系列現象の統計的解析—II
株価変動の統計的解析

赤　池　弘　次

(1954 年 1 月 受付)

Statistical Analysis of Serial Phenomena-II
Statistical Analysis of Stock Price

Hirotugu Akaike

In this report, the author tried to make clear the key points in the statistical analysis of serial phenomena. The former report on the statistical analysis of propagation phenomena by the same author is thought to constitute the part I of this series. Taking the analysis of stock price as an illustration, he sought to trace the process of development of our knowledge by the analysis of serial phenomena.

In analysing the serial phenomena statistically, following points seem to be of importance;

1. to settle proproper time unit
2. to seek multidimensional representatation
3. to select sample efficiently.

Some formulations abont the variable population or the loacl estimation procedure are obtained, which are entirely different in idea from models being adopted dominantly in time series analysis nowadays.

Institute of the Statistical Mathematics

前　書　き

統計数理研究所輯報 第 11 号に発表された伝播現象の統計数理的解析 －I に次ぐものとして，系列現象の統計的解析を，株価分析について試みた結果を報告する．系列現象とは或る外的な変数によつて，順序づけられる事象のつらなりを意味するわけであつて，その最も一般的な例としては，時間と共に変動する現象がある．また連続的な物体からのサンプリングを考えること等もその一例である．従つて，所謂伝播現象も当然一つの系列現象と見做されるわけで，その意味では，輯報第 11 号の報告は，系列現象の統計的解析 －I とさるべきであり，従つてこの報告は，系列現象の統計的解析 －II とした．以下の記述に於ては，輯報第 11 号の報告を，報告 I と記すことにする．

§1. 報告の概要

§1.1 分析の目的

所謂時系列解析は，その問題とされる分野は極めて広く，自然現象，社会現象の何れに於ても，その適切な分析法の発展が強く要求されている．しかし，現在通常試みられる分析法の多くは，自然現象の記述に際して比較的有効であつた数式的模型の形式的解釈に過ぎないものであつて，その社会現象への適用は多くの場合極めて不満足な結果しか与えない．殊に経済現象に關しては外的條

本報告は昭和 28 年度文部省科学研究助成補助金による研究成果の一部である

件の変動が著しく統計的方法の適用そのものについても問題があるとする人もいる．このような状態になつた原因と考えられるものは次の点であろう，即ち，現実の問題そのものの本質に近づくことを試みることを忘れ，既に定式化されたものをその定式化の過程に対する批判無しに形式的にあてはめを行おうとする結果，問題の質的な差を見落すことになつたことであろう．殊に，現象の巨視的状態の説明にその主眼点を置いた自然現象説明のための模型を以て，社会現象に於ける多くの問題のように時間の価値を積極的に考慮するいわば微視的な問題解析の手段としようとするところに最大の原因があるものと言えよう．このような問題を，ひろく系列現象解析の立場より見，統計的方法を適用する時の考え方の発展径路を振返つて見ることにより，そこに新しい接近の仕方を見出すことを試みたい．この分析の目的は，株価変動の分析を例として，この種問題に共通と思われる統計的方法適用の際の問題点を明かにすることによつて，従来の方法の単なる形式的機械的適用が，如何なる点で全く無力であるかを見，統計的方法のより正しい適用を考え，更に新しい統計的方法への要求を明確化し，可能な限りその定式化を試みることにある．

さて当面株価変動の統計的分析の目標とするところは，株式投資に関して有効な知識の獲得を目指すものであることは勿論である．株式投資に際し，注意すべき事項として，各専門家の説くところは極めて多種多様である．それらの株式観測観の底流をなすところの株価変動の本質を統計的方法の活用によりできる限り明瞭に把握することを試みるわけである．

§1.2 株価とは何か

株式市場に於ける取引によつて，株価が構成されるのであるけれども，我々が実際知識を獲得しようとする株価は，株式投資の種々の型に従つて，その表現を変化させなくてはならない．所謂株価そのものは現実に構成されているのだが，それは極めて複雑多様な面を持つているわけであり，我々が実際株式投資に際して注目する株価は，その特定の一側面をなすわけである．結局，複雑な現実の株価を，投資の様式に従つて，最も有効な信号のみを通過させる filter を通して見るわけであつて，投資の様式と filter の構造とは互に他を規定し合つているわけである．この意味で，我々が通常株価と呼ぶものは，具体的な投資に関連して考えられる限り，一つの filter を通した株価を意味することになる．ここでは，従来の多くの解析例に見るように無意識に特定の filter を通した株価のみに注目するという態度を捨てて，filter の設定の仕方と，それによつて得られる知識の型との関聯を全面的に考察することによつて，株価とは何かという問題への接近をこの分析の全過程を通じて試みたい．

§1.3 分析に用いられた資料

分析に用いられた資料は主として，昭和 27 年 3 月 27 日より昭和 28 年 3 月 27 日迄の新聞紙上に掲載されたデータと，昭和 28 年下期及び昭和 29 年上期の会社要覧とである．

§1.4 分析の結果

現在迄の分析結果について簡単に述べると，投資家の株式評価は，各株式会社の利益率と並行しているという事実が最も著しい．

これは，株式の所謂利潤証券説を相当明瞭に裏書きしているものと考えられる．次に，各種専門家の株式評価の基準が極めて多様であり，従つて，一般の投資家もその何れに従うべきかを明瞭にすることが出来ないのであるが，我々はこの分析に於て，考慮する枠が変化するとともに，注目すべき各種條件に対して，賦与すべきウエイトが変化するということの著しい実例を見る．例えば，一株当り再評価積立金が，大なる程株価は高い傾向が僅かに見られるのであるが，これを他の，利益率，配当率，一株当り正味資産，流動比率，固定資産回転率，固定比率，資本金等と同時に考え合わすと，現実の株価構造に於ては，大勢より見ると他の各種要件一定の場合には，一株当り再評価積立金が多い会社の株の方が，株価が低いという傾向が見られる．この事実は，深く株価の構造

を知る人から見ると極めて当然であるにしても，一般の人が，株式評価に際して，どのようにウエイトを置いているかは，このような統計的分析によつて始めて明瞭となる事である．更に投資家の株式評価が，利益率に並行するという事実を，時間的に追跡することによつて，投資者の評価と，利益率との並行性は，各半期の終末の時期に最も著しく，その時期を境として，それ以前，それ以後の何れに於ても並行性は時間的隔りと共に減少することが見られる．この事実は，株式市場に於ける投資者の動きを解明する一端緒となるであろう．取引に際し，手数料の支拂いを必要としないならば，この事実をもとにして考えると，一刻も早く各会社の営業成績を察知することにより，恐らく相当の利益を確保することが可能とされるのではないかと思われる．

