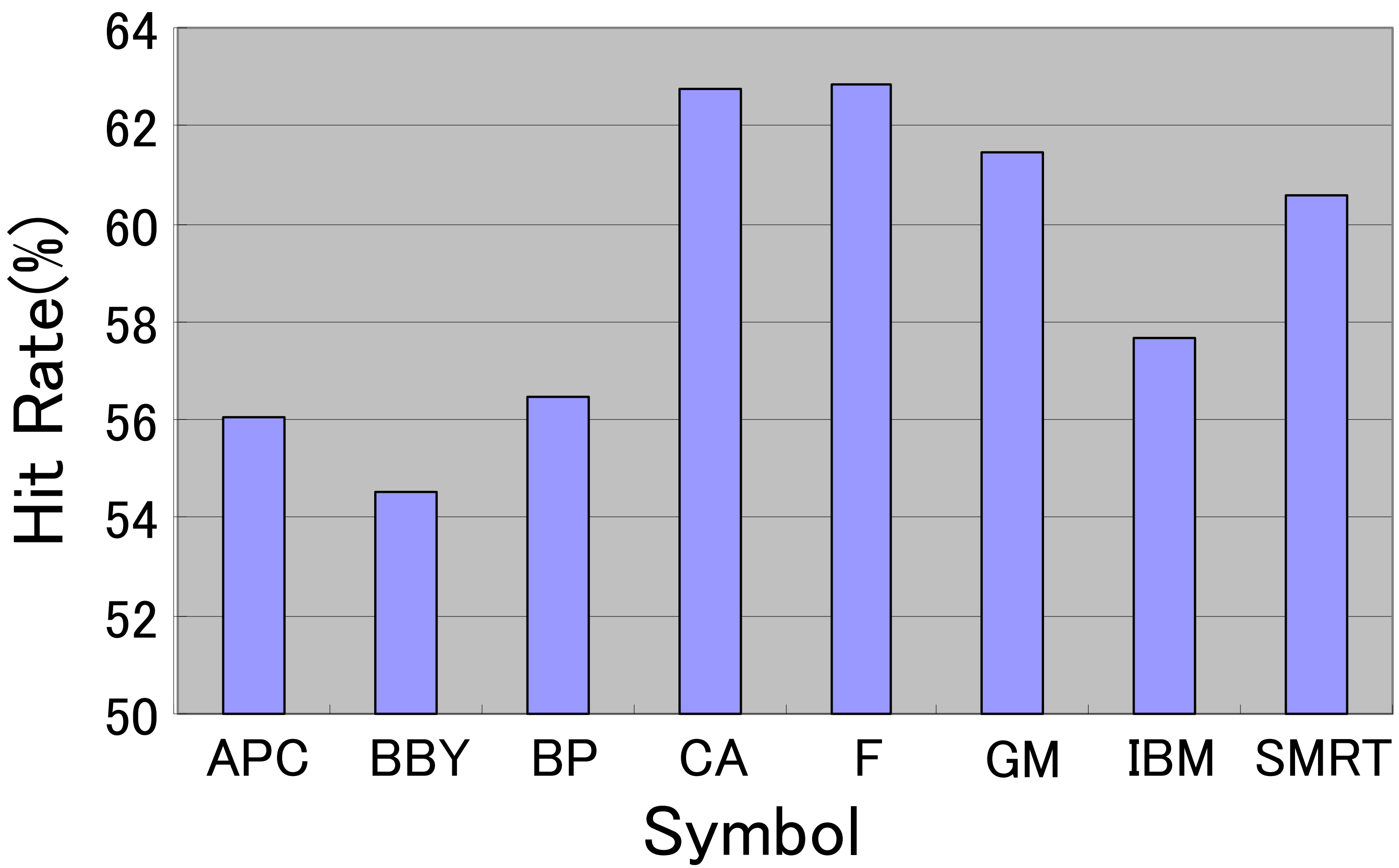
1996-2000



Monthly change (2000)

➡ 70%predictability

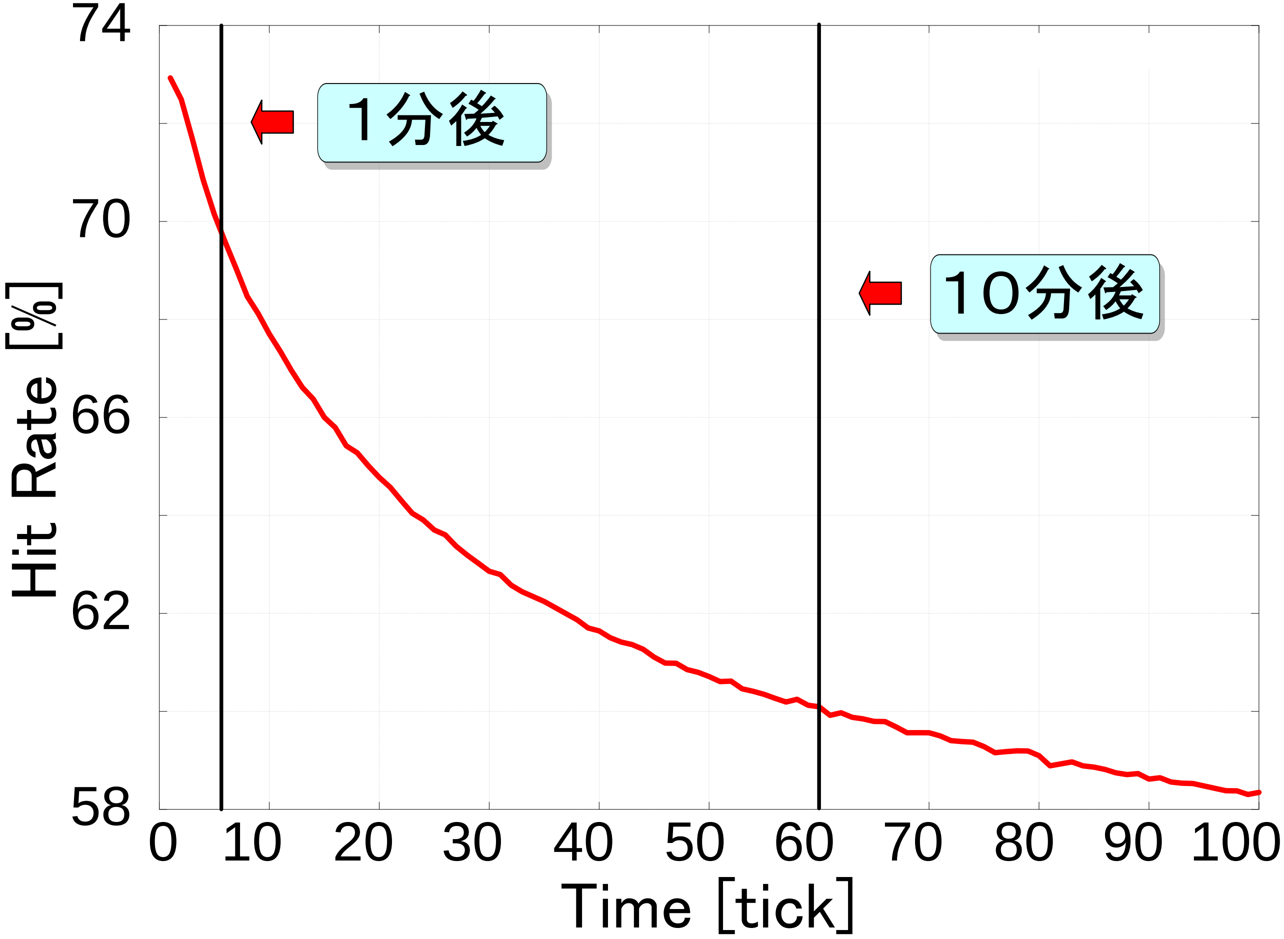
Predictability of 10-ticks ahead



60%能

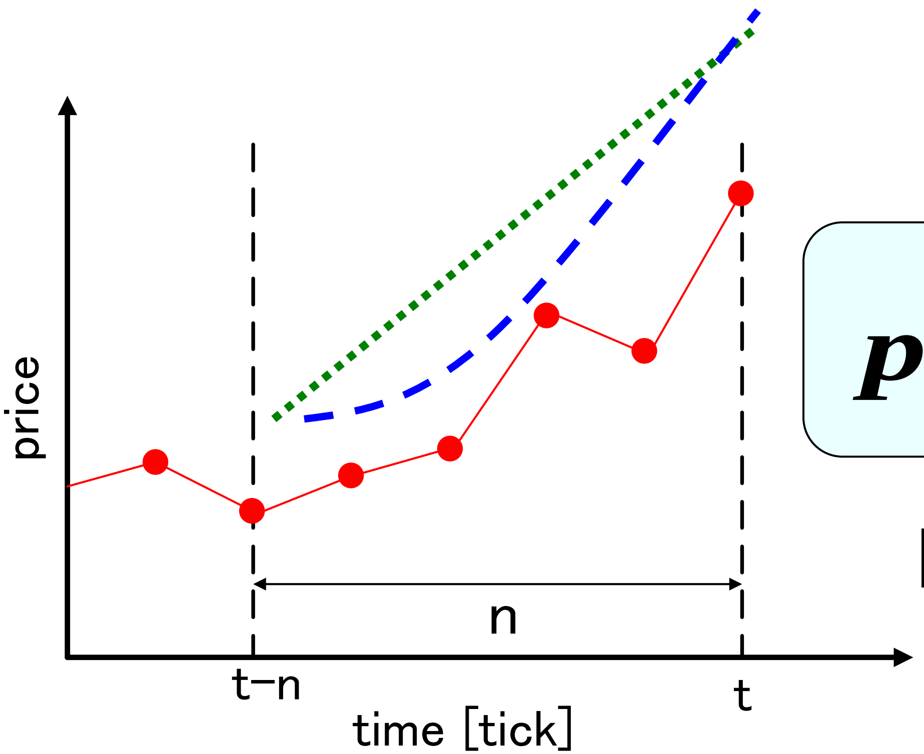
Hit Rate vs. Prediction Term

$/¥ rate(2000)



