

歯科疫学統計  
- 第4報 傾向（トレンド）分析の基礎 -  
瀧 口 徹

A review of oral epidemiological statistics  
- Part IV : The bases of Trend Analysis -  
Toru Takiguchi

要旨：第4報では傾向（トレンド）分析を扱った。本分析は統計学の世界では珍しく目的と手法が多義的なので利用に際して混乱し易いと思われる。そこで次の4つの視点から傾向（トレンド）分析が何たるかを概説した。1. 時系列データの不規則変動を除去して傾向の把握と予測性に重点を置いた分析、2. 最適回帰曲線算出に重点を置いた分析、3. 時系列データの類型化に重点を置いた分析、および4. 量-反応的要因分析に重点を置いた分析、である。事例として①歯科用金属としても重要な金、パラジウムなどの貴金属相場の時系列データ（上記分析1の事例）、②年齢と現在歯数の関係を示す3種類のデータ（分析2の事例）、③タイ国のデング出血熱の20年間の発生率と気象条件との生物学的関連モデルの事例（分析2の事例）、④スリランカ国の20年間の感染性、非感染性疾病発生率の推移データ（分析3の事例）および⑤老人ホームにおける味覚の減退と要介護度との関連データ（分析4の事例）を用いた。本報ではこれらの事例を用いて各種傾向（トレンド）分析法をどのような場合にどのようにして利用したらいいかのノウハウを解説した。

キーワード：傾向（トレンド）分析、時系列分析、移動平均法、コクラン・アーミテージ・トレンド検定、共分散分析（ANCOVA）

はじめに

本シリーズの第1報から3報において各種統計分布に基づくより厳格な統計解析について実戦的視点から概説してきた。第4報においては傾向（トレンド）分析の基礎的手法を紹介する。傾向（トレンド）分析は元々相場の時系列分析由来で

ある。すなわち相場の上昇、下降、停滞および周期性等の数学的な解釈とそれに基づく将来予測から構成されている。しかし現在の傾向（トレンド）分析は現在複数の分野において全く異なった目的と手法で行われているにもかかわらず同じ傾向（トレンド）分析と呼んでいる。例えば保健・医療の分野では要因分析や量-反応曲線（Dose-Response Curve）を示す事象の解析も傾向（トレンド）分析と呼ぶことがある。このように極めて専門的な用語なのにもかかわらず統計の世界では珍しく多義的で混乱し易い。そこで今回下記の4つの視点から多義的な傾向（トレンド）分析を類型化しその統計的基礎を概説する。

【著者連絡先】

〒341-0003 埼玉県三郷市彦成3-86  
深井保健科学研究所  
主席研究員 瀧口 徹  
TEL&FAX : 048-957-3315  
E-mail : taki8020@mth.biglobe.ne.jp

1. 時系列データの不規則変動を除去して傾向の把握と予測性に重点を置いた分析
2. 最適回帰曲線算出に重点を置いた分析
3. 時系列データの類型化に重点を置いた分析
4. 量一反应的要因分析に重点を置いた分析

1. 時系列データの不規則変動を除去して傾向の把握と予測性に重点を置いた分析

1) 事例のプロフィール

(1) 貴金属価格の推移

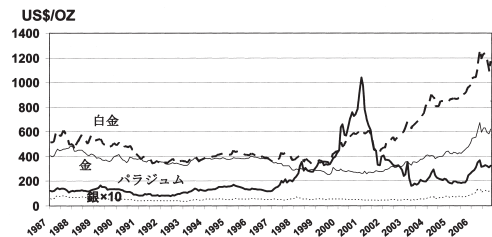
最初の事例は歯科医療に関連が強い貴金属の価格の時系列変動である。歯科関係者であれば数年前(1998-2002年頃)に代表的な特定保険医療材料である歯科用金銀パラジウム合金(以下、金パラ)が急騰<sup>1)</sup>したため採算性の問題から臨床現場が混乱した記憶が新しいであろう。本稿では1987年1月から2006年11月までの約20年間の貴金属素材価格の推移(時系列データ)に焦点を当て、移動平均法の機能と方法間の特徴の違いを示す。

2) 分析手法

○移動平均法(Moving Average)

移動平均法(Moving Average、以下MA法)は時系列データの不規則な上下動、すなわち、ゆらぎ(fluctuation)を除去して平滑化をはかり上昇、下降、停滞および循環等の長期的傾向(トレンド)を視覚的に捕らえる統計手法(web-site WS01)である。1960年代初めに米国のJoseph granville が相場分析の指標として導入したのが最初で、以来、どの時点をどのように重視するかということから様々な手法が開発されてきている。代表的な手法(WS02)は①単純移動平均法(SMA)、②加重移動平均法(WMA)、③三角形移動平均法(TMA)、④指数平滑移動平均法(EMA)および⑤時系列移動平均法(TSMA)である。この他にも様々な移動平均法が開発されているがその変化の原因となる事象と構造変化を推測し将来予測をすることが目的であることはどれも同じである。まず図1(KITCO社 WS03)の金、白金、パラジウム、および銀の単位価格の約20年間の月別推移をみていただきたい。これをみ

ると1998年から約5年の間パラジウムの価格は金を越え、さらに1年半ほどは(パラジウムは酸化防止作用金属という用途からみて白金の代用金属なのにもかかわらず)白金の価格をも凌駕していたことがわかる。しかしながら4種の金属の急騰急落期間以外の任意の比較的短期間(この例の場合は2,3年)をみた場合、そのトレンドは上昇傾向にあるのか、あるいは下降傾向にあるか判定しづらい場合が多いことに気づくであろう。その理由は不規則な変動(ゆらぎ)があることによる。図2は4つの金属のうち金およびパラジウムの1987年1月から2006年11月までの239月の時系列データに対して前述の5つの方法の①、②および⑤のMA法を適用、すなわち単純移動平均法(SMA: Simple MA)、加重移動平均法(WMA: Weighted MA)および時系列移動平均法(TSMA: Time Series MA)の3つのMA法を適用させた結果を示している。ここで、それぞれのMA法を適用するには予め集計単位とする期間を決めなければならない。そこで前二者は計算された平均値の位置を集計期間のちょうど中央にするために通常奇数(3または5日、月、年)を採用する。本例では図2の傾向から3年に相当する35月(奇数)単位を採用した。これに対してTSMA法は前34月のデータから回帰直線式を求め、その式のxに最後の35月を代入して35月目を予測する方法をとった。比較の変動幅が小さい金価格の推移に対して3つのMA法はいずれも視覚的に高い適合性を示している。これにより金の価



Note:

- 1) US\$/OZ: 1金衡オンス(31g)当たり米ドル価格
- 2) 銀×10: 銀は単位価格が他の貴金属と比較して一桁廉価のため同一図に掲載するために10掛けしている。

図1 4種の貴金属価格の推移 (ロンドン市場1987.1月~2006.11月)

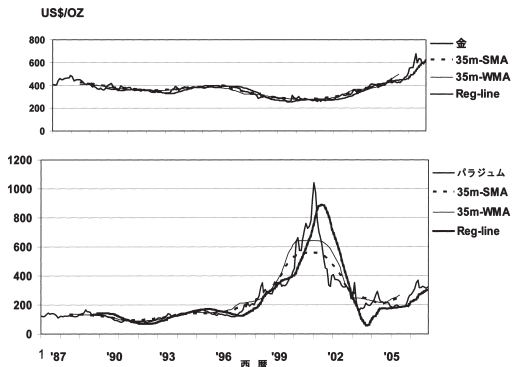
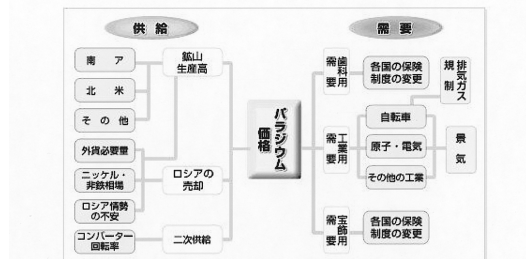