§1.5　系列現象解析に際し特に注意を要する事項

本報告，及び報告-Ｉに於てとりあげられた解析例によつて，系列的な現象一般の統計的解析に際して特に注意を要する点として次のようなものが考えられる．

1.　解析の時間單位の問題

2.　表現の問題

3.　標本設定の問題

これ等の問題を取上げ，それについて見ることがこの報告の主なる目的であつて，それについては，§3，§4，に於てくわしく述べる．

§2.　分析の詳細

以下分析の過程を，その時間的順序に従つて記述する．それによつて，知識の追求がどのような経路をたどつて行われるかを見，知識獲得の為の最も有効な進み方を把握しようというのである．

§2.1　報告Ｉの研究より，株価分析への移行について

報告Ｉ§3 に於て，電界強度記録を線型自己回帰的な時系列として考察することを試みたが，これは，電界強度記録の変動の様相を説明する為のモデルとしての効果はあつたが，それによつて積極的な効果を得ることはできなかつた．

そのような考察の際，中心的な役割をなすのはコレログラムであるので，当然系列相関係数の分布に関する知識が要求される．そのため，報告Ｉ§3 に於て得られたモデルを利用し，ランダムファクターに，乱数表の乱数を用いてモデル系列をつくりその系列について，系列相関係数を計算して見た．モデル系列の作製に際しては，初期値の影響を考慮して，最初の 50 項を除き，その結果継続する 300 項を得た．この 300 項を最初より 100 項づつの 3 組に分け，各組を順番にＩ，II，III，組とし，各組に於て 1 時点おくれの系列相関係数を，初めから n 項，($n=10, 20, \cdots\cdot 100$) に対して求めて見ると，対応する自己相関係数 $\rho(1)$ は $\rho(1)=0.703$ であるが，次のようになる．

組 ＼ n	10	20	30	40	50	60	70	80	90	100
I	0.630	0.650	0.675	0.677	0.675	0.669	0.678	0.679	0.694	0.698
II	0.247	0.398	0.712	0.631	0.662	0.594	0.526	0.486	0.530	0.540
III	0.272	0.317	0.323	0.392	0.359	0.319	0.285	0.328	0.384	0.410
平　均	0.383	0.455	0.570	0.567	0.565	0.527	0.496	0.498	0.536	0.549

又，初めよりの各 10 項毎に 1 時点おくれの相関係数を求めたもの $r(1)_i (i=1, 2, \cdots\cdot 30)$ によると，

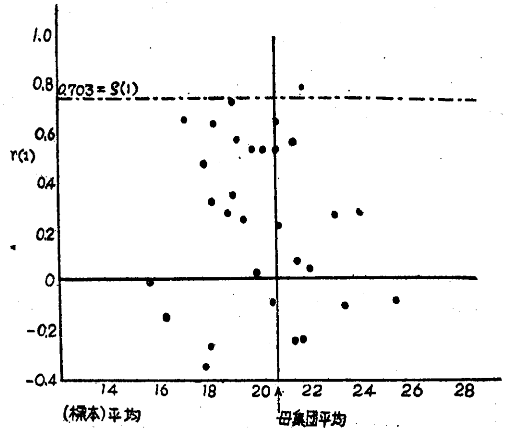
$$\bar{r}(1) \equiv \frac{1}{30}\sum_{i=1}^{30} r(1)_i = 0.247$$

$$S \equiv \sqrt{\frac{1}{30}\sum_{i=1}^{30}(r(1)_i - \bar{r}(1))^2} = 0.325$$

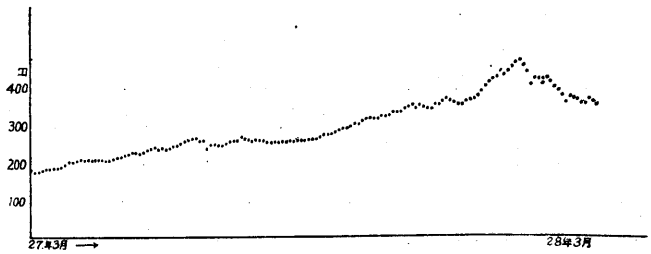
$$\rho(1) - \bar{r}(1) \geqq 7.6 \times \frac{S}{\sqrt{30}}$$

となることが見られた.*

各 10 項毎の平均値と $r(1)_i$ との分布図は次のようになる.



此の結果より見ると明かな通り，所謂時系列解析に際して，形式的に従来の数式模型を現実のデータに無反省に適用しても，見るべき結果を得る場合は少いのではないかと考えられる．このことを東京証券取引所の 225 種平均株価のデータについて見ることにする．先づこの一年間の変動の様子を図示すると次のようになる．各日のダウ式平均株価の概略の図である．



この図より明かであるように，この程度のデータ数に対して，平均株価の系列を所謂定常な時系列とみなすことは殆ど不可能であり，従つて系列相関係数を計算することは意味が無いと考えられたので，先づ時間（日）を独立変数とする一次式を，各継続する 10 日間のデータにあてはめ，その第 11 日目の価を外挿することを試みたのであるが，その結果は，22 例中 16 例迄は單に前日の平均株価を以て，翌日の平均株価の推定値とする方が，外挿値よりも良いことが見られた．このことは，局所的にも單なるあてはめによる外挿が全く無力であることを示す例と考えられる．これ等の結果の示唆するところに從い，平均株価の系列を，むしろ異質的な母集団の連りからのサンプルの系列として考察し，その母集団の変化を明示するような並列する標識を設定することを試みた．その一例として，毎週に於ける日々取引高量の変異係数を計算して，それを平均株価の系列と並行させて見たところ，27 年 9 月に於ける平均株価の曲線の彎曲点の辺に於て，並行する取引高量の週

