



## Osaka Gakuin University Repository

Title	社会科学における相関関係の考え方に関する考察 How to Evaluate Correlations in Social Science
Author(s)	加茂 英司 (Eiji Kamo)
Citation	大阪学院大学 商・経営学論集 (OSAKA GAKUIN UNIVERSITY REVIEW OF COMMERCE AND BUSINESS ADMINISTRATION), 第 43 巻第 2 号 : 33-45
Issue Date	2018.03.31
Resource Type	ARTICLE/ 論説
Resource Version	
URL	
Right	
Additional Information	

# 社会科学における相関関係の考え方に関する考察

加 茂 英 司

## How to Evaluate Correlations in Social Science

Eiji Kamo

### ABSTRACT

In social science many variables are in un-stationary state. Those variables have distinctive characters. Correlations in time-series data and cross-sectional data are different from each other; correlated in time-series data, but uncorrelated in cross-sectional data, and uncorrelated in time-series data, but correlated in cross-sectional data. Furthermore, it is necessary to use monthly data, instead of yearly data, in order to know whether two variables are exactly correlated or not. These characters are explained with three sets of sample variables.

## 目 次

- 第1節 非定常性と疑似相関の違い
- 第2節 非定常性のメカニズム
- 第3節 自殺率は景気循環に反応しているのか

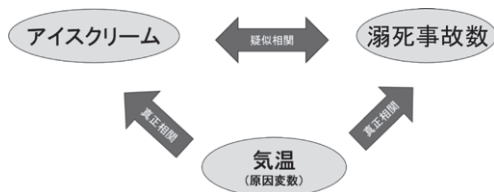
## 第1節 非定常性と疑似相関の違い

### 非定常は時系列では相関、クロスセクション分析では無相関

ウィキペディアには疑似相関という項目がある。そこには「アイスクリームの売り上げ」（以下、単に「アイスクリーム」）と「プールの溺死事故数」（以下、単に「溺死事故数」）の事例を疑似相関として紹介している。

夏になると、アイスクリームの売り上げも、そして溺死事故数も増加する。反対に冬になるとアイスクリームの売り上げも溺死事故数も減少する。もちろん両変数には関係性は無い。関係性が無いのに同じように変化するということから、人々は両変数が気温を原因変数とした疑似相関であると考えることになる。

図表1 アイスクリームと溺死事故数の疑似相関のイメージ<sup>1)</sup>



しかしアイスクリームと溺死事故数は、ともに同じ時系列変化を持つ「非定常性変数」である。非定常性というのはそれぞれの変数が時系列的に自律的な変化を持つ性格のことを意味している。同じ時系列変化を持てば、たとえ関係の無い変数であっても時系列的に相関関係を持つ。反対に、異なる時系列変化を持てば、たとえ関係性のある変数間であっても、時系列的に無相関に変化することになる。

アイスクリームと溺死事故数の持つ季節変動は、冬から夏にかけては上昇トレンドを持つ。温度が高くなるにつれて加算されるが、その程度は互いに無相

1) 著者作成

閑だということが特徴である。夏から冬にかけては両変数とも下降トレンドを持つ。減算されることは同じでも、その程度は無相関なのである。

この性格の興味深いことは、時系列データでしかその特徴が表れないということにある。実際、アイスクリームと溺死事故数のクロスセクションデータでは相関関係は生じない。

図表2は国別のアイスクリーム消費量を表したものである。

最もそれが多いのは、ニュージーランド、アメリカ、オーストラリアなど酪農の盛んな国である。またフィンランド、スウェーデン、デンマークなど北欧の国でアイスクリーム消費量が多いのも、やはり酪農がさかんだからである。その国の気温がアイスクリーム消費量に関係の無いことを示している。

図表2 国別アイスクリーム消費量<sup>2)</sup>

国	国民1人当消費量(リットル)
ニュージーランド	26.3
アメリカ合衆国	18.7
オーストラリア	17.8
フィンランド	13.9
スウェーデン	11.9
カナダ	9.5
イタリア	9.2
アイルランド	9.0
デンマーク	8.7
イギリス	7.7
チリ	5.6
マレーシア	2.0
中国	1.9
日本	0.01

2) [http://www.umai7.com/aisu/2006/09/post\\_3.html](http://www.umai7.com/aisu/2006/09/post_3.html)