図2 金とパラジウム価格傾向(トレンド)各種分析法(移動平均法)の比較

格変動は単位価格400US\$/OZ(1オンス(=31g)当たりの米ドル価格)の前後で数年間隔の緩やかな上下動を繰り返しており、ここ2、3年は比較的急勾配で上昇してきていると読める。一方、パラジウムの場合3つのMA法は視覚的な敏感度からみてTSMA(Reg-line)、WMA、SMAの順で1998-2002年の急騰、急落現象を捕らえている。SMAは移動平均値の前後のデータを均等に扱っているが、WMAは新しいデータほど重み付けを強くしているため図2においても急騰期への適合がSMAより良くなっている。しかし、TSMAの感度は飛び抜けて高くパラジウム価格のような急騰、急落の時系列データのトレンド分析に向いていると判断される。この理由は、前2者は一定の時間幅を右に1単位だけ年を動かしデータ値を1減1増して得られる平均値のみの情報から算出されているため時系列データの凸凹の平滑化にのみ寄与するからである。一方、TSMAは1単位だけ右に移動していく方法はSMA、WMAと同じだが共分散(=相関)の情報によって右肩上がりか、下がりか方向と強度(直線性)を時々刻々予測していくため上昇、下降の変曲点を視覚的に敏感に捕らえることができるためと考えられる。

さて、ここまでの議論は全て過去の傾向を読み取る手法の話であったが、それでは将来予測についてはどう考えればいいのかであろうか。時系列データは過去の軌跡だから過去の傾向はわかっても未来が見えるはずがないのか、それともある程



出典:オリент貿易 <http://orikura.jp/futures/investent/palladium/change.html>

図3 パラジウム価格変動の基礎的要因

度は見えるのか、しかし当たり前はずれのリスクが大きく危険な選択を強いられるのか、緻密な時系列データを取り複雑かつ精密な解析を駆使すれば未来がかなり正しく見えるのか、いずれが正しいか見解がわかれるところである。図3はパラジウムを扱っているオリент貿易(WS04)社がインターネット上に公表しているパラジウム価格変動の基礎的要因図である。ここでパラジウムの生産量は圧倒的にロシアが多く、消費量は日本が最大でロシアの売却戦略に大きく影響されているがその他にも複雑な需要供給関係があると同社は説明している。従って各需給要因の重さの将来の変化を考慮しないで今までの数学的傾向のみで単純に未来のパラジウム価格を予測するのは危険である。冒頭で説明したように傾向(トレンド)分析はその生い立ちからすれば相場の将来予測をするために生まれ発達してきた。しかし将来予測となると経験や専門的な勘といった判断も合わせて必要となり、そうした判断をした場合は論理性、客観性、再現性等の狭義の科学性が保障されなくなるであろう。従って公共的な施策決定において重要なのは将来いつ大きな変動が起きるかという根拠が不確定な、いわゆる予想ではなく「その前兆を掴むこと」および「過去に起きた規模の変動と同様な変動がいつ起きても対応できるようにすること」が重要となる。特に後者は意志決定理論である minimax 戦略<sup>2)</sup>の立場からみて合理的であろう。過去に起きた大変動が同様な規模(あるいは少し上回る規模)で将来起きた場合を想定した対応は、日本が2回の中東オイル危機から学んだ石油備蓄と脱化石燃料化の国策、あるいは大地震ご

とに強化する方向で見直される建築基準等がこのいい例であろう。また歯科界においても例えば、歯科材料成分としての金パラ価格に関してもその考え方が既に導入されている。特定保険医療材料の中核的存在である金パラ価格の極端な相場変動によって歯科医療需要と無関係な理由により売買が行われ、もって歯科保険医療の質を低下させる方向に働く可能性をなくすため価格の高騰時に余り間を置かず実勢価格との差を埋める公的介入が行われている。本稿のテーマに戻っていえばこの種の分析では共分散情報を生かすことが出来る時系列移動平均法(TSMA)の利用価値が高いということを強調しておきたい。

2. 最適回帰曲線算出に重点を置いた分析

1) 事例のプロフィール

(1) 一般的な回帰曲線の事例

事例は表1および図4に示す年齢と現在歯数の関係を示す3種類の分布(group1-3 仮想データ)

表1 年齢と平均現在歯数(仮想データ)

年齢	group1	group2	group3
20	29	31	29
25	30	30	31
30	26	29	20
35	24	28	14
40	26	29	6
45	20	24	8
50	22	23	14
55	19	24	4
60	9	27	4
65	15	25	1
70	11	22	3
75	9	24	5
80	3	20	2
85	2	16	0
90	1	5	0

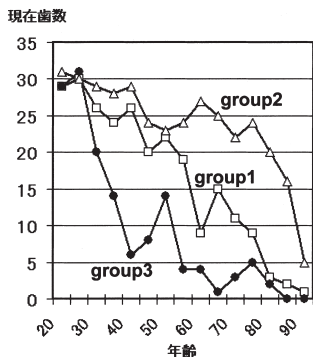


図4 年齢と平均現在歯数(仮想データ)

タ)である。これらをSPSSの「分析-回帰-曲線推定」の6つのモデル(線形、対数、逆数、二次、三次および指数)を用いて回帰式を推定し、それぞれの適合度を比較し最適の回帰曲線を選択するための基本的方法を提示する。

(2) 生物学的モデルに基づく、より複雑な回帰曲線

タイ国の Dengue (出血) 熱発生の時系列分析を事例として扱う。Dengue (出血) 熱: Dengue Haemorrhagic Fever: DHF (以下、Dengue熱)<sup>3)</sup> はマラリア、日本脳炎等とともに蚊を媒介(Vector borne) とするワクチンも有効な治療薬もまだ開発されていないウイルス性疾患であり、特に亜熱帯、熱帯に蔓延し、流行の周期性があつて時に猛威を奮う。症状は高熱に始まり、時に意識障害を引き起こし、血小板の著しい減少が原因の出血を伴う重症例では死に至る例も希ではない。本事例のタイ国<sup>4)</sup> では1987年に大流行し174,285人(人口10万人比325人)が罹患し、うち1,007人が死亡している。この疾患を伝播する熱帯蚊(Aedes aegypti, Aedes albopictus)の繁殖は温度や降雨などの気象条件に大きく左右されるとされているが予見に役立つまでには定量化されていない。そこでタイ政府保健省は全国73省における1978年から1997年の20年間のDengue熱発生に関して大規模な国家レベルの傾向(トレンド)分析<sup>4)</sup> を行い2005年に公表した。この結果を「より複雑な回帰曲線」についての事例として紹介する。

2) 分析法

最適回帰曲線を求める方法は時系列データに限らず断面データの傾向分析をする場合も適用され、事象を従属変数と独立変数の関係で多項式(一次式、多項式)、指数関数などの数式で表すことを目的とした分析である。最適曲線を求めるに当たって注意しなければならない点は次の3つである。