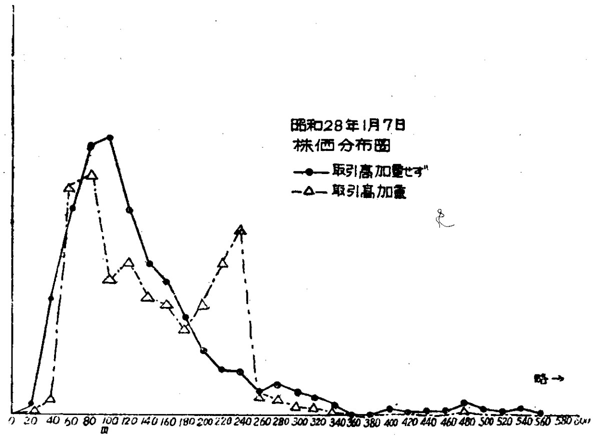
* この問題については次の文献参照

T. C. Koopmans 編 Statistical inference in dynamic economic models: Wiley, 1950 の中の P. 365～383.

L. Hurwicz Least-squares bias in time series.

間変異係数が，極めて低くなることが見られた．

このような事実は，更に数多くのデータについて見るならば，何等かの有効な知識を与えるようになるかもしれない．何れにしても，単に平均値の変化のみを追跡するよりは，このように並列する多数の系列を考えることがより有効であろうということは明かとなつて来た．ここで更に平均株価の定義に関する反省が必要とされた．というのは，単純平均株価に於ては，当日の取引高のウエイトを全然用いていないのであるが，実際には，取引高によるウエイトをつけた平均値を考える方が妥当ではないかというのである，そこで何日かのデータによつて，株価（終値）の単純な分布と，取引高量のウエイトをも考えに入れての分布とを取つて比べて見ると次のことが分る．即ち，単なる株価の分布は，比較的安定した滑かな分布を示すが，ウエイト附きの分布は，遙かに不安定な様相を示している．これより考えると，所謂平均株価として，単純な平均を利用することは妥当だということになろう．ウエイト附きの平均株価はそのふらつきが大きすぎるため，注目する指標としては適当でないと考えられる．むしろ各銘柄について，毎日の株価と取引高とを二次元の標識として考えることが妥当である．両方の分布の一例を下図に示す．



さて，自己回帰型の適用も，時間の単位を適当に定め，相当長い先の株価の推定に利用するならば，かなりの効果を上げることが全くないとは限らない．ここでは隣接する5日間のデータの平均値の系列を作り，その10項によつて，1項おくれの系列相関係数を計算し，それを利用して

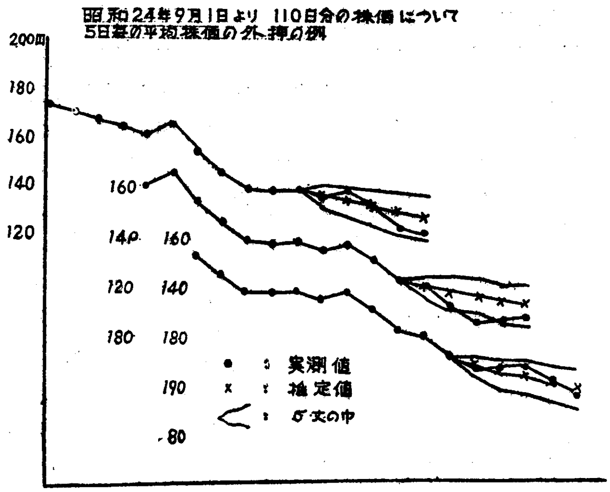
$$X_t = r(1)X_{t-1} + \varepsilon_t$$

なる型をあてはめて，次々に條件つき希望値を計算することによつて，6項程度までの外挿をして見た.*その際各時点に於ける推定値の標準偏差によつて，推定値の両側に巾をつけて見ると，比較的よく合う場合もあることが見られる．

しかしこの結果は，決してこれ迄の定常な時系列のモデルがそのまま有効であることを意味するものではない．推定式の作製に利用しているデータは，極く一部分のデータのみであつて，得られる限りのデータを全部使うことはしていない．この方法は，いわゆる consistency に頼る推定量の作製とはその考え方を異にしていることは明かである．

しばしば時系列の数式的模型によつて，無限に多くの時間的繰返しのデータによつて，あるパラメーターを推定することを考えるような consistency の証明等がなされているのであるが，既に述べたように，巨視的な立場から見る場合ならともかく，経済現象のように，外的條件の変動が著しく効くと見做されるような微視的接近に於て，同一條件の下に於ける無限に多くの時間的繰返しを考えるようなことは，全く非現実的な態度と見なくてはならない．ここでとつた解析の態度は，

* 計算については日本応用力学会編“応用統計学”第7章：小河原正巳：時系列論とその応用．参照．



variable な母集団の系列を想定し，その変動速度の考察に基いて，比較的容易に採用できる近似的な推定方式と考えられるものを利用し，そのような推定の系列的な効果の累積を大にしようというのであつて，いわば“局所的な推定”とでも言うべきものと考えられる．単にデータの数を多く使えばそれ程良いというような考え方ではない．このことについては §4 で改めて述べる．尚この点については，発散するような体系について，ある統計量の consistency を証明するような行き方と比較するとその差が極めて明かであろう.*

さて，一般に株価変動の解析といえば，直ちに，平均株価の推移を予測することを考えるわけであり，平均株価曲線の大きなまがり目を，できる限り早い時機に予知しようとするわけである．従つて誰しも当然各種の市場の模型を考え，その作用を荷重函数として表現し，投資家の投資可能資金等を入力とし，平均株価を出力とするような機構について計算をすることを考えるのであるが，如何なるデータを利用し，如何なる機構を考えるのがよいかを知るためには更に株価そのものの構造について知ることが必要である．単なる時系列模型のあてはめから一歩進んで，株価変動そのものの統計的解析という複雑な問題への接近を試みることにより，ひろく系列現象解析の際に於ける問題点を明かにし可能な限りそれを一般的な形で把握し，定式化することを試みなくてはならない．

§2.2 株価分析の結果

§§2.2.1 目標の明確化

一般に株式投資者の売買の態度を見ると，所謂，大勢，中勢，目先，というような分類がなされる．たとえば，大勢とは半年乃至1年，中勢とは1ヶ月乃至半年，目先とは1ヶ月以内を取引上(商略上）の時間單位として考えることを意味するわけである．これより考えると，大体実際的な行為としては，先づ半年乃至1年を單位として考える長期的な型と，日々の変動に注目して行動する型とを問題にすることが適当であろう．従つて，日々の株価変動の考察と同時に，その長期的な変化について見ることが必要である．

さて，現在利用し得られる資料の項目と，その時間的特性とを表示すると，一応次のようになる．

時間的特性 ＼ 項目	終 値	出来高	前日比	利廻り	225 種 平 均 株 価
略 々 日 單 位	○	○	○	○	○
〃 半 年 單 位					
以上より更に永い單位					

* 参照
T. C. Koopmans 編 Statistical inference in dynamic economic models の中の P. 356～P. 364.
H. Rubin: Csisistency of maximum-likelihood estimates in the explosive case.

