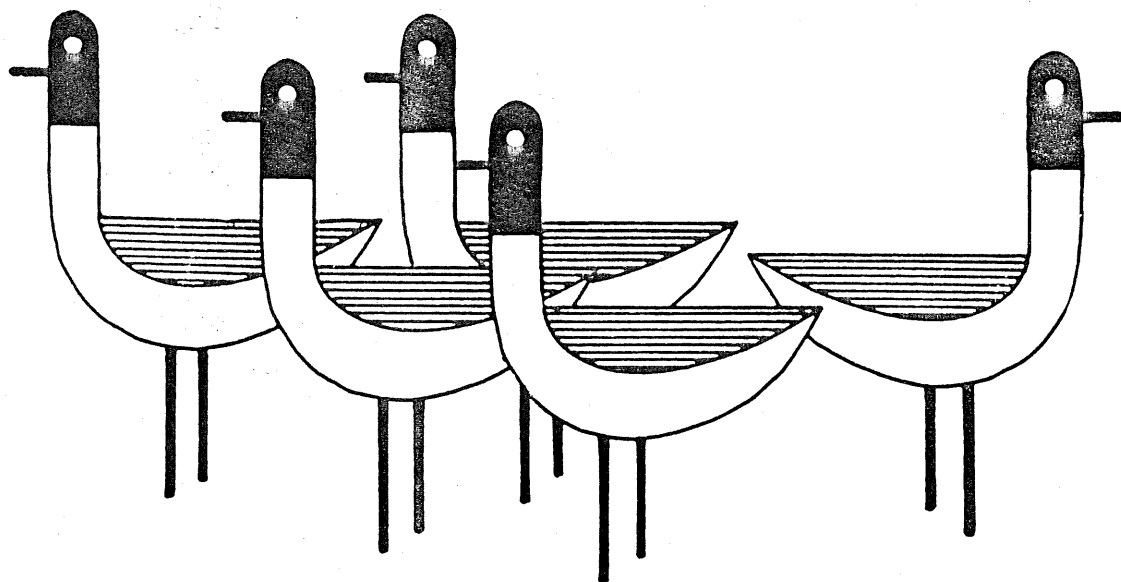


臨床医学における意志決定過程の
基礎的研究



- the disease in different countries, *Acta Dermatovenereol* (stockholm) 50, 293-301, 1970
18. Campos, J.L : Application of the Weibull distribution to some instances of cell survival, and of aging, *Brit. J. Radiol*, 48, 913-917, 1975
 19. Burch, P.R.J : Application of the Weibull distribution, *Brit. J. Radiol*, 49, 564, 1976
 20. Masuyama, M : A mixture of two gamma distributions applied to rheumatoid arthritis, *Rep. Stat. Appl. Res.*, JUSE, 24, 82-85, 1977

班員 柳井晴夫

- (1) H.Yanai, Y.Inaba, H.Takagi & S.Yamamoto; Multivariate Analysis of Cancer Mortalities for Selected Sites in 24 Countries, *Environmental Health Perspective*, Vol 32, pp83-101, 1979
世界24ヶ国の13の癌の部位に関する死亡率を因子分析し、さらにそれぞれの国に関する24の食物指標を用いて食品摂取状況と各部位の死亡率との関連を分析した。
- (2) H.Yanai; A Proposition of Generalized Method for Forward Selection of Variables, *Behaviormetrika*, No 7, pp95-107, 1980
2変数群X, Yについての関連を分析する正準相関分析について、前進選択方式にもとづく変数選択の方式を導入し、さらにその特別な場合としての因子分析における変数選択方式を提案した。
- (3) H.Yanai, Y.Yoshimoto, H.Takagi, K.Maeda & H.Kurita; A Statistical Analysis of Detecting Risk Factors for Stomach Cancer Patients; *MEDINFO 80*, pp974-978, 1980
胃癌に関するケースコントロールデータを数量化二類を用いて分析し、合わせて胃癌のリスクファクターの加乏による変化構造を分析した。
- (4) K.Uchiyama, H.Ishikawa, H.Ito, S.Oda, A.Onda, H.Hashiguchi, H.Yanai, T.Watanabe, H.Kubota, H.Nasuno, Y.Imai, T.Konishi, T.Kubo & M.Yamashita; The Broad Spectrum Inventory System (JMI) applied for the Assessment of Mental Health, *MEDINFO 80*, pp969-973, 1980
身体の異常性をみる尺度、精神の異常性をみる尺度、性格尺度、職場適応に関する尺度、計