もっと良い指標はないのか？

Evaluate velocity and acceleration



$$p(t) = A + Bt$$

Initial Price

Velocity

$$p(t) = \alpha + \beta t + \frac{1}{2}\gamma t^2$$

Initial Price

acceleration

Initial Velocity

2 Dimensionless parameters

$$F = \frac{\beta^2}{\alpha\gamma}$$

$$T = \frac{nB}{A}$$

Job flow

Step 1

Set up conditions

Pre-Process

Step 2

Pattern Classification by dynamical parameters

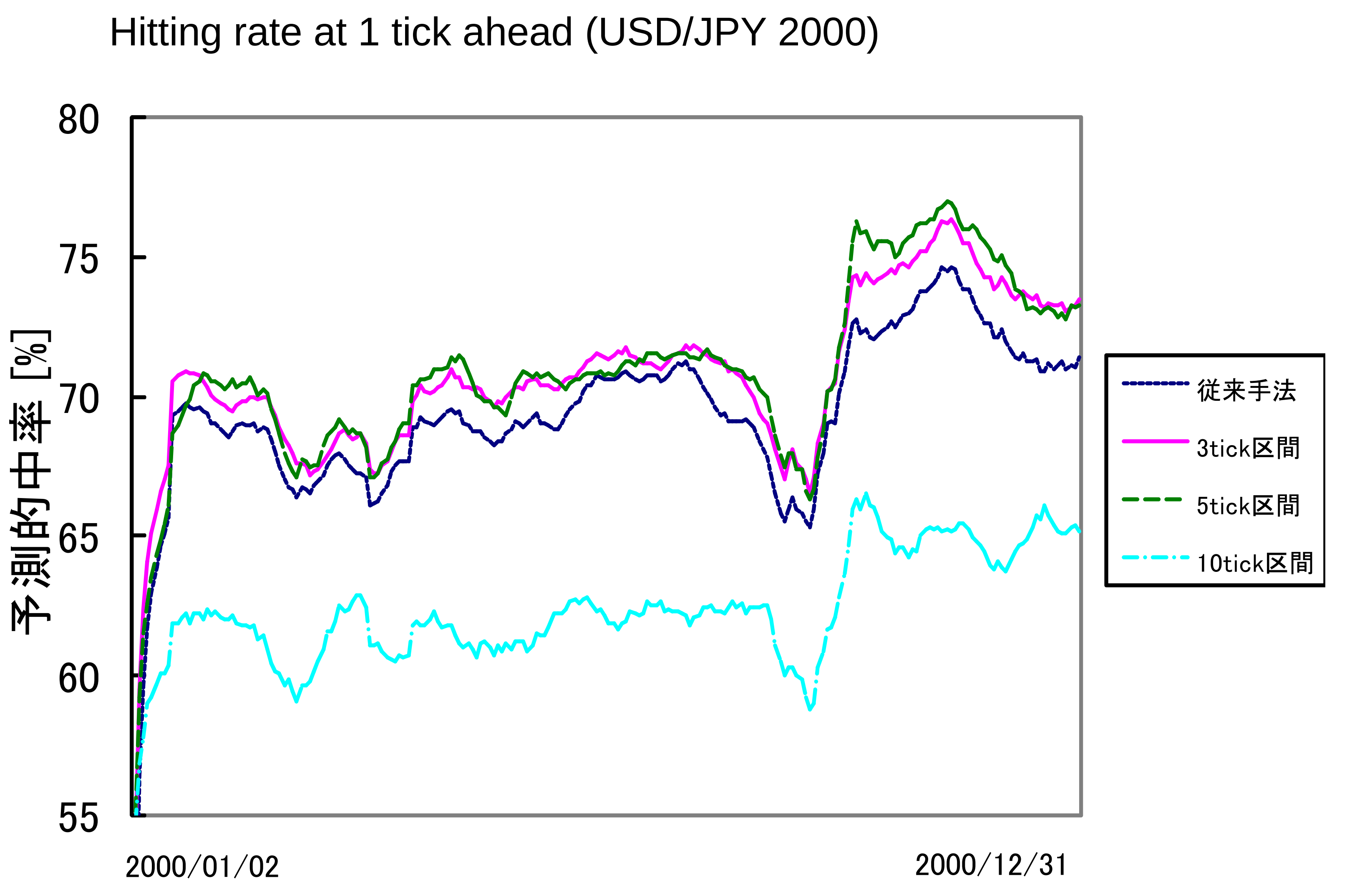
Step 3

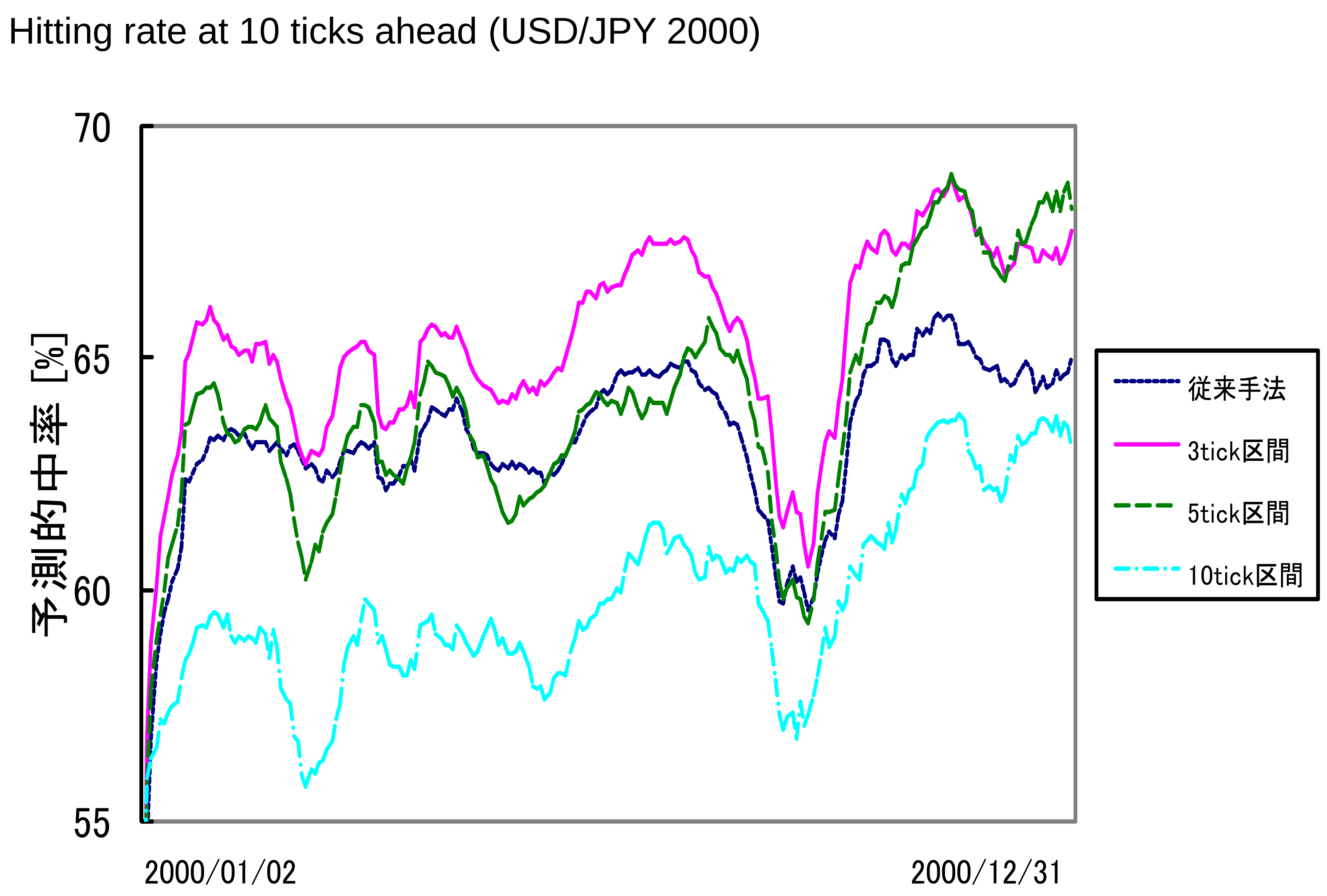