時間的特性＼項目	資　本	配当率	利益率	流動比率	固定比率	固定資産回転率
略々日單位						
〃 半年單位	○	○	○	○	○	○
以上より更に永い單位						

再評価積立金	一株当り取引回数半年平均	株主分布	業　　種	
○	○	○		
			○	

（ここに掲げられた項目は，実際分析に際して，一応は検討された項目であつて，すべての項目を盡してはいない。）ここで我々は一般に株価の表現と考えられる 225 種平均について見ることによつて，株価の変動の大凡の型を知ることにする．（§1 の図参照）

すでに §1 で述べたように，t 日に於ける平均株価を x_t と表現すると，x_t が所謂定常な確率過程の実現値を表わしていると考えることは，この程度のデータをもとにする限り殆ど何等の知識をももたらすことにならないと考えられる．即ち，日を時間の單位とする考察に於ては，定常確率過程の理論を適用しようとすると，系列相関が極めて高く，この程度のデータでは数が全く不足と考えねばならない．

一般に自己相関係数が $\rho(t)$ で与えられる時データの総数が N ならば，相関関係の考察に際してランダムサンプルの場合に於けるサンプルサイズに相当するもの（有効観測回数）n は次式で与えられる．

$$n=\frac{N}{\tau}\ ;\qquad \tau=1+2\sum_{t=1}^{\infty}\rho^2(t)$$

現在のデータに関して n がどの程度であるかを見る為に，上掲のデータより得られる系列相関係数 $r^*(t)$ を見ると次の通りである．

$$r^*(10)=0.94 \qquad r^*(20)=0.91 \qquad r^*(30)=0.88\ .$$

試みに，$\rho(t)=e^{\frac{\log 0.94}{10}\cdot t}$ とおくと

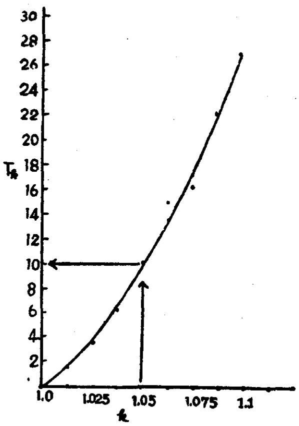
$$2\int_0^\infty \rho^2(t)\,dt \fallingdotseq 162$$

となる．300 項程度のデータでは不十分と考えなくてはなるまい．

このような変動の様相を示す現象の解析に際して，趨性変動（trend）の除去ということが一般に言われるのであるが，現在の例に於て trend を以て株価の平均的動向の表現と考えるならば，この trend を明確に解析することこそ，実は株価分析の最終的な段階を意味するわけであり，形式的に“趨性変動の除去”を試みた所で，その結果得られる所の知識が，何の意味を持つか全く不明となつてしまう危険がある．寧ろ上述の r^* の示唆する所に従つて，先づこの程度のデータについては年間の日々の株価全体が，ようやく一つの“株価”を表現するものと考えて，その動きと他の項目の動きとの関係を考察することから始めることが妥当であると考えられる．この立場は長期的な変動を考察する為に，半年或は 1 年の程度を解析の時間單位としようということを意味する．結局，先づ半年乃至 1 年間を解析の時間單位として，株価の構造を解析することが最初の目標とされるわけである．

§§2.2.2 標識の設定について

前§§の考察よりすると，現在得られている程度のデータ（約1年間分）では，平均株価の動きそのもののみの時間的変動の解析には全く不十分であることが分る．殊に，半年或は1年を以て，解析の時間単位とするならば，データは全く足りないことは明瞭である．ここで，時間的なデータの追跡より，空間的なデータを見ることが問題となつて来る．平均株価に関する知識をもとにして，各銘柄別の株価に着目しなくてはならない．即ち他の要因の空間的変動と株価のそれとの関係づけを試みるわけである．幸いにして，各会社は，大体半年毎に決算を行うので，§§2.2.1 に上げた各項目はそれを時間の単位として取ることにする．ここで，各銘柄の株価の表現としてその意味の上からも，解析に利用する際の便宜の点からも考えて，妥当なもの採用することが先づ第一の問題となる．他の各項目は，夫々各半年の決算報告によるものをそのまま用いることにして，それに対応する株価の指標を如何に設定するかというのである．ここで次のような方法を試みた．或る日の単純平均株価から見て，その k 倍，或は k^{-1} 倍になる迄に要する時間（日数）の平均的な値 T_k として



k	T_k（日）
1.0000	0.0
1.0125	1.5
1.0250	3.6
1.0375	6.3
1.0500	10.1
1.0625	15.1
1.0750	16.2
1.0875	22.1
1.1000	27.0

が得られる．

これを図示すると次の図のようになる．

この図から見ると，$k=1.05$ に対応する T_k が略々 10 日になることが分る．従つて株価の 1.05 倍，或は 1/1.05 倍程度の変化が意味を持つものと考えると（或は，我々の観測の精度として，$k=1.05$ をとると考えると）平均的に見て 10 日間程度の時間間隔を取つて，株価を見れば良いと言えよう．実は，株価の変動の型について，更に十分な研究がなされるならば，恐らく継続する各種の変動の群の連続と見做すことが出来るようになると思われる．即ち株価自体の固有の時間単位が，現在のような解析に対して現われることになると考えられるが，今は一応上のような考えに従つて，各銘柄については，データの 1/10 を系統的に抽出して，それを見ることにする．そこでたとえば，年間平均株価としては，この抽出されたデータの算術平均を以てするというわけである．

§§2.2.3 株価構造の解析

前§§迄を準備として，これから株価の構造について解析を試みる．実は，系列的な現象の解析に際しては，これ迄述べたところの準備的な段階が，最も重要な意味を持つのであつて，それが一応終れば後は比較的容易に解析をすすめることができるように思われる．

0\. 取上げる銘柄

昭和 27 年 12 月 1 日現在東京証券取引所上場銘柄（506 種）とする．

1\. データの等質性

極めて明瞭なことであるが，客観的な條件より明かに異質的であるとされるデータについては，十分に注意する必要がある．今の場合例えば次のようなことがある．

a. 決算期の相異

b. 業種の相異

a. については，3,9 月決算の会社が多い（286 銘柄）ので，先づこの型の銘柄について見ることにした. b. については，銀行保険，鉄道船舶，ガス電力，土地建物，貿易百貨店，興業，を除く他の銘柄を先づ見ることにした. これは実際に物を造つて売る型の会社と，そうでない型の会社とに分けたのである. 尚ガス電力は公益事業として，特殊な様相をもつものと考えて一応除き，後に他の会社の様相との比較を行おうと考えた.

結局このようにして，以後分析に用いられる銘柄は，3,9 月決算で上記以外の業種のもの 198 種となつたのであるが，日数に関して 1/10 のデータを抽出した時，部分的に上場されていない期間のあるもの 15 を除くと，183 種となる. ここで更に欠損会社 6 を一応除いて，177 種となつた. 以後の分析は専らこの 177 種について行われる. このようにデータの等質性に注意してはじめて，各種の関係が明瞭に把えられるようになつた.