① (生物学的な) 仮説の必要性: ある分布に最適な回帰曲線を求める場合、単に数学的な適合のみを希求せず、なぜそうなるかという妥当性のある(生物学的な) 仮説に基づく分布の前提が必要で



ある。例えば年齢とある生理的現象の2次元プロットデータを数式で表す場合、仮にデータの実在する範囲で凹型の二次関数が最適回帰直線だと採用された場合、最右端のデータ値のさらに右部分の高年齢でデータが取れたとしたらそれは上昇に転ずる生物学的妥当性があるかどうか問題である。本事例では現在歯が高齢になると逆に増加に転ずるという摩訶不思議なことは起こりえず凹型の二次関数の生物学的妥当性は無い。一方例えばアルコール量と急性心疾患の関係<sup>5)</sup>のような場合はいわゆるU型あるいはJ型であることが知られており、このような場合には凹型の二次関数は生物学的妥当性を持ち得る。

②過剰適合の制御：多項式の次元を上げていった場合や指数関数、三角関数等と組み合わせた方程式を採用すると数式は見かけ上細かな fluctuation にも限りなく適合していく。しかし有益な独立性の高い情報とともに冗長な情報やノイズ情報も持ち込むことになる。この長所、短所を調整する方法をどうするか。

①誤差分布型の確認：ここで言う分布はデータ値の分布ではなく実測値と推定値の差、すなわち誤差分布である。近年この誤差分布を従来のように正規分布と固定せず最尤法<sup>6)</sup>の手法を用いて他の分布も検討し、より厳密な分布の選択を目指す研究指針もある。しかしデータ値に内部相関が無い場合はランダム抽出であるから誤差分布は正規分布となりその前提での解析が現在でも主流である。しかし時に対数正規分布や別の分布を示すことがある。この場合計算を用意するために用いられている尤度式の  $-2 \log$  最大尤度<sup>6)</sup> は正規分布以外の分布にも使われる。なお、誤差分布を固定して考えない立場でのモデル化の極致が第2報で紹介したGLIM法<sup>7), 8)</sup>でありこれを駆使することによりサンプルの誤差分布として最適の分布が選択される。

(1) 一般的な回帰曲線

表1および図1に示す年齢と現在歯数の関係を示す3種類の分布にそれぞれ最も適合する分布型を求めるためSPSS (12.0J) の「分析-回帰-曲

表2 6モデル式の適合度比較

説明変数の数:k	group					
	線形 linear	対数 logarithm	逆数 inverse	2次 quadratic	3次 cubic	指数 exponent
R <sup>2</sup>	0.9388	0.8656	0.7240	0.9467	0.9480	0.7557
R <sup>2</sup> の順位(降順)	3	4	6	2	1	5
モデル適合度F検定	p=0.0000	p=0.0000	p=0.0001	p=0.0000	p=0.0000	p=0.0000
-2Log Likelihood	69.75	81.56	92.35	67.69	67.30	104.28
AIC	73.75	85.56	96.35	73.69	75.30	108.28
AICの順位(昇順)	2	4	5	1	3	6
R <sup>2</sup>	0.6889	0.5908	0.4731	0.7796	0.9034	0.6166
R <sup>2</sup> の順位(降順)	3	5	6	2	1	4
モデル適合度F検定	p=0.0001	p=0.0004	p=0.0024	p=0.0001	p=0.0000	p=0.0031
-2Log Likelihood	81.64	85.76	89.56	76.44	64.39	88.77
AIC	85.64	89.76	93.56	82.44	72.39	82.77
AICの順位(昇順)	3	4	6	2	1	5
R <sup>2</sup>	0.7450	0.8519	0.8884	0.8796	0.9005	x
R <sup>2</sup> の順位(降順)	5	4	2	3	1	x
モデル適合度F検定	p=0.0000	p=0.0000	p=0.0000	p=0.0000	p=0.0000	p=0.0000
-2Log Likelihood	91.57	83.42	79.03	80.31	77.46	x
AIC	95.57	87.42	83.03	86.31	85.46	x
AICの順位(昇順)	5	4	1	3	2	x

Note:  
 1) -2Log Likelihood: -2×最大対数尤度 (第3報 p45-47参照)  
 2) AIC: Akaike's Information Criterion: 赤池の情報量基準(規準) (第3報 p45-47参照)  
 3) Groupのx: データ値に0があるため計算不能

線推定」のモデルのうち視覚的に近似すると思われる6つのモデル(線形、対数、逆数、二次、三次および指数)を用いて回帰式を推定した。それぞれの適合度を表2に示す。ここでモデルの適合度(goodness of fit)を表す指標は重相関係数: R<sup>2</sup>、回帰式のF検定、およびAICである。一つの目安としてR<sup>2</sup>が0.5以上と大きいこと(=寄与率50%以上)は実測値と最小二乗法で求められた回帰モデルによる推定値間の誤差が小さい、すなわち適合が高いことを示している。表2に示すgroup 1の解析では6モデルの全てでR<sup>2</sup>は0.72-0.95と高く、F値もp=0.0001~0.0000と高度に有意で、これらの適合度指標はいずれのモデルに対しても使用可のゴーサインを出している。しかし、ベストはどれかという問いにはこれらの適合度指標は残念ながら明確に答えてはくれない。どれでもいいなら最もシンプルな直線式がいいという見方、実測値と推定値間の相関係数のべき乗すなわち寄与率: R<sup>2</sup>が最も大きい3次式がいいという見方、それぞれに一理あって収拾がつかない。このことを解決してくれるのが赤池情報量基準(規準): Akaike's Information Criterion: AICである。

$$AIC = -2 \times \log(\text{最大尤度}) + 2 \times k$$

(k: 自由パラメータ数 = -2最大対数尤度 + 2k (or (k+1)))

ただし最小二乗法を使う場合は右辺は (k+1)

このAICは誤差分散が正規分布の場合に限らず各種分布に基づくモデルの適合度を評価できる極めて汎用性の高い方法で、データ値から母数(母

集団を表す平均値、分散等)を推定する最尤法と適切な情報量の概念を組み合わせたものがある。データ値からすなわち左項の最大対数尤度で最も適切な(最尤の)母数を求めつつも右項で変数を過剰に使うことによる情報のダブリやノイズの増加を調整し適切な情報量としたものである。なお、AICの詳細は第3報<sup>6)</sup>等を参照されたい。ここで表2に戻っていただいてgroup 2においては3次式が、group 3においては逆数式においてAICが最小となり最適なモデルということになる。group 2においては最大R<sup>2</sup>が最小AICであるが、group 1およびgroup 3においてはR<sup>2</sup>が最大でもAICはそれぞれ3位と2位で最小でないことに注目いただきたい。図5、6、および7は3つのgroupにそれぞれ最も適合がいい回帰曲線を示しており本法の妥当性が視覚的にも確認いただけると思う。

(2) 生物学的モデルに基づく、より複雑な回帰曲線

次の事例はプロフェールで紹介したデング熱に関するタイ国の研究<sup>4)</sup>であり、前述した最適な回帰式を求めるときの注意点である(生物学的な)仮説を重視したモデルである。当該研究では気候要因の重さを計量するため下記の3つのモデルが

デング(出血)熱発生変動モデル モデルA=定数+時系列傾向 モデルB=定数+時系列傾向+循環効果+残差 ただし、モデルBの図は掲載省略 モデルC=定数+時系列傾向+循環効果+気候要因+残差
--