Construct Prediction Strategies

Step 4

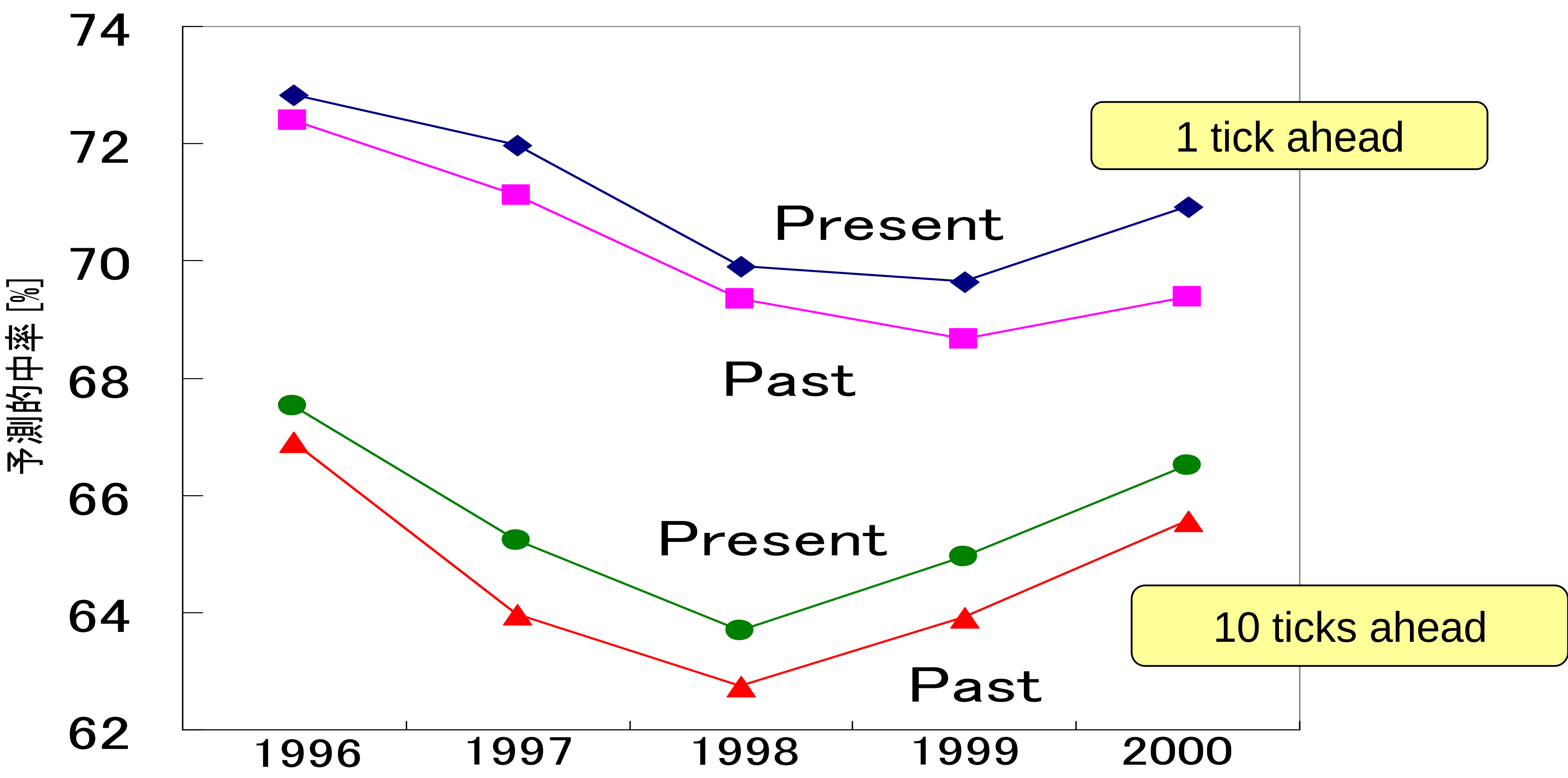
Prediction

Predictions


Hitting rate at 1 tick ahead (USD/JPY 2000)
予測的中率 [%]
80
75
70
65
60
55
2000/01/02
2000/12/31
従来手法
3tick区間
5tick区間
10tick区間

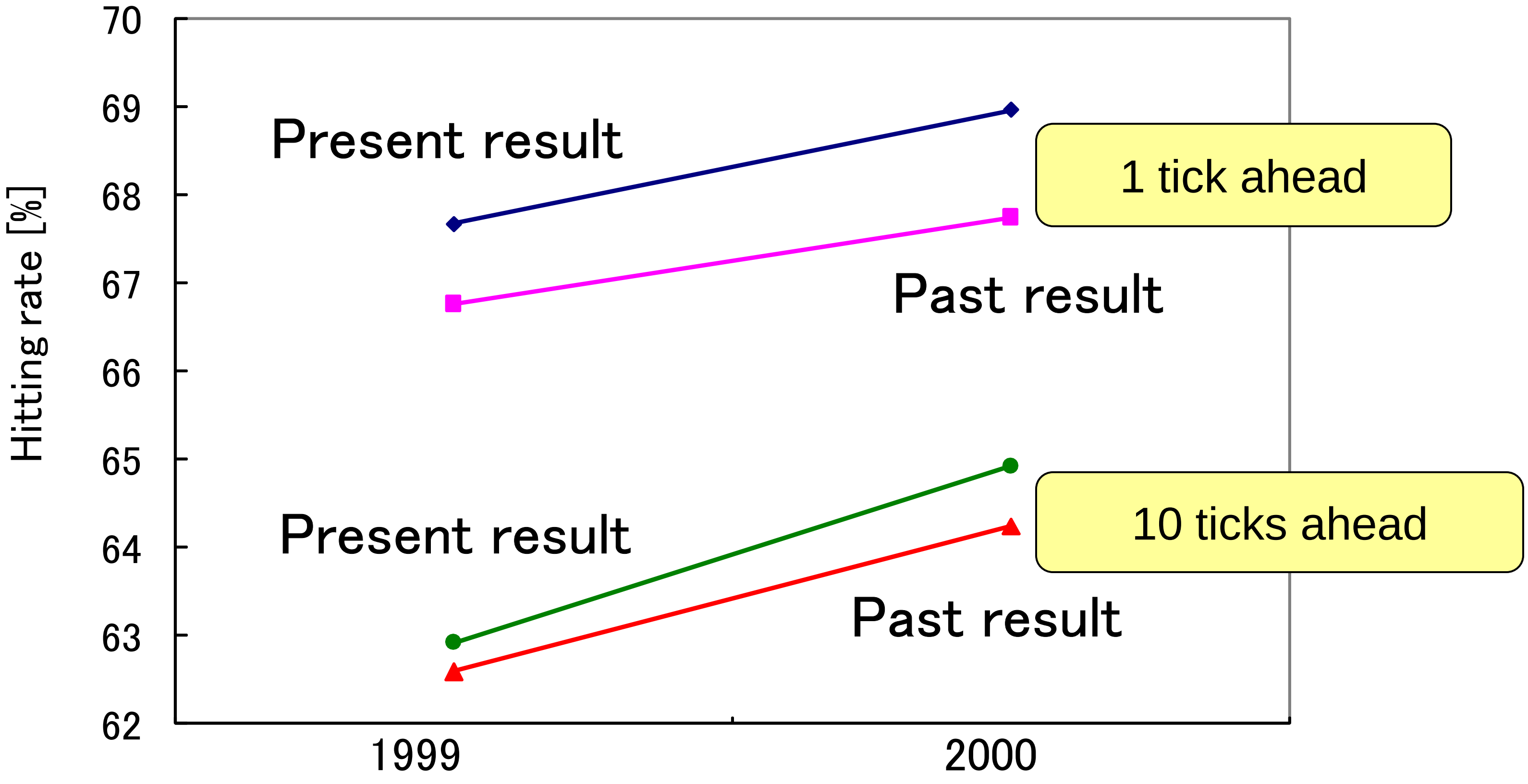

Hitting rate at 10 ticks ahead (USD/JPY 2000)
70
65
60
55
予測的中率 [%]
2000/01/02
2000/12/31
従来手法
3tick区間
5tick区間
10tick区間

Improved Prediction Rates



Hitting rates improved for all years ('96-'00)