2. 年間平均株価と他の項目との関係（1）

年間平均株価を x_0

利益率 （昭和 27 年9月期）を x_1

配当率 （ 〃 ）を x_2

一株当り正味資産（ 〃 ）を x_3 とする.

r_{ij} を x_i と x_j の相関係数

$R_{ij,k}$ を x_k の影響を考えた時の x_i と x_j の偏相関係数

$R_{ij,kl}$ を x_k, x_l, の影響を考えた時の x_i と x_j の偏相関係数

$R_{0,123}$ を x_0 と x_1, x_2, x_3, との重相関係数とする.

現在のデータによると次のようなことが分る.

$$r_{0.1}=0.641 \qquad r_{0.2}=0.507 \qquad r_{0.3}=0.376$$
$$r_{1.2}=0.663 \qquad r_{1.3}=0.487 \qquad r_{2.3}=0.393$$
$$R_{02,1}=0.170 \qquad R_{03.1}=0.106$$
$$R_{02.13}=0.133 \qquad R_{03,12}=0.080 \qquad R_{01.23}=0.434$$
$$R_{0.113}=0.653$$

各 x_i の β- 係数 a_i は

$$a_1=0.516 \qquad a_2=0.137 \qquad a_3=0.070$$

となる.

これらの結果より，このような表現方法に関する限り，利益率（x_1）が，平均株価の中心的な表現となつていることが一目瞭然となる.

この判断に従つて，利益率が株価はどのような関係にあるかを，もう少しくわしく見ることにする.

3. 利益率の変化の速さ

利益率の相対的な変化の程度を見る.

$r_{28,3:27,9}$ が，昭和 28 年 3 月の利益率と，昭和 27 年 9 月に於けるそれとの相関係数を表わすものとする. 他の記号も同様な意味を表わすものとする. 177 種について見ると，（欠損の場合には除去）

$$r_{28,3:27,9}=0.586 \qquad r_{28,3:27,3}=0.280 \qquad r_{28,3:26,9}=0.018 \qquad r_{28,3:26,3}=-0.013$$
$$r_{27,9:27,3}=0.577 \qquad r_{27,9:26,9}=0.233 \qquad r_{27,9:26,3}=0.018$$
$$r_{27,3:26,9}=0.688 \qquad r_{27,3:26;3}=0.233$$
$$r_{26,9:26,3}=0.589$$

昭和 26 年 9 月の利益率を考慮しての昭和 27 年 3 月の利益率と昭和 26 年3月のそれとの偏相関係数を見ると，

$$R_{27,3:26,3;\,26,9} = -0.294$$

同様にして

$$R_{28,3:26,3;\,27,3} = -0.084$$

更に同様な種々の係数が考えられる.

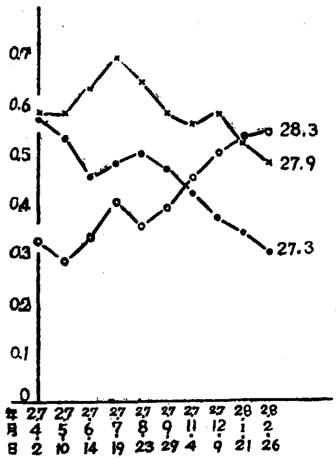
上掲の数字が比較的定安であることを考えると, これ等の数字の意味を考えることは, 相当役に立つことであろう.

4. 各決算期の利益率と日々の株価との関係

さて, 2, 3, に於ける数字より, 利益率と株価の関係が幾分明瞭となつたから, これと, 日々の株価との関係を見る. §§ 2.2.2 で抽出した 1/10 のデータの 1/3 の日数について見る. 日々の株価と, 昭和 27 年 3 月, 9 月, 28 年 3 月, の利益率との相関係数 (株価も利益率も対数で表現した場合) の変化の模様は次の通りである. この結果は投資者の相対的評価と, 利益率との関係を相当明瞭に表わしているといえよう. この関係は, これから先の分析に示唆するところが大であると考えられる. 即ちこのような関係をもとにして, 各種情報の流れと, 投資者の動きの関係の精密な調査分析を計画することが可能となる. この問題を更に解明して行くためには投資者の心理的な動きに関する考察が必要となつて来る.*

x 月の利益率と y 日の株価 (各々対数) の相関係数

y \ x 年 月 日	27 年 3 月	27 年 0 月	28 年 3 月
27. 4. 2	0.57	0.58	0.32
5. 10	0.53	0.58	0.28
6. 14	0.45	0.63	0.33
7. 19	0.48	0.69	0.40
8. 23	0.50	0.64	0.35
9. 29	0.47	0.58	0.39
11. 4	0.42	0.56	0.45
12. 9	0.37	0.58	0.50
28. 1. 21	0.34	0.52	0.53
2. 26	0.30	0.48	0.54



5. 年間平均株価と他の項目との関係 (2)

27 年 9 月のデータについて, 年間平均株価と, 利益率, 配当率, 一株当り正味資産, 流動比率, 固定資産回転率, 固定比率, 一株当り再評価積立金, 資本, との重相関関係を見ると次のようになつている. 先づ各項目の相関係数は次表の通りである.

	利益率	配当率	一株当り正味資産	流動比率	固定資産回転率	固定比率	一株当り再評価積立金	資本	年間平均株価
1. 利益率	1.00	0.81	0.49	0.27	0.21	0.23	0.27	0.16	0.62
2. 配当率		1.00	0.51	0.32	0.10	0.20	0.17	0.26	0.56
3. 一株当り正味資産			1.00	0.18	−0.17	−0.09	0.80	0.19	0.43
4. 流動比率				1.00	−0.08	0.49	0.16	0.08	0.31
5. 固定資産回転率					1.00	0.46	−0.28	−0.25	0.21
6. 固定比率						1.00	−0.10	−0.23	0.19
7. 一株当り再評価積立金							1.00	−0.01	0.18
8. 資本								1.00	0.05
9. 年間平均株価									1.00

* 参照 Katona, "Psychological analysis of economic behaviour", Mcgraw Hill, 1952.

各項目の年間平均株価に対する β-係数は次のようになる.

利益率	配当率	一株当り正味資産	流動比率	固定資産回転率	固定比率	一株当り再評価積立金	資本
0.486	0.076	0.527	0.220	0.116	−0.082	−0.371	−0.114

重相関係数は，約 0.70 である.

§3. 系列現象の統計的解析に関し一般的な結果

前§に於ける解析結果は極めて簡單明瞭であつて，それのみに注目するならば何等の困難も認められない．しかしながら，一般に極めて複雑と考えられる株価構造について，これ迄の結果を比較的短時日に獲得する為には，系列的な現象の解析に際して問題となる種々の困難に遭遇したわけである．その問題点について以下に述べる.