比較検討された。

本研究は気候要因である降雨量と最高気温の月変動の影響に焦点を合わせたものである。この分析は当然決定的な関係がある高温多雨の熱帯、亜熱帯の気候がデング熱の発生に関係があるかないかを議論しているのではなく、降雨量や気温が例年の変動幅を超えて上昇した場合デング熱の流行を警鐘する必要があるかを見極めようとしたものである。その結果を図8に示したがここにはモデ

ルAおよびCの結果のみ図示してある。モデルAの場合緩やかな凸型の曲線となるが規則的な上下変動を全く説明できていないのは明らかである。これに対してモデルCの場合余弦式(cos θ)と温度変化の不規則要因を加味することで規則的循環のかなりの部分を説明できるようになっている

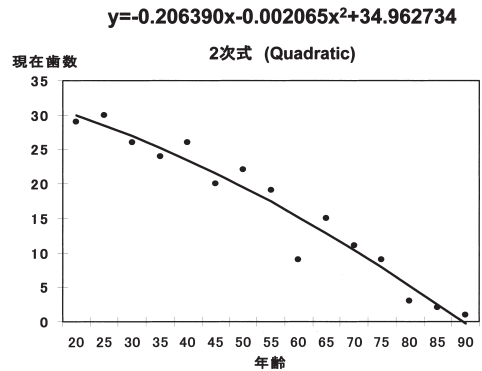


図5 group 1最適回帰曲線

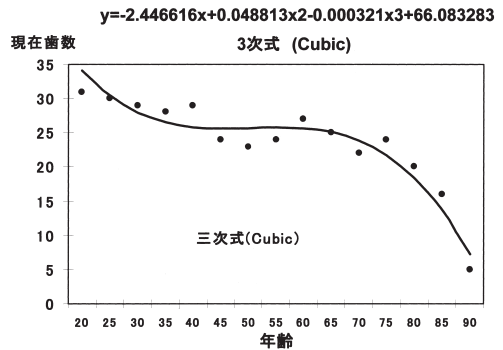


図6 group 2最適回帰曲線

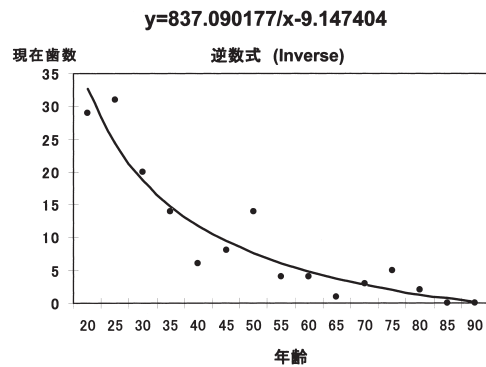


図7 group 3最適回帰曲線

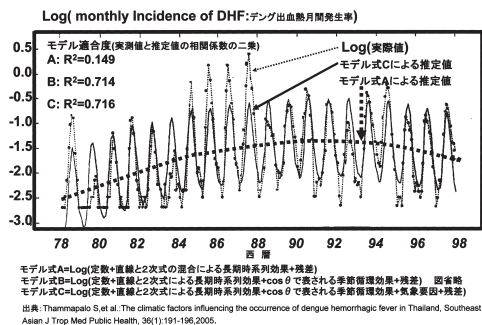


図8 タイ国20年間のデング出血熱の発生状況と傾向  
モデル式

のがわかる。実測値とモデルによる推定値の単相関係数： $R^2$  ( $r^2$ ではないことに注意) はモデルA、B、Cそれぞれ0.149、0.714および0.716であった。つまりモデルの適合度という視点からだモデルCに新たに取り入れた気象要因は $R^2$ を0.02しか増加させずモデル適合性に対する貢献度はかなり小さい。すなわちモデルBで足りるという結論となる。従って本テーマを時系列データへの回帰線の適合問題としてのみみた場合は気象条件をパラメータから外したモデルBが最適ということで話は完結する。少なくともタイ国においては長期的時系列傾向と同一年内の循環効果が決定的な要因で、全国一律でみるとデング熱の発生率の予測精度に降雨量増加と気温の上昇がほとんど貢献しない。しかしこのモデルからタイ国のデング熱対策は気象条件の変化を考えた対応が必要ないと結論づけてはいない。本研究は時系列データの回帰曲線で表す最適モデルのパラメータに気象の変化要因が採用されなかったからといって要因として軽視することには誤りであることを同時に示して興味深い。すなわち、降雨量と最高気温の変化に関して各省ごと、あるいは省の集合体である地域 (region) ごとに検証すると興味深い結果が出てきた。それは降雨量の増加と気温の上昇の影響が逆であるという点である。降雨の影響については降雨量が1mm増えた場合の影響はかえってデング熱の発生を下げるという結果がでている省が散見 (5/73省) された。一方気温が $1^{\circ}\text{C}$ 上昇した場合は8省においてデング熱発生率は有意に増加

していた。ごく一部の省 (8/73) が有意なものにもかかわらず国全体を地理的な4地域区分中でみると半分の2地域で明らかに温度の上昇に反応してデング熱の発生率が有意に増加している。このため通年の範囲を超える過剰な降雨があったからといってそれがデング熱の急増と関係しないので殊更の注意喚起は必要ないが、温度が上昇した場合注意喚起が必要と結論付けられている。この理由は先行研究によると気温が上昇すると原因となる蚊の生存期間、繁殖力、小型化による飛翔力増加などが生じてデング熱が急増する危険があるためとされる。他の研究でも温度上昇はデング熱流行の重要な要因とされていて日本人としては地球温暖化による同疾患流行域の北上を懸念させる。ただ降雨量が増えたときデング熱の発生率がなぜ低下するかの理由は「大雨だとボウフラが洗い流されてしまうから」など、もっともらしいが巷の推測の域をでていないようである。本事例は全体では影響が小さくなるか相殺されてしまうけれども、ある集団においては同じ方向で足並みが揃って従属事象に影響しているという隠れた要因 (本例の場合は地域差) がありうることを教えてくれる。なお、こうした隠れた要因は統計的には一般に層別解析か重回帰分析の場合偏相関係数で検出できるが本事例のように多項式や正弦 (余弦) 式、指数関数等を組み合わせた回帰直線の適合を調べる場合、比較的小さい要因を見逃さないためには層別解析を併用することが現実的で望ましい。

### 3. 時系列データの類型化に重点を置いた分析

#### 1) 事例のプロフィール

##### (1) スリランカ国の各種疾患発生率の時系列データ

近未来の保健・医療政策の方向性を考える上で疾病構造の変化は重要な情報である。そこで筆者が現在JICA政策顧問として勤務しているスリランカ保健省が公表<sup>9)</sup>している感染性、非感染性疾患59種のうち代表的疾患の1980-2003年の発生率 (入院患者としての) のデータを増減傾向および変動の傾向をもとに疾病を類型化するための

事例とした。

2) 分析法

分析法を表3に示す。59疾患は西暦：xと入院有病者数：y (Inpatient Morbidity) との間の単回帰式の係数：aを用いて3群に分けられた。さらに個々の疾患ごとに回帰直線で説明できない不規則変動 (fluctuation) を回帰直線からの残差平方和である  $(1 - r^2)$  値を用いた100  $(1 - r^2)$  で表し表3に示す基準で5段階の水準 (最大5、最小1) を決めた。なお、fluctuationは辞書の一般的定義としては①不規則な変化、②波のように (undulate) 上下降する動き、とされまた医学においては脳波、心拍、呼吸、血圧、末梢血流量等の生体信号における一見不規則な変動を指し「生体ゆらぎ」と呼ばれ時系列データにカオス理論を導入した解析<sup>10)</sup> 等が試みられてきている。しかしここで用いた不規則変動 (fluctuation) の指標は上記のように相関分析で用いられる寄与率 $r^2$ の逆を用いた単純で扱いやすいものとした。これは回帰直線で表せない分散で、誤差分散と周期変動 (もしあったとしたら) 等が合わさったものであ