56 尺度, 596 項目からなる質問紙形式による総合的健康テストの開発手順について紹介したものの。

- (5) 山本俊一, 稲葉 裕, 柳井晴夫, 高木広文; ガンの疫学統計の方法論に関する研究, 厚生省「ガンの疫学」分担研究報告, 1980 年 3 月
これまでに筆者らが行った癌の疫学に関する研究を紹介し, さらにこれらの研究の問題点を考慮した。
- (6) 柳井晴夫, 高木広文; 計量診断と多変量解析, 大阪大学 BME 研究会, 抄録集, pp16-21, 1980
- (7) 柳井晴夫; ノルムをつくる手続きと問題点特集「正常と異常のはざま」, 教育と医学, 6 月号, pp17-25, 1980
- (8) 大山 正, 武藤真介, 柳井晴夫; 行動科学における統計学, 朝倉書店, 1980

疫学データ解析のための相関分析に関する基礎的研究 千葉大 柳井晴夫 順天堂大・医 稲葉 裕 東大医 高木広文

疫学データ解析のための方法論として, 多変量解析の各種の技法が頻繁に用いられるようになってきたが, その反面, 多変量解析の結果を導く前提となる相関関係についての基礎的研究は必ずしも十分に行われていない。ここで 3 つの角度から, これまで筆者らが行った相関分析に関する研究の概要を記述する。

(1) 相関係数と相関図について

2 変数 x, y のピアソンの積率相関係数がゼロであっても, x と y の間に曲線相関が存在する場合もあるし, データが 2 つの群に分かれ, 一方では正の相関, 他方では負の相関をもつ結果, 全体から打ち消されて, ゼロの相関を持っている場合もある。

このような場合, 相関図を描くことが大切であることはいうまでもないが, 相関図にいくかなる場合でも回帰直線を描くことは好ましくない。 x と y に $x \rightarrow y$, または $y \rightarrow x$ のいずれの関係も成立しない場合は, むしろ棄却楕円を描くべきであろう。楕円の長軸を $a(x - \bar{x}) + b(y - \bar{y}) = 0$, 短軸を $c(x - \bar{x}) + d(y - \bar{y}) = 0$ とすると, x, y の 2 変数間の共分散行列の固有値の小さい方に対応する固有ベクトルが (a, b) を与え, 大きい方の固有値の固有ベクトルが (c, d) に対応する。

(例 1) 20, 30, 40, 50, 60 代の血圧 (最高, 最低) の相関図を血圧楕円として示したのが図 1 である (このような基礎データを多く集積することは計量医学における基本的研究として必須のものであろう)。

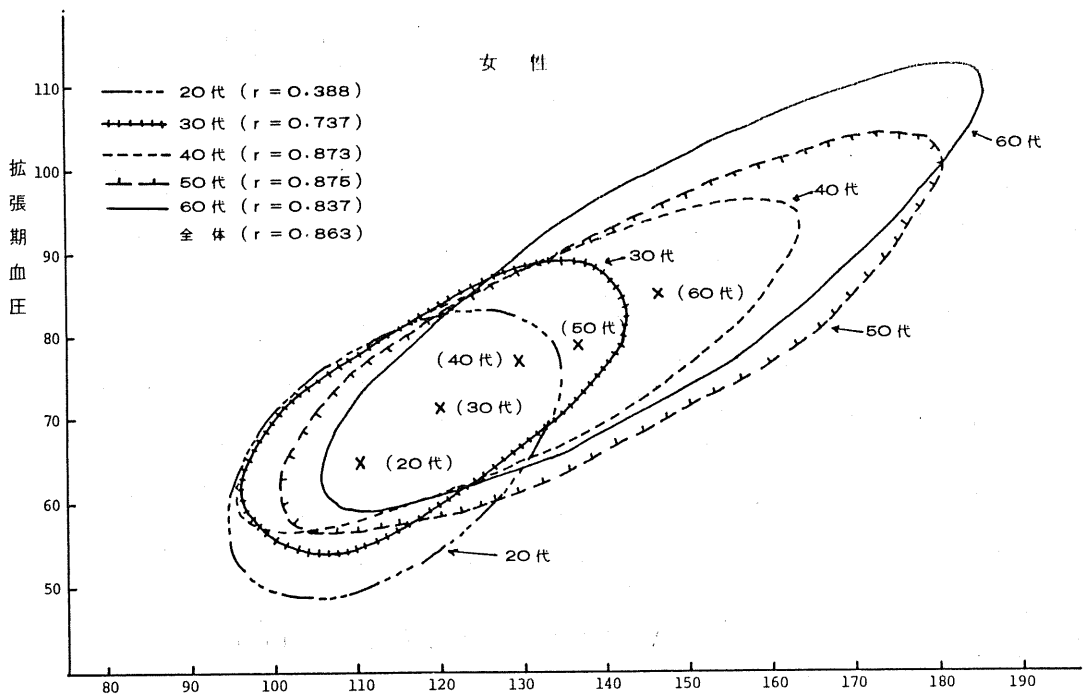


図1 収縮期血圧と拡張期血圧との相関図

表1 悪性新生物の7つの部位に関する相関係数

(上段は24ヶ国, 下段括弧内は日本46都道府県
のデータを用いて計算したもの)