時間と共に変動する現象の統計的解析に際しては，既に種々の数学的モデルが考えられているのであるが，我々はここで改めて統計的解析の目標の特性を明確にすることによつて，現実の問題解析に際し注意を要する点を明かにしたい．統計的解析の目標の持つ時間的特性によつて分類すると大体次のようになる.

イ．時間の価値を全然考える必要が無い場合.

例えば，或る集団の構造を説明することのみを目的として，個々の要素の時間的な動きを考える場合.

ロ．時間の価値が積極的に問題となる場合.

社会現象の解析は，殆どすべてこの場合である.

ハ．上記両者の中間に位する場合.

有限時間を限定して，その中での現象の生起を問題にする時，等.

上記の三分類は極く粗雑なものであるが，これによつて次のことが分る.

即ち，時間と共に変動する現象を考察する際には，速度の概念を明確に持たなくてはならない．しかもこの速度の概念は，分析の対象の固有の速度と，我々がその対象に対して働きかける際に利用可能な種々の方法の固有の速度との比較に於てのみその意味を持つのであつて，結局解析に際して現実の問題の呈する種々の時間的特性は，我々のとる調査方法或は更に一般的な働きかけの速さと，対象そのものの変化の速さとの相対的な関係によつて定まつて来るものと考えられ，上記三分類はこの意味で明瞭なものである．單に観念的に費用の計算に時間の項を入れて，そのウエイトの変化として考えるならば，上記の三分類は殆どその意味を持たないのであるが，ウエイトの変化は單なる数量的な変化のみを意味するものではなく，現実の解析の態度の全面的変更を意味するものとなるのであつて，上記のような分類はその意味で有効なものであろう．これ等の考察は，既存の数式的モデル適用の前に先づなされなくてはならないものであつて，統計的解析に際し最も重要な段階であると考えられる.

以上の点に留意しつつ，分析の方向を振返つて見ると，知識の獲得のために百数十枚の相関図その他の図表が作製されていたのであるが，それを時間的順序に従つて見ることによつて，知識の領域の拡張の仕方の一例が得られるように考えられる．その示唆するところによると，結局時間と共に変動する現象の統計的解析に際しては，次の諸点が最も重要であると考えられる.

1. 時間單位（解析上の）設定の問題
2. 標識設定（表現）の問題
3. 標本設定の問題

1. は，時間と共に変動する現象の解析に於て特に意味を持つものである．特に計算集計の費用の増大と，それによつて得られる知識の有効性の増大との關係が問題になるわけであつて，系列的な

現象に於て共通の問題たる数字化或は量子化の問題である．この問題に対しては，

1. 実際的な行為からの要求を明確にすること，行為の時間的特性を明確にすること．
2. 利用可能な情報の供給源を，その時間的特性に従つて分類すること．
3. 既得のデータにより対象の時間的特性を出来る限り明確にすること．
4. 1,2,3 各項の要素の時間的特性の比較を行うこと．
5. 費用（時間を含む）を考えること．

が常に必要である．

2. は一般に統計的解析の成否を定める問題であつて，その決定には解析目標の十分な理解が要求される．特に時間と共に変動する現象の解析に於て注意すべき点は，知識の増大を最初から唯一の標識の系列について得られる時間的繰返しだけに頼ることをしないで，可能な限り，同種類の多数の系列と併せて見ることである．即ち，單に時間的繰返しだけでなく，空間的な繰返しによつて知識の獲得を試みることが大切である．このことは全く当然なことであるが，一つのモデルを頭に置いて現象に接近する際に，往々忘却され易い点でもある．これ迄にとり上げられている多くの数学的モデルの場合のように，時間の価値を考えていない場合には，多く定常確率過程として空間的集団と，その中の一点の時間的変動の様相との関係づけは，エルゴード性によつてなされていたのであつて，両者を区別して考える必要はないように定式化されているのであるが，このようなモデルを，実際に一つの系列のみを追跡する場合に形式的に適用することには問題があると考えられるし，* 我々が時間の価値を積極的に問題とする際には，集団の時間的変動に対する考察のより高い精度が要求されることになるから，單純に時間に無関係に一定な構造を持つ集団を想定することは妥当でない．集団の構造の時間的変化と，個々の系列の時間的変化の相対速度を考えたり，或は一般的に変化して行く集団の構造中に時間と共に変化しないものを見出すことが，最も重要な問題となるわけである．**

多次元的な系列として考察することの重要性は，またそのような考察によつて，系列の一時点に於ける空間的特性と，時間的特性との関係づけを考えることが出来るようになる点にある．この点は時間と共に変動する現象を，時間の価値を積極的に考えて解析しようとする際に局所的な推定の問題に関聯して極めて重大な意義を持つ．この点を簡單に数式的に表現することを後章で試みる．

尙本報告に用いられた解析例の多くに於て，最も重要な統計量は相関係数であつたが，従来の計算機によると相関係数の計算には多くの時間が必要とされ，従つて，多次元的な系列の統計的処理に際しては，極めて多くの時間が要求される．この方向の解析法を有効なものとする為には，是非とも相関係数の算出其の他の計算を迅速確実に行う計算機の利用を考えなくてはならない．本報告に於ける解析例の示す極めて重要な系列現象解析に於ける問題点の一つは“相関係数の計算を迅速確実化するような計算機の利用を考えること”であろう．この点については目下その実現に努力中である．統計的方法の適用分野の拡張のためには，このような努力が是非必要である．

3. については，時間と共に変動する現象の解析に際して特に注意を要する点である．時間と共に変動する現象を微視的に追跡する際は，観測の結果にもとづいて一定の傾向を見出す為には，相当な精度を要求されるので，少しでもちらばりを増大させるような原因はこれを除去しなくてはならない．従つて，サンプルによつて追跡する際には，特にこの点に留意して，得られる限りの知識と

* この点については，実際，問題となることが多いものと考えられる．Symposium on informatin theory の report の中で，D. Gabor が Communicationtheory and physics という題の下で少しくこの点にふれている．

** このような問題の一つの数式化として，Neyman 及び Scott による“structural parameter”“incidental parameter”の考えがある．尙 J. Wolfowitz : Estimation by the minimum distance method, Annals of the institute of statistical mathematics Vol. V, No. 1 の §5. The problem of sstimation of structural parameters がある．

過去のデータを利用することによつて通常のサンプリングに於ける層別の場合以上に嚴密な分類が試みられなくてはならない．この点に関聯して，2. の表現の問題の一部として，系列的なデータの多数を，如何なる表現を利用することによつて有効な分類を行うことが出来るかという問題が生起する．即ち，2. について上述したように，系列現象の解析に際しては，その多次元的な表現を考慮することが重要であるが，更にこの時間的空間的ひろがりを同時に考察することが問題となる．一次元の系列から多次元の系列へ，或は更に系列の多次元表示へと進むのであるが，形式的な表現は容易であつても，具体的な問題に対する系列の多次元表示の方法は，今後の問題である．この方法の発展なくしては，3. の問題の十分な解明はできないのであるがこれに関連して次の §4 に於て型の認識に関係する問題をとり上げて述べて見る．