総じて乳製品を食べる歴史の浅いアジア諸国では気温に関係なく、消費量は少ない。常夏のマレーシアではさぞアイスクリームをたくさん食べているかと思えばそうでもなく、世界でも最もアイスクリーム消費量の少ない国の一つである。これもまたアイスクリーム消費量と気温に関係の無いことを示している。

そのため、当該国の平均気温別にアイスクリーム消費量をクロスセクションデータで見れば、両変数は無相関になる。

そもそも疑似相関は、疑似とはいえ列記とした相関関係のあることを意味している。しかしアイスクリーム消費量と溺死事故数には相関関係は存在しない。存在するのは、単に同じ時系列変化を持つことによって、上昇トレンドと下降トレンドのタイミングが同じだということに過ぎないのである。

次に紹介する事例はれっきとした疑似相関である。疑似相関なので、クロスセクションデータでは相関関係が生じていることに特徴がある。ところが両変数とも異なる時系列変化を持つ非正常性変数なので、時系列変化では無相関になることに注意してもらいたい。

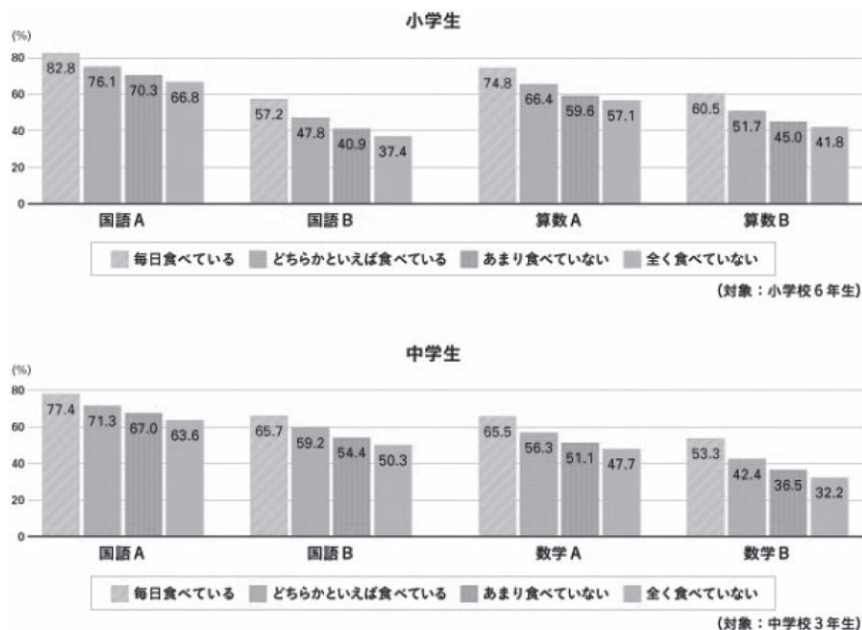
### 朝食摂取率と成績

「朝食摂取率」と「成績」が強い結びつきを持つことは経験的にわかっている。そこで全国の小中学校ではアンケート調査が良く行われることになる。

図表3の調査結果は、朝食摂取率別に個々の生徒の成績を調べたものである。各科目の左から「毎日食べている」「どちらかといえば食べている」「あまり食べていない」「全く食べていない」、つまり朝食摂取率の高いグループから低いグループに並べている。

朝食摂取率の高いグループでは成績が良く、それが低いグループになるほど成績も悪くなる。まさに朝食摂取率と成績には正の相関関係があるという根拠になっている。

もちろん朝食を食べるだけで成績が上がるというのはにわかには信じられな

図表3 朝食摂取率別成績<sup>3)</sup>

い。つまり私たちは、両変数が疑似相関であることを経験的に知っている。

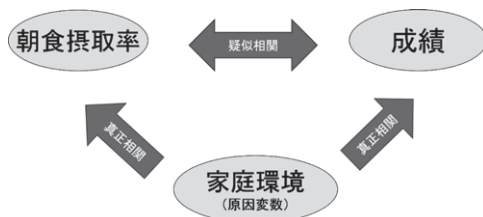
良好な家庭環境の子弟であるほど朝食摂取率も高く、成績も良い。反対に劣悪な家庭環境の子弟ほど朝食摂取率も低く、成績も悪い。誰かが家庭環境を指数化したわけではないから、この関係を証明できるわけでもないが、経験的に多くの人は疑似相関だと考えているのである。