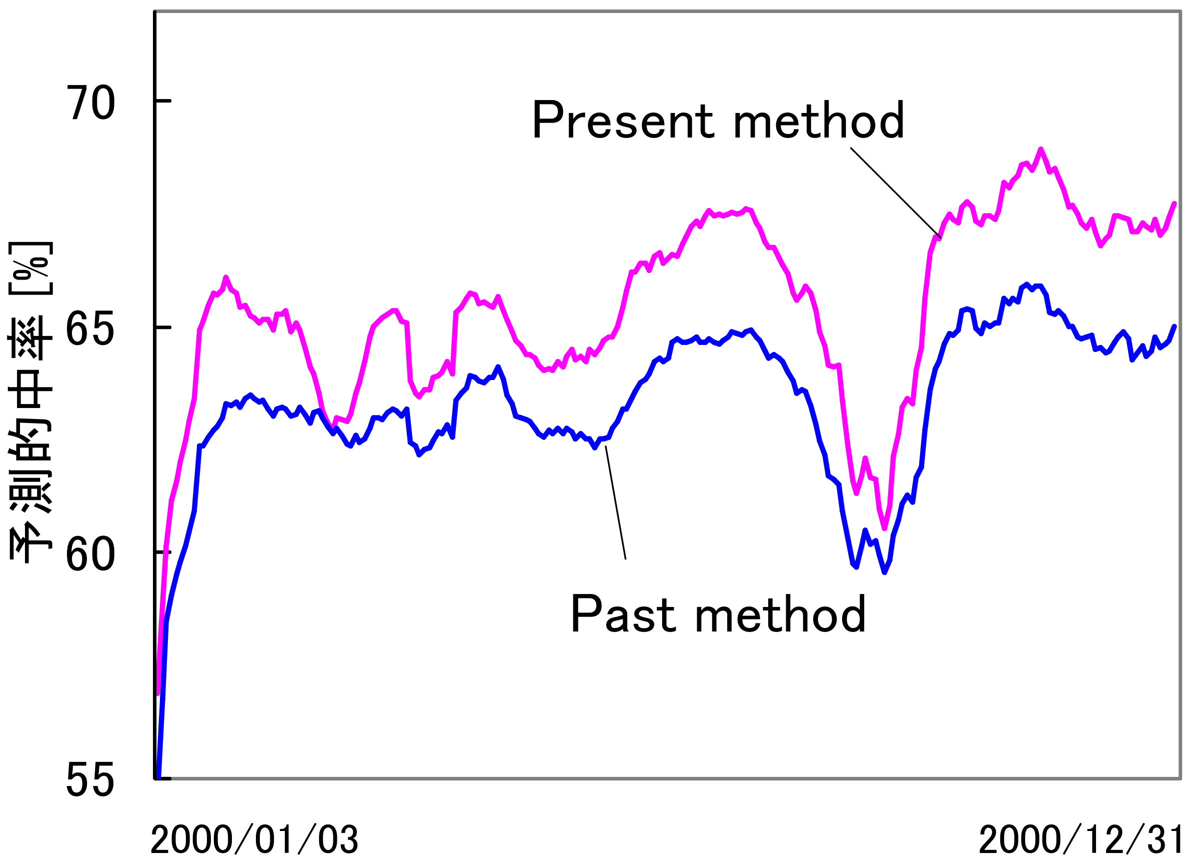
Improved hitting rate USD/JPY



Current result vs. old result in 1999-2000

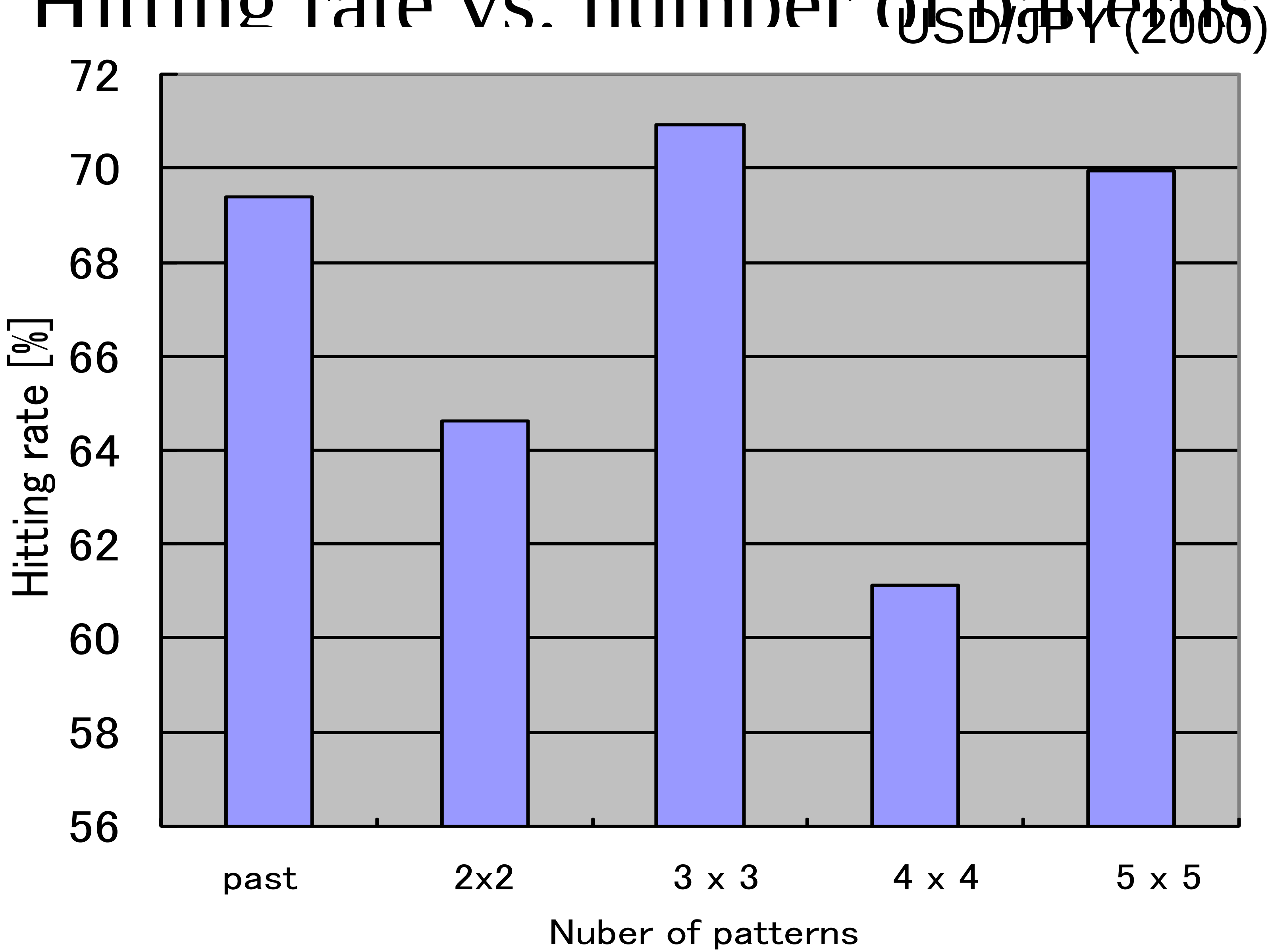
Hitting rate in a year of 2000

(Predicting 10tick ahead data of USD/JPY)



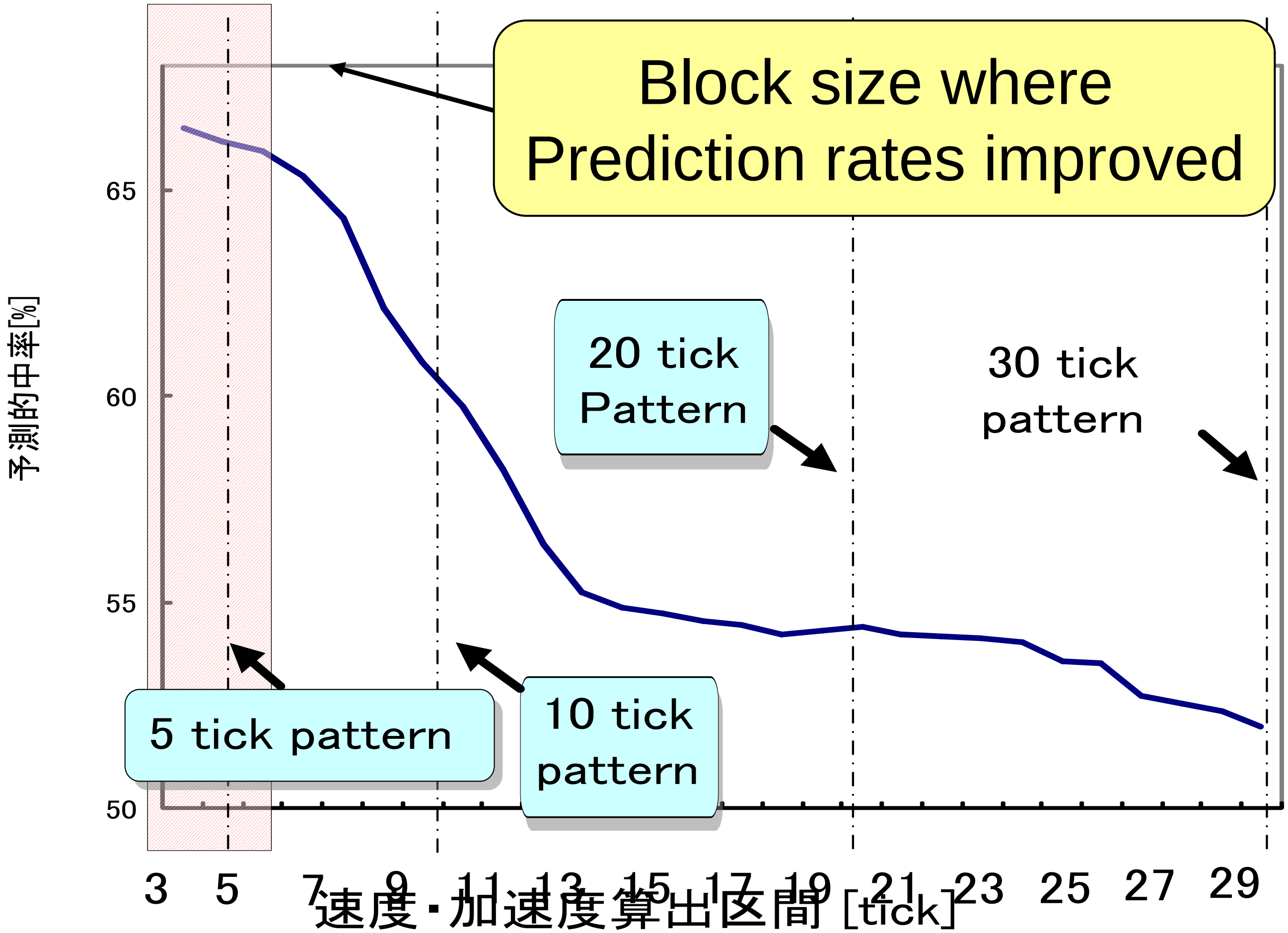
Hitting rate in a year (2000)