表3 スリランカにおける疾病別入院患者の推移 - 1980~2003年の傾向の類型化 -

データ出典: Medical Statistics Unit, Health Bulletin, Sri Lanka, 2003
1) 対象疾患: 公表されている59疾患は年間入院患者総数を人口10万人あたりに補正した数値。増加傾向疾患は全て表示、その他の停滞、減少傾向を示す疾患は代表15例のみ表示。
2) <b>NCDs</b> : 非感染性疾患: Non-Communicable Diseases
3) <b>I</b> : 調査年数 (1980-2003)
4) (疾患の) 順位: 単相関係数 $r$ を降順にソート
5) 傾向(トレンド): +++: 係数: $a > 0$ and $p < 0.001$ ,    ▼▼▼: 係数: $a < 0$ and $p < 0.001$ ++: 係数: $a > 0$ and $p < 0.01$ ,    ▼▼: 係数: $a < 0$ and $p < 0.01$ +: 係数: $a > 0$ and $p < 0.05$ ,    ▼: 係数: $a < 0$ and $p < 0.05$
6) 変動: <b>Fluctuation</b> の指標化: $100 \times (1 - r^2) \geq 80.0 \rightarrow 5$ $79.9-50.0 \rightarrow 4$ $49.9-30.0 \rightarrow 3$ $29.9-10.0 \rightarrow 2$ $< 10.0 \rightarrow 1$ ただし、 $r^2$ は単相関係数の二乗値

る。表4-aは明らかな増加傾向を示した15疾患を示し、表4-bには減少傾向もしくは不規則変動4、5を示した30疾患を相関係数の大ききで降順に並べて示している。結果をまとめると下記の5つとなる。

- 〈1〉 59疾患のうち15疾患が統計学的に明らかな有病者数の増加傾向を示したが、特にHIV/AIDSは $r=0.98$ というほとんど直線の増加傾向 (fluctuation=1) を示した。
- 〈2〉 増加傾向を示した15疾患のうち8疾患は非

表4-a 増加傾向を示す15疾患

順位	疾患名	感染・非感染	r	100*(1-r <sup>2</sup> )	傾向 Trend	変動 Fluctuation
1	HIV/AIDS		0.98	3.5	++	1
2	喘息	NCD	0.97	6.7	+++	1
3	肝疾患 (非感染性)	NCD	0.96	8.4	+++	1
4	循環器系疾患 (IHD: 虚血性心疾患性を含む)	NCD	0.95	9.9	+++	1
5	神経系疾患	NCD	0.94	10.8	+++	2
6	悪性新生	NCD	0.94	11.0	+++	2
7	泌尿・生殖器系疾患		0.94	11.7	+++	2
8	高血圧性疾患	NCD	0.94	12.0	+++	2
9	傷害、中毒およびその他の外的要因由来疾患		0.94	11.2	+++	2
10	先天性奇形、形成不全および染色体異常	NCD	0.93	13.3	+++	2
11	周産期の不調		0.92	14.9	+++	2
12	糖尿病	NCD	0.88	23.0	++	2
13	消化器(GIT diseases)		0.85	27.3	++	2
14	妊娠、出産および産褥期の疾病		0.81	33.8	+	3
15	敗血症		0.72	47.9	+	3



表4-b 停滞、減少傾向を示す疾患15例

順位	疾患名	感染・非感染	r	100*(1-r <sup>2</sup> )	傾向 Trend	変動 Fluctuation
16	デング(出血)熱		0.53	71.7		4
17	結核		0.50	75.1		4
18	呼吸器系疾患		0.41	83.4		5
19	精神神経疾患	NCD	0.33	89.0		5
20	細菌性髄膜炎		-0.05	99.8		5
21	ポリオ		-0.45	79.9		4
22	狂犬病		-0.45	79.6		4
23	ウイルス性肝炎		-0.64	59.4		4
24	マラリア		-0.67	54.9		4
25	フィラリア		-0.70	51.4	▼	4
26	ジフテリア		-0.82	32.7	▼▼	3
27	栄養不良	NCD	-0.90	19.6	▼▼	2
28	日本脳炎		-0.90	18.2	▼▼	2
29	貧血	NCD	-0.98	4.8	▼▼▼	1
30	百日ぜき		-0.99	2.2	▼▼▼	1

感染性疾患：NCDsである。

- 〈3〉NCDsのうち特に循環器疾患、ガン、高血圧、糖尿病などのいわゆる生活習慣病（LRDs: Lifestyle-Related-Diseases）はこの20年間で直線的に増加してきている。
- 〈4〉蚊を媒介とする伝染病のうちフィラリア、日本脳炎はこの20年間で明らかに減少しているが、マラリアの場合相関係数：rも回帰直線の係数：aも統計学的に僅かに有意水準に達しておらず(p=0.07 and 0.06)、その理由は特に1987年の時点で猛威を奮った等不規則変動が大きい（Fluctuation=4）ことに起因している。しかし近年は事実上根絶に近い。一方、デング熱の場合は上記を含め蚊を媒介とする4つの代表的な病気の中で唯一有意に近い増加傾向を示すがやはり年ごとの不規則変動が大きい。
- 〈5〉百日咳と貧血は相関係数：rがそれぞれ-0.99, -0.98とほとんど直線的減少傾向を示している。前者ワクチンの普及、後者は経済状態の好転が影響した栄養状態の改善がかかわっていると考察される。

以上、時系列データの類型化に重点を置いた分析には一般に複数の事象が増加か減少かあるいは停滞かという単純な分類が必要になる。直線回帰以外の方法を使うと適合は良くなるが2次曲線の場合は「減少し、再び上昇（またはこれらの逆）」、3次関数の場合は「減少し、停滞しまた減少（またはこれらの逆）」、指数関数の場合は「最初急激に上昇、その後緩やかに上昇（またはこれらの逆）」等となり類型化が単純でなくなる。両方法の長短は議論のあるところであるが単相関係数の絶対値が0.7以上になる場合であれば単純な増減分類を用いる方が利点が多いであろう。

#### 4. 量-反応的要因分析に重点を置いた分析

##### 1) 事例のプロフィール

##### (1) 要介護度と味覚の関係

味覚と要介護度の関係をみるため某老人ホーム入居者約200名を対象とし、6段階の要介護度判定（平成11年5月時点の判定表使用）と現在歯数の検査、味覚にかかわる項目、身長、体重、基礎疾患等について聞き取り調査を行ったデータを事

例とした。

2) 分析法

単要因傾向分析  
 コクラン・アーミテージ検定（トレンド検定）  
 多要因傾向分析  
 共分散分析：ANCOVA：Analysis of Covariance

下記の2つの分析法を用いた。

要介護度が高いと味覚の減退、すなわち食事が美味しいという感覚や味覚が減退したりする問題を生じやすいという仮説の検定を行う目的で表5のTable 1 (2×4表) 示す $\chi^2$ 検定を行い $p=0.0053$ の高い有意性を得た。この結果をもって仮説が確認されたと考えていいだろうか。問題になるとしたら何が問題であろうか。表5の解析を見せられた医歯学の研究者は推測するに下記のような疑問を提示するのではなかろうか。