朝食摂取率と成績は両変数とも非定常性変数であり、異なる時系列変化を持っている。

図表5のグラフは朝食欠食率の推移を表したものである。欠食率は摂取率の対義語であって、欠食率が高いほど摂取率が低いことを表している。欠食率は

3) <https://papimami.jp/7830/>

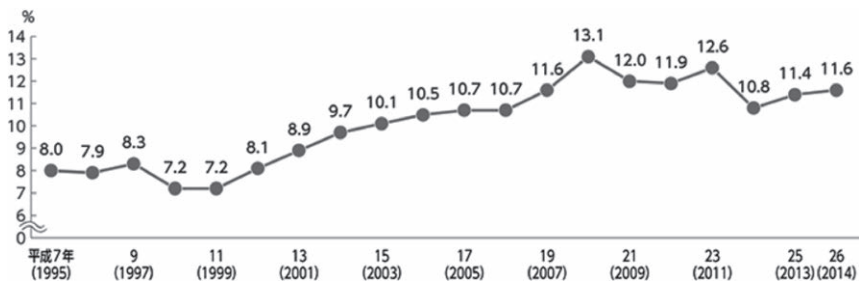
図表4 アイスクリームと溺死事故数の疑似相関のイメージ<sup>4)</sup>



一貫して上昇トレンドを持っていることから、朝食摂取率は一貫して下降トレンドを持つことを示している。

ところが成績はといえば、ゆとり教育が導入されれば成績は下降トレンドを持ち、脱ゆとり教育が始まれば再び成績は上昇トレンドを持つ。このようにそもそもトレンドが異なるために、時系列的には両変数は無相関に変化することになるのである。

図表5 朝食欠食率<sup>5)</sup>



4) 著者作成

5) [http://www.maff.go.jp/j/wpaper/w\\_maff/h27/h27\\_h/trend/part1/chap1/c1\\_3\\_01.html](http://www.maff.go.jp/j/wpaper/w_maff/h27/h27_h/trend/part1/chap1/c1_3_01.html)



## 第2節 非定常性のメカニズム

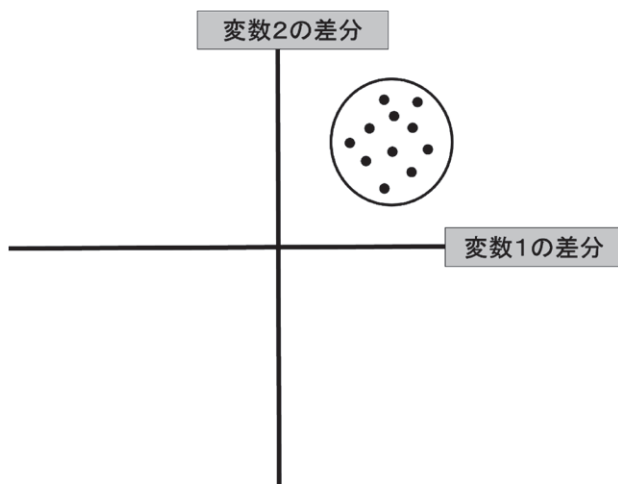
アイスクリームと溺死事故数のように関係性も無く、また疑似相関でもないのにどうして時系列データでは相関関係が生じるのか。朝食摂取率と成績のように関係性が有るのに、どうして時系列データでは無相関になるのか。