Hitting rate vs. number of patterns



Prediction Rate vs. Pattern Length

- USD/JPY(2000) / Prediction at 10 ticks

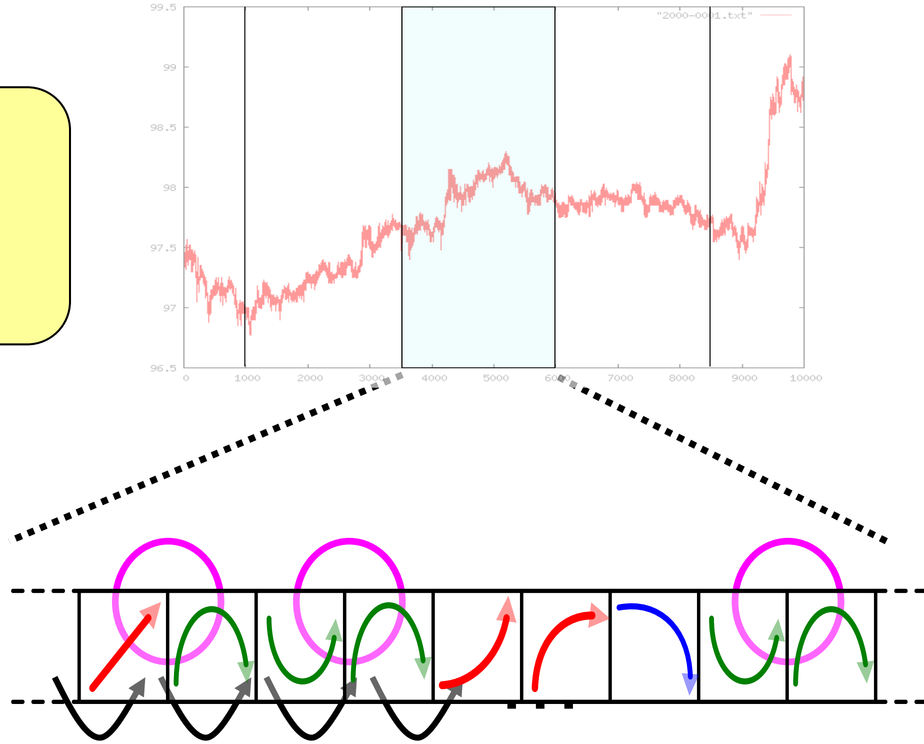


Classifying time series pieces by SOM

① Cut tick data into pieces

② Pattern classification

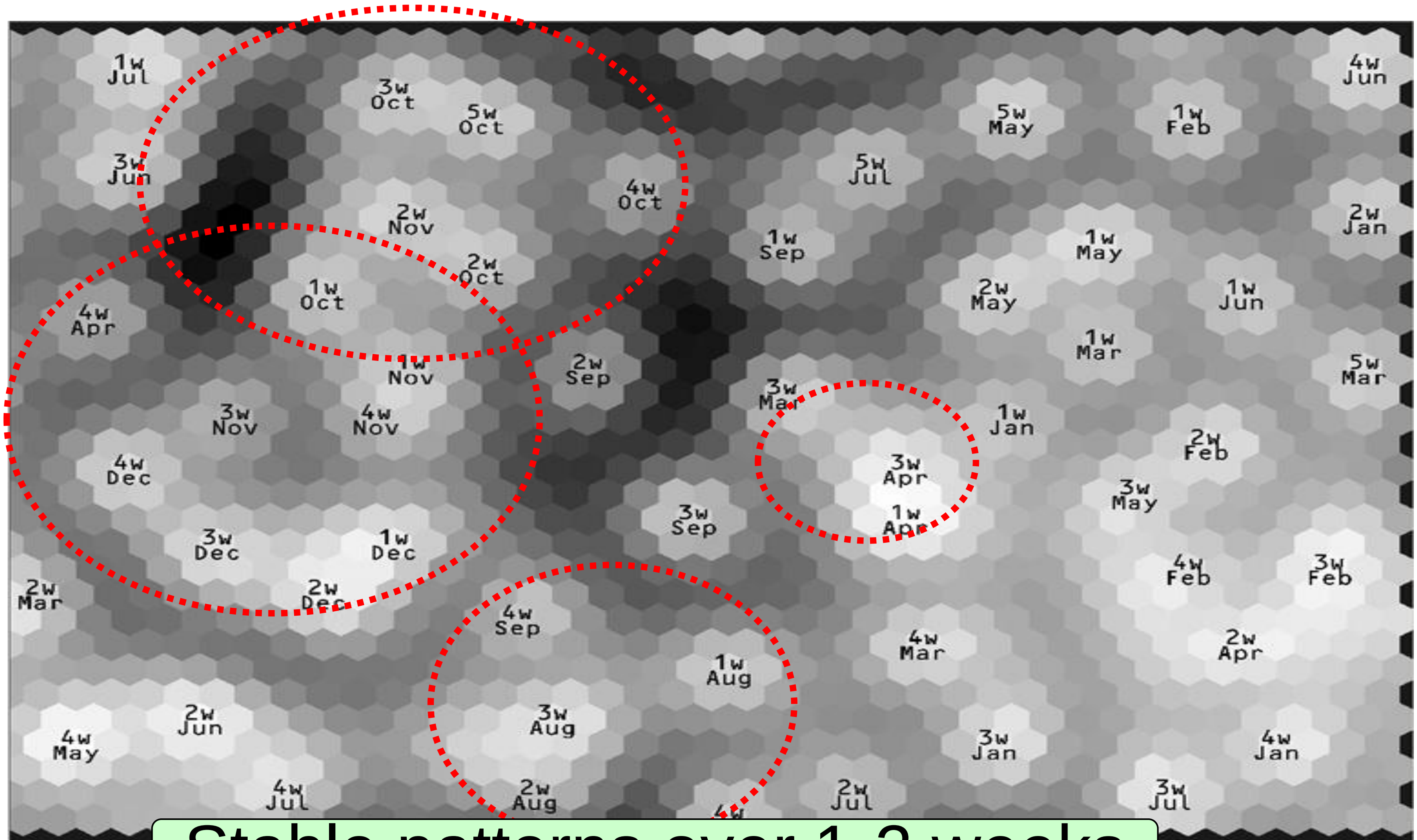
③ SOM classification
of time series pieces



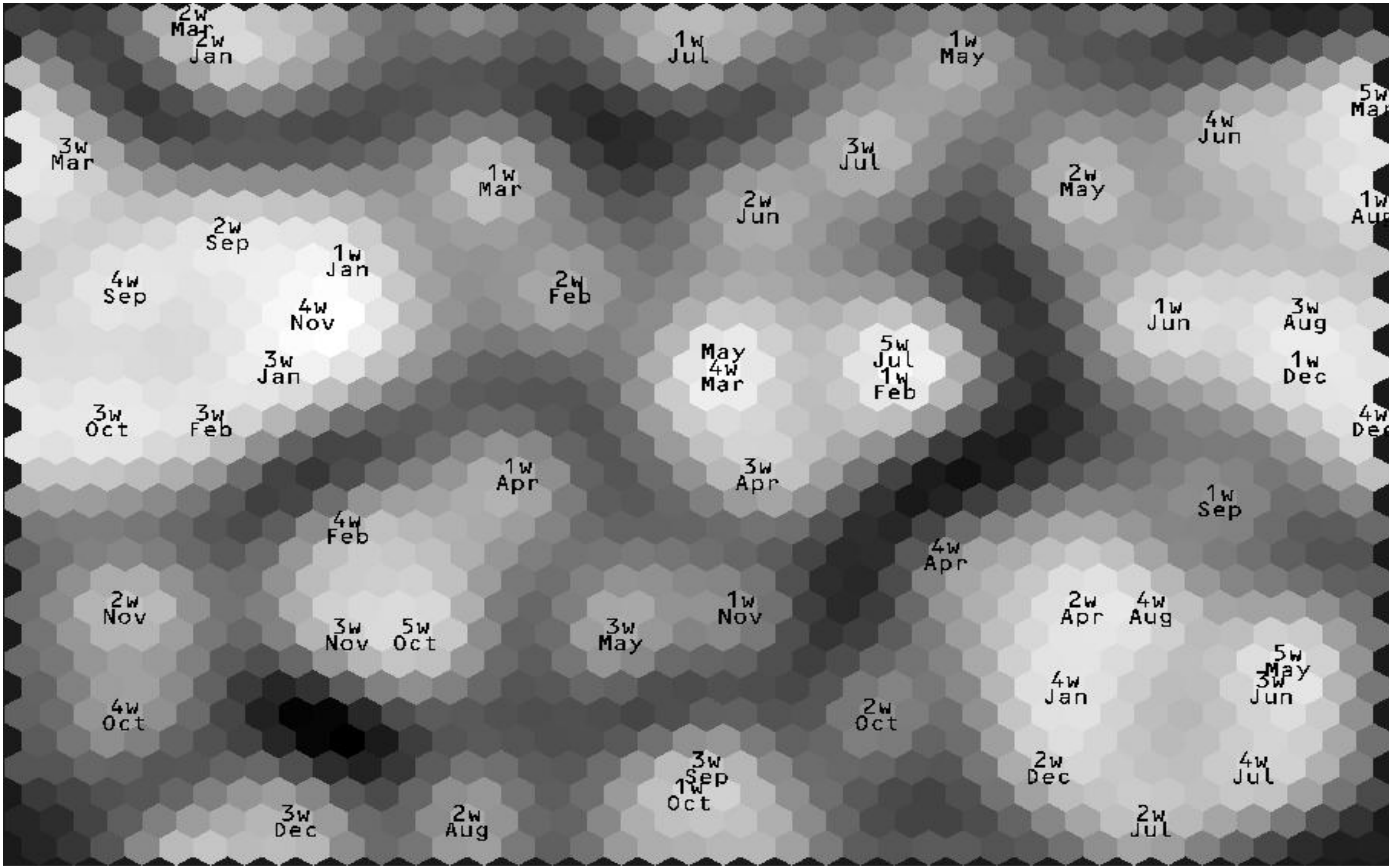
Prob. occurrence of 9 patterns
9 × 9 = 81 transition prob.