- a) 男女間で味覚の差があるのでは？
- b) 高齢になると味覚能力が低下するのでは？
- c) 自分の歯があると無いでは味覚が違ってくるのでは？
- d) 基礎疾患が多いと味覚に影響するのでは？

これらの疑問はごく自然でかつ解析上重要な点である。疫学的にいうとこれらa)、b)、c) およびd)の疑問中に示される要因は要介護度という事象の背景にあって要介護度とは別の作用で味覚に影響を与えている可能性がある。このことからa) - d) は交絡因子 (confounding factor) 候補ということになる。しかしその議論の前に統計手法の選択に問題は無いであろうか。ここで表5下表のTable 2は上表のTable 1の表側の第1水準と第4水準を入れ替えたものであるが $\chi^2$ 値は全く同じ12.7038である。これは $2 \times n$ 表で $n$ が3以上の場合 $\chi^2$ 検定はデータ頻度が凸凹しているということが偶然か否かを判定しているのであって右肩上がり、または下がりの傾向（トレンド）を判定しているわけではないからである。すなわち傾向（トレンド）は通常の $\chi^2$ 検定だけでは確定できない。そこで登場するのが名称が少し厳めしいコクラン・アーミテージ検定 (Cochran-Armitage

Trend Test) 通称トレンド検定である。表6はその計算式、表7はその検定結果とその解釈を示している。表7において①傾きの有意性 $\chi^2$ 検定、と②直線からの乖離の有意性の $\chi^2$ 検定、2つを行ったところ①は $p=0.0017$ で傾き0という帰無仮説は棄却されること、そして②は $p=0.2360$ で直線とい

表5  $\chi^2$ 検定表の水準を入れ替えても $\chi^2$ 値は変わらない

		食事の美味しさと味覚				
		状態良		状態不良		
		0	1	2	3	計
要介護度 = 0	36	6	2	3	1	46
要介護度 ≥ 1	67	38	29	4	4	138
対象高齢者数	103	44	32	5	5	184
		$\chi^2 = 12.7038$		$p = 0.0053$		
		d.f = 3				

		状態良		状態不良		
		0	1	2	3	計
要介護度 = 0	36	1	3	6	46	
要介護度 ≥ 1	67	4	29	38	138	
対象高齢者数	103	5	32	44	184	
		$\chi^2 = 12.7038$		$p = 0.0053$		
		d.f = 3				

第2水準と第4水準の値を入れ替えても $\chi^2$ 値は変わらない

**Note:**  
 a) 食事の美味しさ: 0:おいしく食べられる 1:あまり美味しさを感じない 2:おいしくない  
 b) 味覚の状態: 0:問題無し 1:味がよくわからない 2:味がほとんどわからない  
 上記2指標の相加効果は $a+b=0-4$ の5水準であるが実際値は0-3のため表側の水準は4段階

表6 コクラン・アーミテージ検定の計算式

1. ケース総数	$n = \sum n_i$
2. 陽性総数	$r = \sum r_i$
3. 平均陽性率	$p_m = r/n$
4. 平均水準 (外的基準)	$X_m = \sum n_i X_i / n$ $X_i$ : 個々の値
5. トレンドが有る場合の比率の予測値	$e p_i = a + b X_i$ $b = \frac{\sum n_i (p_i - p_m)(X_i - X_m)}{\sum n_i (X_i - X_m)^2}$ $= \frac{(\sum n_i p_i X_i - n p_m X_m)(\sum n_i X_i^2 - n X_m^2)}{\sum n_i (X_i - X_m)^2}$ $a = p_m - b X_m$
6. 検定1: 傾き0の $\chi^2$ 検定	$\chi^2_T = b^2 / \{ p_m(1 - p_m) \} \sum n_i (X_i - X_m)^2$
7. 検定2: 直線性の $\chi^2$ 検定	$\chi^2_O = \sum n_i (p_i - e p_i)^2 / p_m(1 - p_m)$
8. 検定3: 群間比率の差の $\chi^2$ 検定	$\chi^2_H = n^2 / r(n-r) \{ \sum n_i^2 / n_i - r^2 / n \}$

表7 コクラン・アーミテージ検定結果と解釈

<b>Cochran-Armitage Trend Test (2×4)</b>				
エクセル関数: CHIDIST( $\chi^2$ 値, 自由度)				
	要因 $\chi^2$ 値	自由度	p	
$\chi^2_T$ :	傾き	9.8162	1	0.0017
$\chi^2_O$ :	直線からの乖離	2.8876	2	0.2360
$\chi^2_H$ :	合計 (非一様性)	12.7038	3	0.0053
<b>結論:</b> 要介護度が高くなるに従って直線的に味覚の減退傾向が有意				
<b>検定結果の解釈:</b>				
① $\chi^2_T$ が有意 ( $p < 0.05$ ) → 傾きが0の仮説は棄却 → 群の比率は、外的基準値により増減				
② $\chi^2_O$ が有意であれば ( $p < 0.05$ ) → 直線性の仮定は棄却 → 直線では表現しきれない				
③ 上記①と②の結果を総合的に判断する。 $\chi^2_T$ が有意 ( $p < 0.05$ )で、かつ $\chi^2_O$ は有意でない ( $p > 0.05$ ) 場合、各群の比率は直線式で表現可				

う帰無仮説は棄却されず「直線性」が有意であること、をそれぞれ示している。これら2つの検定結果から傾きを持った直線的な関係が有意、すなわち。要介護度が高くなると味覚が減退する傾向（トレンド）が確認できた。それではこのことを結論としていいかどうかについて考えてみると前述のようにこの段階で前述の3つの交絡因子の調整し結果を歪める背景因子の影響を無くすることが必要となる。これらの調整は実は結果の質を決定的に左右する重要な意味を持っており調査をする前と調査をした後でその手法が異なる。調査前の対策は分散分析法の生みの親 R. A Fisher の実験計画法そのものである。理論的には a. 性、b. 年齢および c. 現在歯数が比較する群と水準間で偏らないように無作為（randomize）に配置すればいいわけである。そうすれば a、b、c が交絡因子であっても群間に構成比率の差がでないことで結果に影響を与えないことになる。臨床疫学における RCT 法：Randomized Control Trial<sup>11)</sup> はまさにそれである。問題は現実的手法か、医学であれば倫理上の問題を惹起しなしかということである。農業においてある優れた品種を選ぶ問題と違って医学研究のフィールド観察データの場合理想的な実験配置はほとんど不可能といえる。そこで事後的に統計的（数学的）な手法を用いて調整するわけである。図9はSPSSの一般線形分析（GLM）の一つである共分散分析（ANCOVA: Analysis of Covariance）の入力画面である。某地区老人ホーム入所者データ男女184名のデータから従属変数（D）に「味覚減退」、固定因子（F）に「性別」、

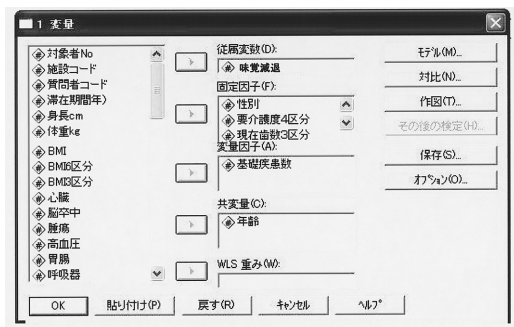


図9 SPSSによる共分散分析変数入力画面

「要介護4区分」、「現在歯数3区分」、変量因子（A）に「基礎疾患数」および共変量（C）に「年齢」を挿入してある。ここで解釈時の混乱を回避するため分散分析と共分散分析の専門用語である（1）固定因子（Fixed Factor）と変量因子（Random Factor）、（2）（Covariate）および（3）タイプ I、II、III、IV平方和、について予め説明しておくこととする。