§4.　定式化への試み

これ迄にのべられた問題について，その数式的表現を最も簡単な例について述べる．これらの例に於ける考え方を更に発展させ，複雑な統計量に迄及ぼすことが今後の問題と考えられる．ここに述べる例はその考え方に於ては，全く従来中心的に論じられている数式的モデルとは異つている．この方向の研究を更に押進めるつもりである．

1.　次のような問題を考えて見る．

X_t を平均 m_t，分散 σ^2，なる確率変数とする．但し $m_t=t\cdot\delta$ 各 X_t は互に独立とする．

この時　$\overline{X}\equiv\dfrac{1}{n}\sum\limits_{t=1}^{n}X_t$　によつて，X_{n+1} を推定することを試みる．

$\overline{X}$ の平均は $\{(n+1)/2\}\delta$，分散は σ^2/n となる．

$\overline{X}$ と X_{n+1} との差の二乗の平均値を見ると，

$$E(X_{n+1}-\overline{X})^2=\frac{\sigma^2}{n}+\left(\frac{n+1}{2}\right)^2\cdot\delta^2+\sigma^2$$

$\sigma=k\cdot\delta$ なる場合には

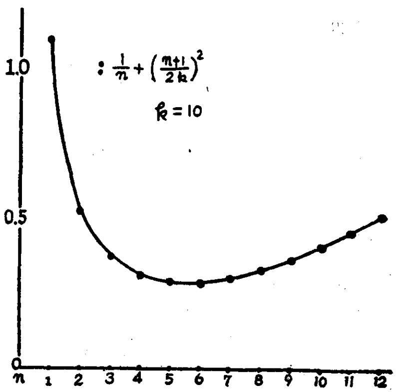
$$E(X_{n+1}-\overline{X})^2=\left\{\frac{1}{n}+\left(\frac{n+1}{2k}\right)^2\right\}\sigma^2+\sigma^2$$

右辺の第一項に注目すると，

$$x^2(x+1)=2k^2$$

なる x に近い n，（$[x]$，又は $[x]+1$）で最小になることがわかる．

$k=10$ なる場合について $1/n+(n+1/2k)^2$ を図示して見る．



これより明かなように，母集団の変化を考える場合でも，その変化の速さによつては，母集団の変化が無いものと考える場合に適当とされる統計量が相当有効に利用出来，且その時には，ある意味で局所的な有効性しか持たないことがわかる．この簡単な例を以て，先に述べた時間の単位の問題，及び局所的な推定の問題の定式化への試みと考えることが出来よう．

2.　多次元的な系列として考察することの重要性を次のような例について見よう．

比較的短時間（少数の t）について次のような関係が成立するものと考えられる場合がある．

$$n_t=a\cdot m_t+b+\varepsilon_t\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\text{(A)}$$

$$X_i^{(t)}=m_t+\xi_i^{(t)}$$
$$Y_i^{(t)}=n_t+\eta_i^{(t)}$$

ここに　$\varepsilon_t, \xi_i^{(t)}, \eta_i^{(t)}$ は

$$E(\varepsilon_t)=0, \quad \sigma^2(\varepsilon_t)=\sigma_\varepsilon^2$$
$$E(\xi_i^{(t)})=0 \quad \sigma^2(\xi_i^{(t)})=\sigma_{\xi t}^2 \qquad \rho(\xi_i^{(t)}\cdot\eta_i^{(t)})=\rho$$
$$E(\eta_i^{(t)})=0 \quad \sigma^2(\eta_i^{(t)})=\sigma_{\eta t}^2$$

なる確率変数で，ε_t は t が異なれば独立，$(\xi_i^{(t)}, \eta_i^{(t)})$ は i が異なれば独立なるものとする．

このようなモデルを考えると

$$Y_i^{(t)}-aX_i^{(t)}-b-\varepsilon_t=\eta_i^{(t)}-a\xi_i^{(t)}$$

となる．$i=1,2,\cdots n$，とすると，最小2乗法を適用して a の推定値として

$$\hat{a}=\frac{\sum_{i=1}^{n}(X_i^{(t)}-\bar{X}_t)(Y_i^{(t)}-\bar{Y}_t)}{\sum_{i=1}^{n}(X_i^{(t)}-\bar{X}_t)^2}$$

を得る．

この結果は，経済に於ける各種の弾力性係数の測定の際に於ける空間的（同時的）測定と，時間的（逐次的）測定の間の関係を示すものとも考えられよう．モデル（A）が比較的短い時間しか成立しない場合，このような接近法によらなくては a を推定することは出来にくいであろう．しかし，このモデルの有効性については更に検討を要する．

3.　先に利益率と株価の相関係数の時間的推移の型を見て，それから利益率と投資家の評価との関係を推定したが，系列的な現象の解析に於ては，このような型を見ることによつて各種の知識を得る場合が多い．そこでこのような問題をもう少しはつきりした形にして，系列的な現象解析に際して，一般的に有効な知識としたい．この問題は，結局型の認識に関係して来るわけで，心理学的な考察が要求されることになるのであるが，ここでは一応極めて単純な一つの試みとして次のように考えて見る．即ち，一般に型を見るのは，系列的な現象の実現値を x_t とするとき，t が diacrete な場合には

$$\Delta x_t,\ \Delta^2 x_t, \cdots,\ \Delta^n x_t, \cdots \quad .$$

t が continuous な場合には

$$\frac{dx_t}{dt},\ \frac{d^2x_t}{dt^2}, \cdots,\ \frac{d^nx_t}{dt^n}\cdots .$$

等の継続的な様相によると考える．

このような型で一番簡単なものは，例えば

$$\Delta_1 x_1 = x_2 - x_1 > 0$$

なる型を確認しようとする場合であろう．そこで次のような問題を考える，電球の纖條の温度と，電流量との関係を見る場合に，多数の電球について，電流量 i_1 なる時の纖條の温度 T_1，電流量 i_2 なる時の纖條の温度 T_2 を測つて一般に $T_2>T_1$ なる関係を見ようとする時次の二つの方法が考えられよう．その一つは，電流 i_h なる時の温度 T_h が，平均 $\bar{T}_h$，分散 σ_h^2，なる確率変数をなしていると考え，各 i_h に於て，ランダムサンプルのサンプル平均 $\bar{t}_h$ によつて $\bar{T}_h$ を推定し，その推定値 $\bar{t}_h$ の差より $\bar{T}_2-\bar{T}_1$ を見ようと考える場合であり，他の一つは，ランダムサンプルたる各電球（k）について，$T_2^{(k)}-T_1^{(k)}$ を計り，そのサンプル平均 $\overline{\Delta t}=\bar{t}_2^*-\bar{t}_1^*$ によつて $\bar{T}_2-\bar{T}_1$ を推定しようと考える場合である．この場合直ちに分るように，$T_2^{(k)}$，$T_1^{(k)}$，の相関の様相によつて両者

の推定の精度の比が定まる.