Result of classification

- SOM classification of weekly pieces in 2000



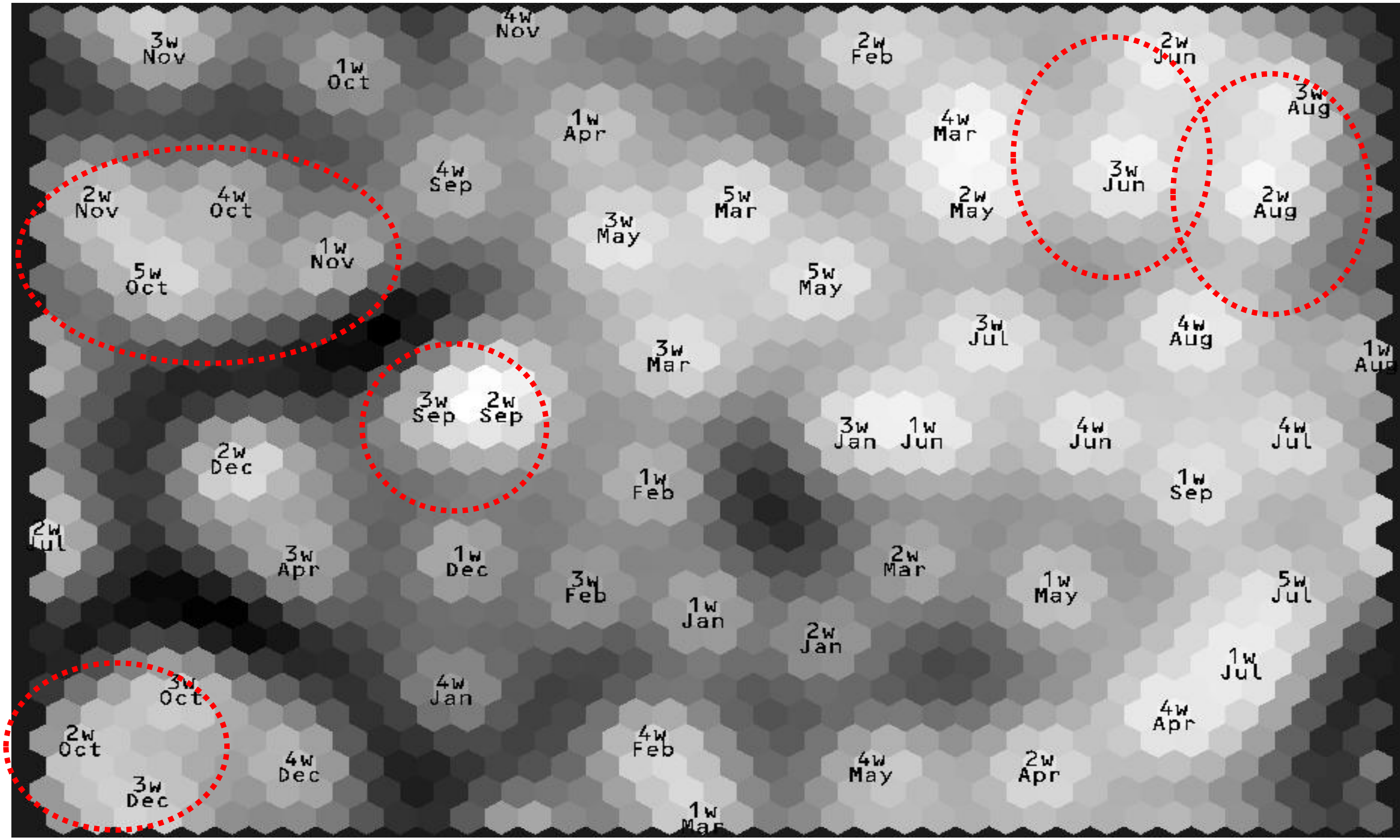
Stable patterns over 1-2 weeks



2w Mar
2w Jan
1w Jul
1w May
5w
4w Jun
3w Mar
1w Mar
3w Jul
2w May
2w Jun
1w
2w Sep
1w Jan
4w Sep
2w Feb
4w Nov
1w Jun
3w Aug
3w Jan
May
4w Mar
5w Jul
1w Feb
1w Dec
3w Oct
3w Feb
4w
1w Apr
3w Apr
1w Sep
4w Feb
4w Apr
2w Nov
1w Nov
2w Apr
4w Aug
3w Nov
5w Oct
3w May
5w May
3w Jun
4w Jan
4w Oct
2w Oct
3w Sep
1w Oct
2w Dec
4w Jul
3w Dec
2w Aug
2w Jul

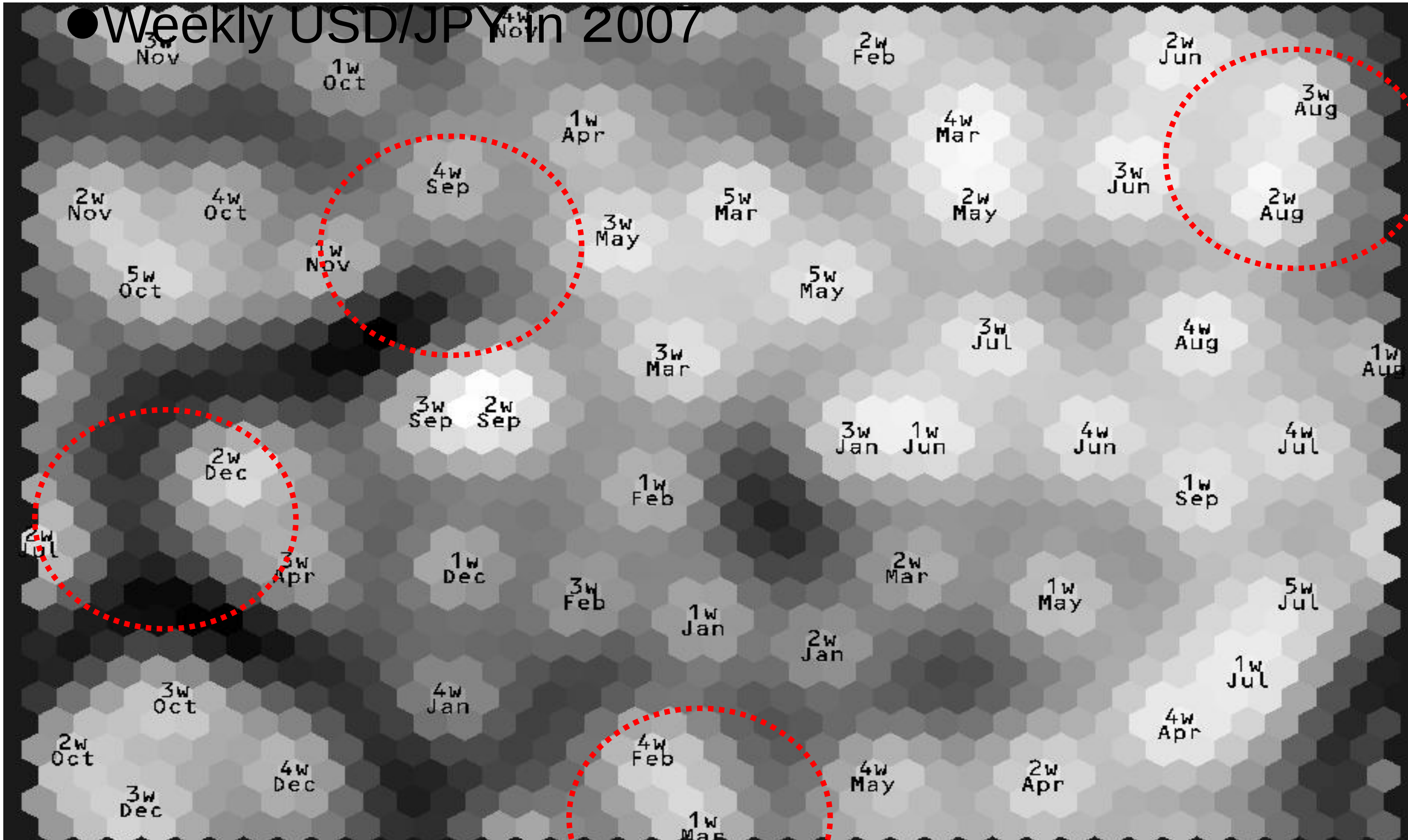
Result : SOM Classification

- Weekly USD/JPY in 1999



Result： SOM classification

- Weekly USD/JPY in 2007



Cross Correlation Spectra compared with Random Matrix Theory

Applied on tick data

NYSE-TAQ 1994 vs. 2002

Use of Random Matrix Theory

How can we separate significant information from the flood of randomness ?