#### （1）固定因子と変量因子について

固定因子として取り上げられたのは性別、要介護度4区分、および現在歯数3区分の3つである。これらはいずれも限られた数の水準（区分）からなり別の標本に対して同じ分析を適用してもその水準が変化しない（あるいは変化させる必要がない）。これに対して変量因子に分類した基礎疾患数は対象者数が多ければ理論的には無限個の水準が想定され、本調査では対象者をランダムに抽出し調査した結果、頻度が多い基礎疾患による組み合わせ数がたまたまこうなると考えられる。また、例えば基礎疾患数が2と3と4の場合の影響の違いなどという水準間の違いがどう影響するかという点につき特定の関心が無い（あるいは学問的意味合いが小さい）場合である。なお、扱いは分散分析（ANOVA: Analysis of Variance）の場合も同様である。

#### （2）共変量について

共分散分析（ANCOVA: Analysis of Covariance）といわれる所以はこの共変量（Covariate）であり分散分析では使用せず共分散分析にのみ使用される。別名、従属変数に対する影響を統制する対象となる剰余変数（Extraneous Variable）と呼ばれ、研究目的とする指標（独立変数）の背景にあって従属変数と相関がある指標、すなわち交絡因子（confounding factor）の可能性のある指標である。図10は分散分析と比較して共分散分析で検出力が高まる理屈を示している。図10の左右の図において4点のデータ値は同一である。分散分析では誤差に分類されるバラツキ（ $\sum d_1^2 + d_2^2 + d_3^2 + d_4^2$ ）が平均値からの偏差の二乗の総和であるが共分散分析では共変量と従属変数との

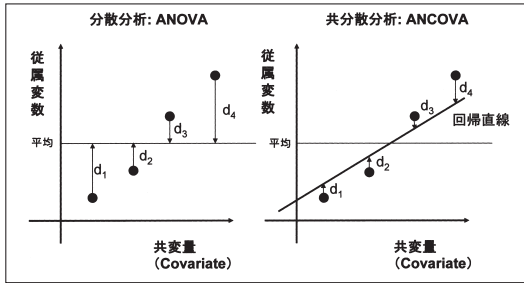


図10 分散分析と共分散分析

回帰直線からの偏差の二乗の総和であり

誤差=回帰で説明できる系統誤差+偶然誤差に仕分けるためF検定を行う際分母となる誤差項が小さくなり検出力が高まるという理屈である。

(3) タイプI、II、III、IV平方和について

分散分析、共分散分析において調べる要因が互いに無相関(ベクトルが直交関係)であれば問題ないが実際には何らかの相関関係がある場合が多い。この偏差平方和の加法性である「全分散=要因1+要因2+要因3+・・・+誤差分散」が単純には成立しない。そこでその相関の部分の影響をコントロールする方法が幾つか開発されておりそれが「タイプI~IV平方和」である。以下、その違いの説明である。

タイプI平方和: 主効果と交互作用を順番にモデルに挿入する方法(逐次平方和法)、例えば要因1に要因2を、次に要因3をモデルに挿入した場合それぞれの平方和の増加分をその要因の偏差平方和とする方法

タイプII平方和: タイプIと異なり予め全ての主効果をモデルに投入し、そこからある要因を除去したときの平方和をその要因の平方和とする方法。次に全ての主効果と全ての交互作用を挿入したモデルから一つづつ交互作用を除去し、その平方和の差をその交互作用要因の平方和とする方法

タイプIII平方和: タイプIIと違い全ての主効果、交互作用をいっぺんにモデルに投入し、そこからある要因を除去したときの平方和の差をその要因(交互作用要因を含む)の平方和とする方法

タイプIV平方和: タイプIIIの特殊型でデータが空

白あるいは欠損のセルがある場合に適用

本題に戻って共分散分析の事例の結果を表8に示すが主効果だけのタイプI平方和法で計算されている。その理由は表8の結果を得る前に予め各要因間の交互作用の有意性を確認し有意性がなかったためである。ここで3つの固定因子、1つの変量因子および1つの共変量のうち要介護度4区分のみがp=0.0058で有意となった。性別、現在歯数3区分、基礎疾患数および年齢はいずれも有意ではなかった。従って「要介護度が高くなると味覚が減退する」という傾向が共分散分析法でその要因を調整して場合でも確認された。なお、この傾向は味覚を従属変数とした変数増減法による重回帰分析(変数選択基準: Fin=Fout=2.0)を行っても同様に要介護度4区分のみが選択された。

今回本事例に重回帰分析を行っても結論は同じであった。しかし必ずしもいつも両方適用できるわけではない。また両者が適用できたとしてもとりわけ第2、第3等の重みが小さい要因について結果が同じになるとは限らない。その理由は従属変数が連続変数であるかあるいは適切に出来るだけ細かく数量化<sup>12)</sup>した数値で行う重回帰分析と連続数変数を小数のカテゴリー化して行う分散分析で情報が微妙に異なってくること、および独立変数として分析に使用できる数値の尺度<sup>13, 14)</sup>の違いと仮説モデルの違いによる。ここでいう仮説モデルの違いは両者の違いを考えると最も重要な違いである。分散分析は比較的少数の要因に関して分散の加法性に注目して誤差分散を分母とし

表8 共分散分析(ANCOVA)結果

ソース		タイプI平方和	自由度	平均平方	F値	有意確率
切片	仮説	82.223	1	82.223	69.546	.0041
	誤差	3.439	2908	1.182		
性別 1M2F	仮説	.049	1	.049	.070	.7917
	誤差	119.741	171.250	.699		
要介護度4区分	仮説	9.090	3	3.030	4.303	.0058
	誤差	123.391	175.455	.703		
現在歯数3区分	仮説	.319	2	.159	.227	.7968
	誤差	121.889	173.827	.701		
基礎疾患数	仮説	5.864	6	.977	1.399	.2175
	誤差	118.736	170	.696		
年齢	仮説	.719	1	.719	1.030	.3116
	誤差	118.736	170	.696		

- Note:
1. 従属変数: 味覚の減退(美味感覚の減退を含む)
  2. タイプI平方和: 本文説明参照
  3. 前提条件であるグループ間の誤差分散の均質性の確認 Levene 検定 F=0.6528 p=0.9740
  4. 要介護度4区分: 日本の要介護判定区分(1+1.5)を 01,2,3(4+5)の4段階とした
  5. 現在歯4区分: 1現在歯0-9歯 210-19歯 320歯以上
  6. 基礎疾患数: 心臓病、脳卒中、腎臓病、高血圧、糖尿病等老人に多い14疾患数の総計
  7. 各要因間の交互作用はいずれも有意でないのでモデルから除外してある



た分散分析値 (F 値) によって各要因の重みの有意性を判定する各要因の分散和モデルである。一方、重回帰分析や多重ロジスティック回帰分析は実測値と回帰式の誤差を最小化することを主眼としており回帰式の係数の有意性 (t 検定、または F 検定) と適合度の変化の有意性 (F 検定) で最適モデルを選択する回帰式モデルだという違いである。なお、本稿で扱った共分散分析は分散分析と重回帰分析の中間のいわばハイブリット的分析である。この2つの分析法に限らず2-2) - (2)のタイ国の Dengue 熱の事例で解説したように同じデータでも採用した分析法の性質の違いと交絡因子の調整の違いが生じて主因以外の一部の結果が異なることがあることに注意して分析法を選択<sup>15)</sup>することをお勧めしたい。