T_2, T_1, の相関係数を ρ とすると,

$$D^2(\bar{t}_2-\bar{t}_1)=\frac{\sigma_2{}^2}{n^2}+\frac{\sigma_1{}^2}{n_1}$$

$$D^2(\overline{\Delta t})=\frac{1}{n}\cdot(\sigma_2{}^2+\sigma_1{}^2-2\rho\sigma_1\cdot\sigma_2)$$

$$D^2(\bar{t}_2)=\frac{\sigma_2{}^2}{n^2}\ ,\qquad D^2(\bar{t}_1)=\frac{\sigma_1{}^2}{n_1}$$

$$D^2(\bar{t}_2{}^*)=\frac{\sigma^2}{n}\ ,\qquad D^2(\bar{t}_1{}^*)=\frac{\sigma^2}{n}\ .$$

ここで次のような場合がある.

実際与えられたデータは, i_1 で N 箇, i_2 で N 箇であるが, その中 i_1, i_2, 共同一の電球によるものが n 箇ある. このような時 $\bar{T}_2-\bar{T}_1$ の推定をどのように行うのが良いかというのである. 全部のデータのサンプル平均を $\bar{t}_1$, $\bar{t}_2$ とし, 一の電球について測定された組の中のデータのサンプル平均を $\bar{t}_1{}^*$, $\bar{t}_2{}^*$, とすると, 夫々の場合にデータがランダムサンプルによつて得られたものと考えると,

$$D^2(\bar{t}_2-\bar{t}_1)=\frac{1}{N}\left(\sigma_1{}^2+\sigma_2{}^2-\frac{n}{N}2\rho\sigma_1\sigma_2\right)$$

$$D^2(\bar{t}_2{}^*-\bar{t}_1{}^*)=\frac{1}{n}(\sigma_1{}^2+\sigma_2{}^2-2\rho\sigma_1\sigma_2)$$

今特に $\sigma_1{}^2=\sigma_2{}^2=\sigma^2$ と考えると

$$D^2(\bar{t}_2-\bar{t}_1)=\frac{2\sigma^2}{N}\left(1-\frac{n}{N}\rho\right)$$

$$D^2(\bar{t}_2{}^*-\bar{t}_1{}^*)=\frac{2\sigma^2}{n}(1-\rho)$$

$D^2(\bar{t}_2-\bar{t}_1)$ と $D^2(\bar{t}_2{}^*-\bar{t}_1{}^*)$ との大小を見ると

$$\begin{aligned}D^2(\bar{t}_2-\bar{t}_1)-D^2(\bar{t}_2{}^*-\bar{t}_1{}^*)&=\frac{2\sigma^2}{Nn}\left(n-\frac{n^2}{N}\rho-N+N\rho\right)\\&=\frac{2\sigma^2}{Nn}\left(1-\frac{n}{N}\right)\{\rho(N+n)-N\}\\&=\frac{2\sigma^2}{Nn}\left(1-\frac{n}{N}\right)\{n\rho-N(1-\rho)\}\end{aligned}$$

従つて,

$$\frac{\rho}{1-\rho}\geq\frac{N}{n}\rightleftarrows D^2(\bar{t}_2-\bar{t}_1)\geq D^2(\bar{t}_2{}^*-\bar{t}_1{}^*)\ .$$

これより見ると $\rho>0.5$ なる場合には注意を要する. 例えば, $\rho=0.6$ とすると

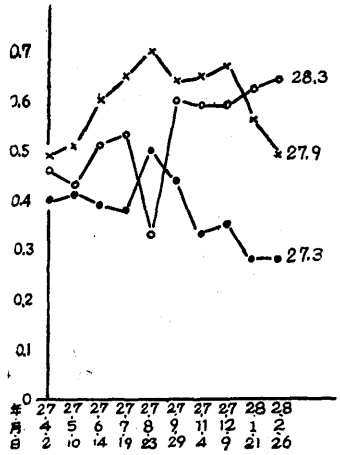
$$N>1.5n$$

でない限り, 後者の方が推定の精度が良い. 時間と共に変動する現象, 或は更に一般的な系列現象の解析に於ては ρ の大なるものが多い. そのような場合のデータを処理して一定の型を見出そうとする時に注意しなくてはならない点であると考える. 利益率と株価の相関係数の推移を見る場合

についても次のような例がある．各日に於ける株価と，利益率との相関係数の算出に用いるデータを一定の銘柄に定めずに，取引の有無に従つて，算出に用いる銘柄を変えた場合には次のような結果になつた．

x 月の利益率と y 日の株価の相関係数

x ＼ y 年　月　日	27 年 3 月	27 年 9 月	28 年 3 月
27.　4.　2	0.40	0.49	0.46
5. 10	0.41	0.51	0.43
6. 14	0.39	0.60	0.51
7. 19	0.38	0.65	0.53
8. 23	0.50	0.70	0.33
9. 29	0.44	0.64	0.60
11.　4	0.33	0.65	0.59
12.　9	0.35	0.67	0.59
28.　1. 21	0.28	0.56	0.62
2. 26	0.28	0.49	0.64



この計算の際には，両項目ともにそのままの数であつて，対数をとつてはいない．従つて数字の性質が，前掲のものとは異るから，全く同一視することは出来ないが，この計算に用いられたデータ数の方が各点に於て前掲の場合より必ず多くなつているにもかかわらず，前掲のものに比して，遙に不十分と考えられる結果しか示さないのは，表現方法と同時に，恐らく銘柄を一定にしないということの上述したような影響が表われたものと考えられよう．これ等の結果から，系列現象解析の計画に際して，一般に有効と考えられる事柄として，

“一定の関係を見出すためには，系列現象の解析に於ては，特に一貫したデータの利用について慎重でなくてはならない．系列相関が相当高いときには特に一定のサンプルの追跡に留意しなくてはならない，多くの場合，部分的に余分なデータを取入れることは損失を招くことになりやすい”

と言えるであろう．而して，社会現象の多くに於ては，その変化の速度は比較的遅く，従つて，系列相関が，通常の解析の時間単位の程度では，相当高いものが多いと考えられる．そのような調査に於て，精密な知識を獲得しようとする際には特にこの点の注意が必要であろう．

統計数理研究所