A good recipe is given by

V. Plerou, P. Gopikrishnan, B. Rosenow, L.A.N.Amarmal, T. Guhr, H.E. Stanley,

"Random matrix approach to cross correlation in financial data",

Physical Review E 65, 066126, 2002.

Recipe 1 Correlation Matrix

Cross correlation between Stock- i and Stock- j

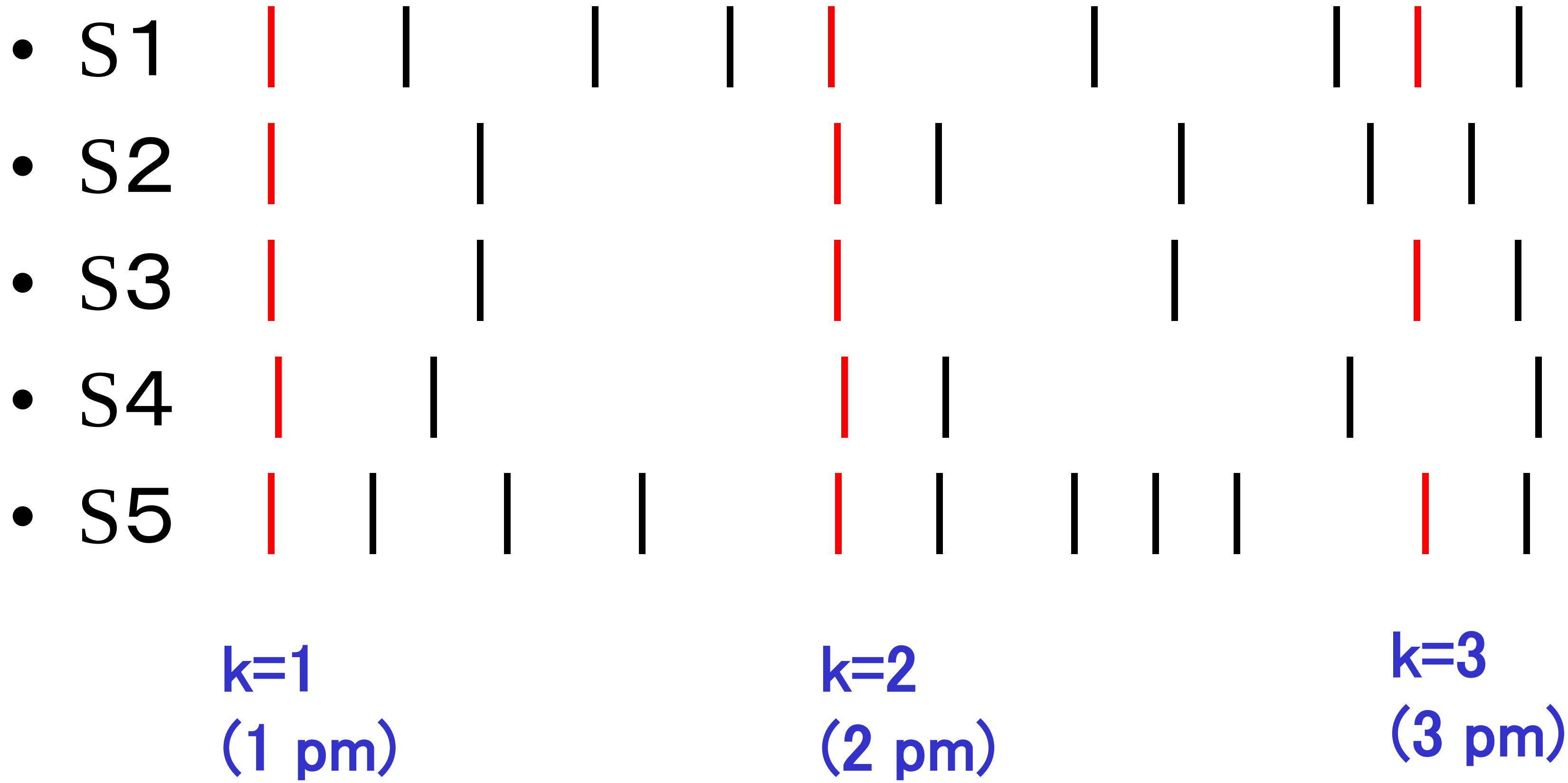
$$C_{i,j} = \frac{1}{L} \sum_{k=1}^{L} x_{i,k} x_{j,k}$$

$$x_{i,k} = \begin{vmatrix} x_{1,1} & x_{1,2} & \cdot & x_{1,L} \\ x_{2,1} & x_{2,2} & \cdot & x_{2,L} \\ \cdot & \cdot & \cdot & \cdot \\ x_{N,1} & x_{N,2} & \cdot & x_{N,L} \end{vmatrix}$$

← Price time series of Stock-1

Data at every tick (k=1,..,L) for all N is needed!

Trouble of tick-wise stock prices



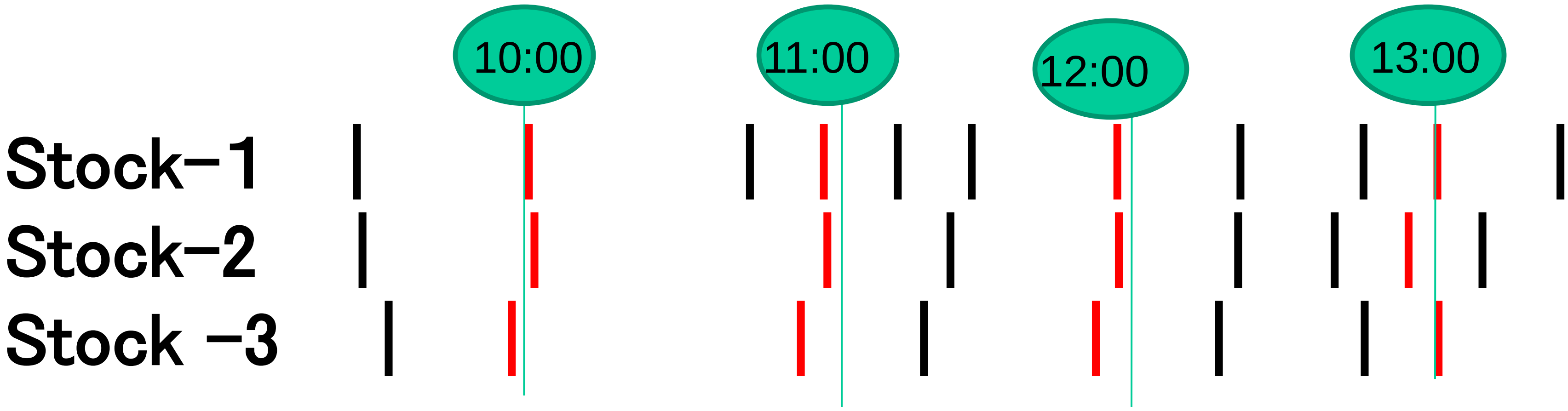
1) Time intervals between ticks are not regular
2) Not all symbols have values at every k

N↑ means L↓

- L=1,512 for every 1 hour

(10 am-3 pm = 6 pts × 252days = 1,512)

- 1994 NYSE-TAQ : N=419
- 2002 NYSE-TAQ : N=569



Use the nearest prices to every hour

7 step recipe

1) Get price time series $S_{i,k}$ (i=1,.., L) for k=1,.., N stocks

2) Compute return time series

$$X_{i,k} = \ln(S_{i,k+1} / S_{i,k}) \approx \Delta S_{i,k} / S_{i.k}$$

3) Compute $x_{i,k}$ by normalizing $X_{i,k}$ (mean=0, variance =1)

4) Compute cross correlation matrix $C_{i,j} = \frac{1}{L} \sum_{k=1}^{L} x_{i,k} x_{j,k}$