## 5. おわりに

今回は傾向 (トレンド) 分析をテーマに概説した。事例に示したいずれの分析も耳慣れないかも知れないやや特殊な手法である。しかし主として動的な時系列データを扱い高級感のある切れ味を実感できる分野と思われる。冒頭でも触れたように目的と方法が全く異なる分析が同じ傾向 (トレンド) 分析と呼ばれることがしばしば混乱を生んでいると思われる。統計用語は法律用語ではないので仕方がない側面もあるが、多くの統計分析法、例えば相関分析、主成分分析、因子分析、分散分析、判別分析等は多義的ではなく一義的である。その大きな理由として「相関」、「主成分」、「因子」および「分散」の統計学的定義が固定されているためと考えられる。「判別」は日本語では多義的であるが元々の英語 discriminant が特殊なため固定できていると考えられる。しかし傾向 (トレンド) 分析の場合、傾向や Trend、Tendency の各言葉は一般的過ぎることに原因があるように思われる。例えば「要因分析」のような用語を用いた場合その方法は極めて多義的となるのに類似している。このように今回の傾向 (トレンド) 分析を学習または利用する場合は目的、用法が多義的であることを念頭に、まず目的にあった手法の選択

が必要である。本稿がその際参考になれば幸いである。

謝辞：本総説を執筆するにあたり各段階でご指導いただいた歯科疫学研究会の顧問である重松逸造、蓑輪眞澄、高江洲義矩、境 脩、伊藤学而、佐々木英忠の各先生に深謝申し上げます。さらに貴重なご助言をいただいた深井穂博同研究会会長をはじめとする幹事の方々に感謝申し上げます。また事例データ収拾と解析に協力頂いたスリランカの Dr. A. Usha S. Perera, Dr. L. Shereen M. Y. Perera 両先生に感謝申し上げます。

## 参照 Web-site (2006/12/20 現在)

- WS01) [http://en.wikipedia.org/wiki/Weighted\\_moving\\_average](http://en.wikipedia.org/wiki/Weighted_moving_average) : 移動平均法  
 WS02) [http://www.opticast.co.jp/opt/gcom/kakuzuke/mikata/mikata\\_idou.htm](http://www.opticast.co.jp/opt/gcom/kakuzuke/mikata/mikata_idou.htm) : 各種移動平均法  
 WS03) <http://www.kitco.com/LFgif/ag1987D.gif> : 貴金属相場月別長期データ  
 WS04) <http://orikura.jp/futures/investment/palladium/change.html> : パラジウム：変動の基礎的要因

## 文 献

- 1) Christopher W. Corti, Richard J. Holliday: Commercial Aspects of Gold Applications: From Materials Science to Chemical Science, Gold Bulletin 2004, 20-26, 2004.
- 2) 鈴木義一郎：ブルーバックス 現代統計学小辞典、東京、第1刷、1998、227-228頁。
- 3) 瀧口 徹：JICA Sri Lanka - Dengue (出血) 熱 Q and A (日本語版、English ver)、スリランカ、第1刷、2005、7-12頁。
- 4) Thammapalo S, Chongsuwiatwong V, McNeil D, Geater A: The climatic factors influencing the occurrence of dengue hemorrhagic fever in Thailand, Southeast Asian J Trop Med Public Health, 36 (1) 191-196, 2005.
- 5) Dharam P. Agarwal: Review Alcohol and Alcoholism, Medical Council on Alcohol Vol. 37, No. 5, 409-415, 2002.
- 6) 瀧口 徹：歯科疫学統計 - 第3報 重回帰分析、多重ロジスティック回帰分析モデルの適合度判定指標の解釈 - - SPSS, STATA の利用に際して - ,

- Health Science and Health Care, No 1 vol 5 : 35-49, 2005.
- 7) P. McCullagh, J. A. Nelder : Generalized Linear Models. Chapman & Hall/CRC., USA, 2nd ed, 1999, pp.7-8, 30-31, 33-36.
- 8) 瀧口 徹 : 歯科疫学統計 - 第2報 一般化線形モデルの意義と潮流 - Health Science and Health Care, No 1 vol 4 : 29-36, 2004.
- 9) Ministry of Healthcare & Nutrition: MOH&N, Sri Lanka : The Health Bulletin 2003.
- 10) 劉 僖根, 鈴木秀明 : カオス性からみた心拍変動の加齢変化, Bulletin of Toyohashi Sogo College, No. 1, 161-170, 1997.
- 11) 瀧口 徹 : EBMのための(臨床)疫学・統計学的基礎 (3), 第3章 EBMの基礎 : 臨床疫学の最近の潮流とポイント, 障害者歯科学雑誌, 23巻 : 443-458, 2002.
- 12) 杉山高一 : 多変量データ解析入門, 朝倉書店, 東京, 第1刷, 1983, 38-39頁.
- 13) 瀧口 徹 : EBMのための(臨床)疫学・統計学的基礎 (1) 第1章 統計の基礎 : 一部のデータから全体を推定する, 障害者歯科学雑誌, 23巻 : 1-10, 2002.
- 14) 瀧口 徹 : 歯科疫学統計 - 第1報 各種統計分布の相互関係と利用の潮流 - Health Science and Health Care, No 1 vol. 4 : 19-27, 2004.
- 15) 瀧口 徹 : EBMのための(臨床)疫学・統計学的基礎 (2) 第2章 疫学の基礎 : 流行病の法則性を見つけ予防する, 障害者歯科学雑誌, 23巻 : 89-98, 2002.

## A review of oral epidemiological statistics

– Part IV : The bases of Trend Analysis –

Toru Takiguchi

(Fukai Institute of Health Science)

Key words : Trend Analysis, Time Series analysis, moving average, Cochran-Armitage test for trend, Analysis of Covariance (ANCOVA)

In the fourth review, trend analysis was carefully dealt with. This analysis is very rare case in terms of the purposes and statistical methods are polysemous. Therefore some perplexity may ensue in using this method. Hence, trend analysis was interpreted in relation to the four view points noted below in summing up the analysis.

The analysis, which accentuates:

1. both, on the assessment of the trend and the prediction of the near future by means of eliminating the fluctuations.
2. on the calculation of the regression curve which is best fitted to the actual data.
3. on the classification of the time series data.
4. on the analyses of factors which indicate the dose-response relationship.

The case studies on trend analysis are interpreted by using five characteristic data sources noted below:

- a) The time series data on the market price of precious metals (e.g. Gold, Palladium which are important for dental treatment in Japan) - Case Study for the analysis 1
- b) The three types of data on the relationship between age and the average number of present teeth - Case Study for the analysis 2
- c) The time series data for twenty years, on the incidence of Dengue Haemorrhage Fever: DHF in Thailand - Case Study for analysis 2
- d) The time series data for twenty years, on the incidence of communicable and non-communicable diseases in Sri Lanka - Case Study for analysis 3
- e) The relationship between the grade of nursing care for elderly and the loss of the individual sense of taste in a nursing home - Case Study for analysis 4

The know-how (when and how to use) of each trend analysis using these case studies mentioned above in the fourth Review were interpreted.

Reprint requests to T. TAKIGUCHI, Fukai Institute of Health Science, 3-86, Hikonari, Misato-shi, Saitama 341-0003, Japan

TEL:048-957-3315 / FAX:048-957-3315 / E-mail: taki8020@mth.biglobe.ne.jp