非定常性変数に関するこれらの不思議なメカニズムを理解するキーワードは「差分」である。

差分とは、ひとつ前の期の値と今期の値との差異のことである。変数1の差分と変数2の差分の分布図を描いたので、それを見てほしい。

図表6は関係性の無い二つの変数の差分の分布を表したものである。差分の分布が無相関であるのは両変数に関係性のないことを示している。

図表6 関係性の無い変数の差分の分布図<sup>6)</sup>



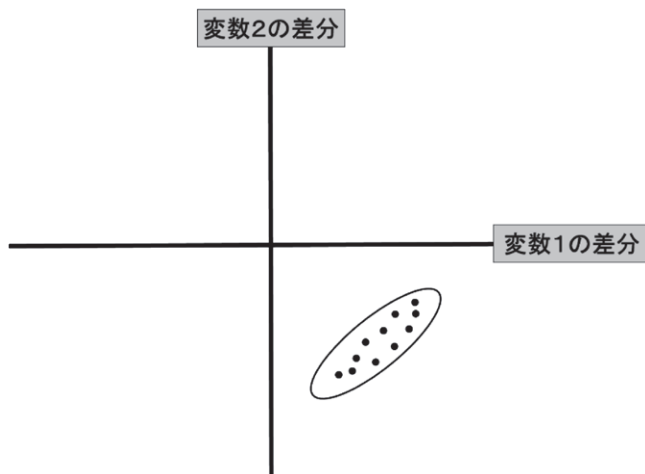
6) 著者作成

しかしその分布の符号を見てほしい。変数1の差分は左がマイナス、そして右がプラス。また変数2の差分は下がマイナス、そして上がプラスである。図はどちらの変数の差分もプラスの値をとっていることを示している。この場合、変数1も変数2も上昇トレンドを持つ。つまり時系列データでは両変数とも右肩上がりであり、そのために同じ時系列変化を持つことになる。これが無関係なのに同じ時系列変化を持つメカニズムである。

次の図は正の相関関係のある二つの変数の差分の分布である。変数1の差分はマイナス、変数2の差分はマイナスである。変数1は下降トレンド、そして変数2は上昇トレンドを持つことを意味している。つまり両変数は異なる時系列変化を持つことになる。

これが、関係性があるのに異なる時系列変化を持つメカニズムである。

図表7 関係性のある変数の差分の分布図<sup>7)</sup>



7) 著者作成

### 第3節 自殺率は景気循環に反応しているのか

これまでの検証でわかったことは、変数間の関係性を検証しても非定常性変数がどのような時系列変化を持っているかを確認することはできないということである。つまりひとつひとつの変数の時系列変化はそれ自体を検証しなければならぬということなのである。

では私たちはどれくらい正確に時系列変化を検証しているのだろうか。その興味深い事例に自殺率の時系列データがある。

#### 長期トレンドが景気循環にあわない自殺率はどうして？

自殺率が景気循環に反応するかどうかについて、様々な議論があった。

図表8は自殺率が循環サイクルを持つと主張する根拠になっているグラフであり、自殺率と失業率の時系列データを見たものである。一般に循環サイクルを持つ変数は、60年代には右肩上がりに上昇し、対照的に90年代には右肩下がりに低下している。

もちろん失業率は景気循環とは逆向きに反応するので、60年代には右肩下がりに低下し、90年代には右肩上がりに上昇をする。その意味では、失業率はきわめて景気循環に適切に反応している。

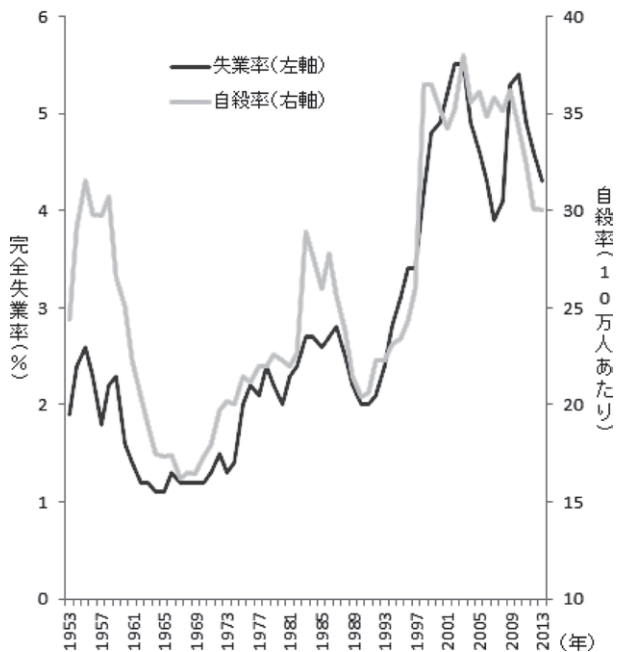
自殺率とはいえば、やはり失業率と同じように60年代には右肩下がりに低下し、90年代には右肩上がりに上昇をしている。したがって自殺率もまた景気循環に反応していると考えることができる。