5) Solve eigenvalue problem of

$$CV = \lambda V$$

6) Compare the eigenvalues with corresponding RMT

7）Discrepancy is the SIGNIFICANT INFORMATION

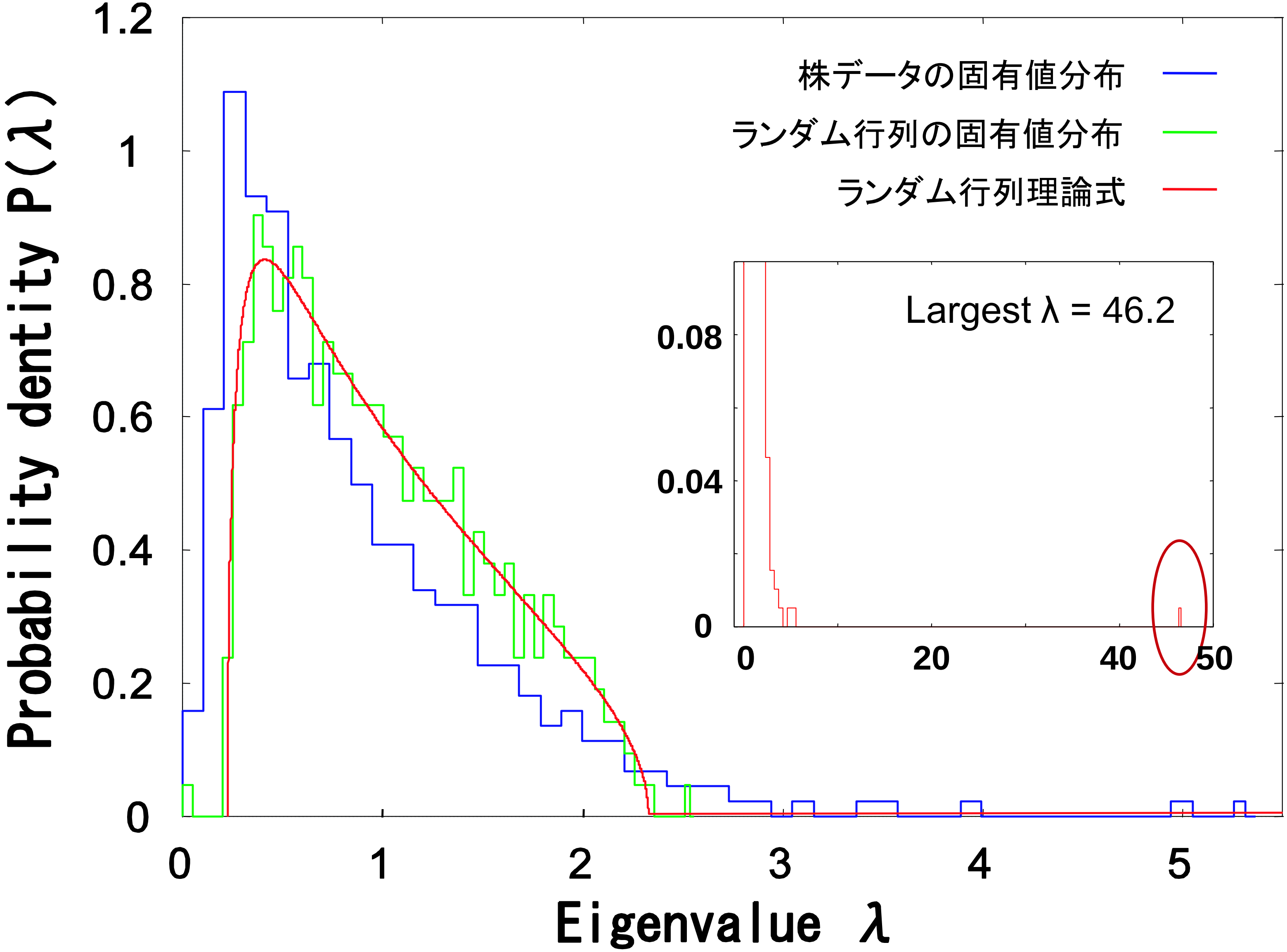
RMT tells us..

$$\mathbf{N \to \infty, \ L \to \infty, Q = L / N = const.}$$

Eigenvalue (λ) distribution is given by

$$P_{RMT}(\lambda) = \frac{Q}{2\pi} \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{\lambda}$$

$$\lambda_\pm = 1 + \frac{1}{Q} \pm 2\sqrt{\frac{1}{Q}}, \quad (\lambda_- < \lambda < \lambda_+)$$



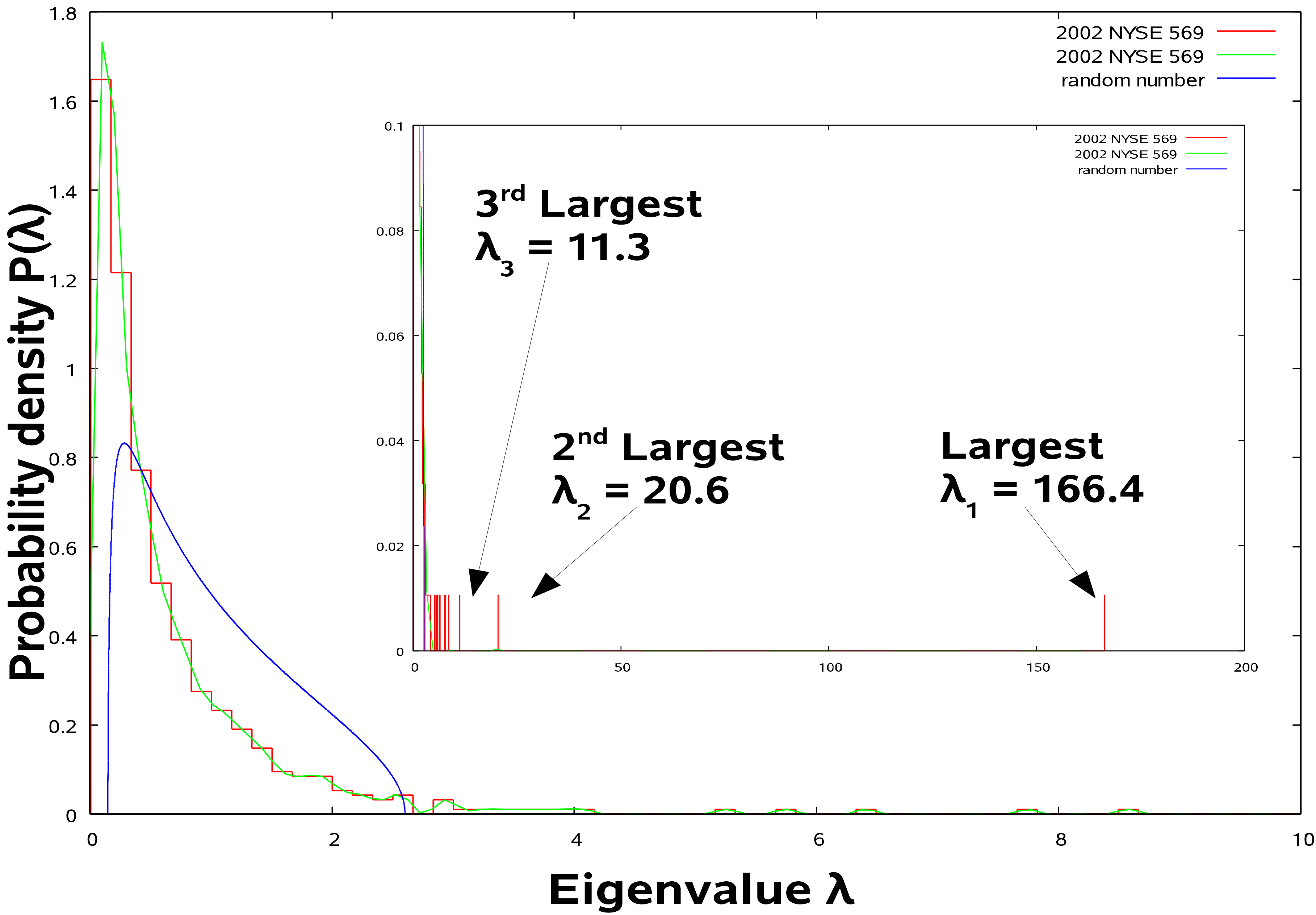
$\lambda_3 = 5.0, \lambda_4 = 3.9, \lambda_5 = 3.5, \lambda_6 = 3.4, \lambda_7 = 3.1$

SIGNAL:

NOISE: λ_8

Eigenvector Configulations(94)

EIGENVALUE	**EIGENVECTOR**		
λ_1 = 46.2	BIG COMPANY	BIG COMPANY	BIG COMPANY
λ_2 = 5.25	GOLD MINING	---	---
λ_3 = 5.04	SEMICONDUCTOR	SEMICONDUCTOR	SEMICONDUCTOR
λ_4 = 3.90	SEMICONDUCTOR	SEMICONDUCTOR	SEMICONDUCTOR
λ_5 = 3.51	OIL	OIL	OIL
λ_6 = 3.41	---	---	---
λ_7 = 3.11	PAPER	PAPER	---
λ_8 IN NOISE	---	---	---
λ_9 IN NOISE	FINANCE	FINANCE	---
λ_{10} IN NOISE	AUTOMOBILE	AUTOMOBILE	COMMUNICATION


Probability density P(λ)
Eigenvalue λ
2002 NYSE 569
2002 NYSE 569
random number
3rd Largest
λ3 = 11.3
2nd Largest
λ2 = 20.6
Largest
λ1 = 166.4

Eigenvector Configulations(02)

EIGENVALUE	**EIGENVECTOR**		
λ_1 = 166.4	BANK	BANK	BANK
λ_2 = 20.6	FOOD	FOOD	FOOD
λ_3 = 11.3	ENERGY	ENERGY	ENERGY
λ_4 = 8.6	FOOD	FOOD	FOOD
λ_5 = 7.7	ENERGY	ENERGY	ENERGY
λ_6 = 6.5	ELECTRIC	ELECTRIC	ELECTRIC
λ_7 = 5.8	FOOD	FOOD	FOOD
λ_8 = 5.3	RETAIL	RETAIL	RETAIL
λ_9 = 4.1	METAL	METAL	METAL
λ_{10} = 4.0	COMMUNICATION	COMMUNICATION	COMMUNICATION

1 hour data vs. daily data

1. Qualitatively almost same obtained
2. More random in 1 hour data
3. 1994 are consistent with Gaussian RN, while 2002 data seems different

   (need filtering ?)

Future perspectives

4. Analysis of all the years between 1994 - 2002 ongoing
5. Other markets, e.g. S&P500, TSE, etc.
6. Other kinds of data, e.g.,
- Demography
- Climate
- Earthquake, etc.