しかしこれは単純に10万人あたりの自殺者数を計算した数字に過ぎない。

自殺率はきわめて年齢に依存する数字である。

10歳未満の子供では、自殺率はきわめて低いことは言うまでもない。年齢が高くなるにしたがって自殺率が高くなり、中高年ではきわめて高い水準になる。したがって、自殺率に変化がなくても、中高年の人口比が高くなるにした

図表 8 失業率と自殺率<sup>8)</sup>  
男性の失業率と自殺率の推移



資料：総務省『労働力調査』、厚生省『人口動態統計』  
作成者：舞田敏彦 (@tmaita77)

がって、自動的に全体の自殺率も高くなってしまふ。

そこであらためて人口構成の調整を行った標準化自殺率というデータを見る必要がある。図表 9 は標準化自殺率の推移を表したものであるが、戦後の一時期こそ高い水準にあったものの、高度経済成長期の到来とともに劇的に低下している事が特徴である。その後、バブル期に一時的に低下するとはいえ、むしろ安定して推移していることに着目してもらいたい。

8) [http://tmaita77.blogspot.jp/2015/02/blog-post\\_6.html](http://tmaita77.blogspot.jp/2015/02/blog-post_6.html)

つまり失業率は長く安定して推移し、90年代になってようやく大きく変化をし始めた。対照的に自殺率のほうは、戦後から高度経済成長期にかけて大きく低下して、その後は長期にわたって安定して推移している。自殺率と失業率が互いに独立して変化している以上、自殺率は景気循環に「反応していない」ように見える。

図表9 標準化自殺率の推移<sup>9)</sup>



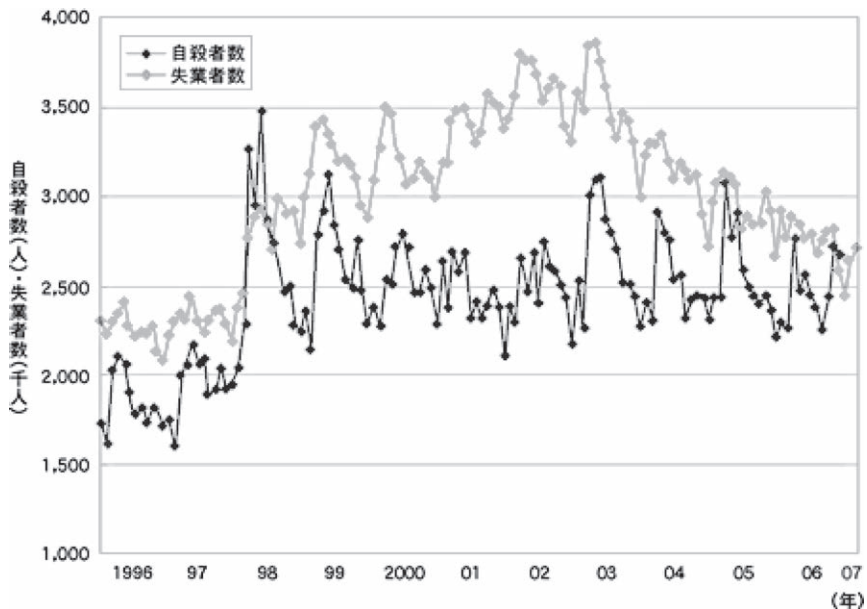
これまでに見てきたデータは全て「年次データ」である。ところが景気循環に反応しているかどうかを判断するためには「月次データ」を見る必要がある。

図表10は、月次データで見た失業率と自殺率である。失業率はもちろん景気循環に反応していることがわかっているが、自殺率も同じように反応していることがわかる。つまり両変数はともに同じ時系列変化をもっていることがわか

9) <http://blogs.itmedia.co.jp/sakamoto/2013/10/sales-tax-suicide.html>

るグラフである。反対から見れば、景気循環に反応していることは月次データを見なければ、わからないということでもある。

図表10 月次データで見た場合の失業者数と自殺者数<sup>10)</sup>



出所：厚生労働省『労働力調査』『人口動態統計』（2006年以降は概数）

10) <https://research.nttcoms.com/database/data/000585/>