1	食道	1.000 (1.000)					
2	胃	0.344 (0.385)					
3	腸	-0.221 (0.278)	-0.729 (0.548)				
4	直腸	-0.175 (0.405)	-0.339 (0.589)	0.754 (0.574)			
5	肺	-0.001 (0.203)	-0.107 (0.244)	0.516 (0.540)	0.589 (0.377)		
6	皮膚	-0.188 (-0.266)	-0.429 (-0.607)	0.462 (-0.513)	0.097 (-0.502)	0.114 (-0.118)	
7	白血病	-0.387 (-0.155)	-0.603 (-0.347)	0.423 (-0.073)	0.329 (-0.255)	0.138 (0.229)	0.392 (0.254)
		1)	2)	3)	4)	5)	6)

相関係数のもう一つの問題点は標本の選び方によって相関係数が著しく変わる点である。
(例2) 表1に世界24ヶ国のデータによって計算される6つの部位の訂正死亡率間の相関係数
と我が国の都道府県別に計算される間接法に基づく訂正死亡率間の相関係数を示す。胃と腸の
場合, 外国では-0.728, 日本では0.548と著しい差がみられる。

一般に、地域を単位として相関（地域相関）を求める場合、地域の大きさの区分によっても相関の大きさが異なり、地域の単位が大きくなると相関係数も大きくなる傾向がみられる。

(2) 変数選択の方法

判別分析や重回帰分析のような外的基準のある場合の分析手法を用いる場合、多くの変数を分析に用いると、いわゆる多重共線性が生じ、判別係数が偏回帰係数の値が不安定になりやすい。したがって効率的に変数を選択する方式が必要とされることになる。変数選択の問題としては、(a)どのような順序で変数を判別式や重回帰式に合わせるか、(b)いくつの変数を選択すべきか、に分けられる。第1の点は重回帰分析における種々の変数の選択の手法（前述選択、後進選択、変数増減、減増法）があるが、この手法は二群の判別、多群の判別、さらには正準相関分析における変数選択や、因子分析において潜在因子を外的基準とした場合の変数選択にも適用可能となる。第2の点としては、重回帰分析におけるP.S.S., A.I.C, Cp, 統計量や自由度調整済み重相関自由度調整済みマハラノビスの汎距離を用いた方法がある。これらの方法の優劣については、まだ十分に検討されておらず、今後の課題である。

（例3）Y.G. 検査の12尺度について、バリマックス法に因子分析を行い、得られた4つの因子について前進選択にもとづく変数選択を行なった結果を表2に示す（文献2参照）。表の上から順に選ばれた尺度名が示されている。

表 2

Result of forward selection method in factor analysis

	1	2	3	4	commu- nality	s_j	$s_{(j)}$
1 D	0.684	-0.298	0.254	-0.326	0.728	0.728	0.728
2 A	0.183	0.821	-0.040	0.098	0.715	0.698	1.426
3 Co	0.751	0.186	-0.189	0.199	0.674	0.451	1.877
4 C	0.466	-0.102	0.534	-0.018	0.512	0.265	2.142
5 R	0.032	0.457	0.209	0.438	0.445	0.190	2.333
6 G	-0.082	0.677	-0.030	-0.117	0.479	0.170	2.503
7 S	-0.135	0.795	0.012	0.030	0.651	0.099	2.602
8 N	0.808	-0.157	0.045	0.059	0.684	0.088	2.690
9 Ag	-0.056	0.404	0.407	0.026	0.332	0.057	2.747
10 O	0.837	0.073	-0.020	0.011	0.707	0.053	2.800
11 T	0.157	0.092	-0.171	0.471	0.284	0.034	2.834
12 I	0.383	-0.535	0.185	-0.000	0.461	0.020	2.854

Note: $s_{(j)} = s_1 + s_2 + \dots + s_j$

上記の手法は、多くの医療検査項目から検査全体のもつ情報量を失うことなく、効率良く少数の検査項目を選択する際に有用となろう。

(3) 因果関係の分析をめぐる

相関関係の分析の究極の目的はやはり因果関係の分析といってよい。とくに疾病の原因を探

究する疫学の立場からは、実験疫学と違って、厳密な実験条件を設定した因果関係の解明方法をとることができにくいので、各変数間の相関関係のもつ情報を最大限に生かした因果関係の推定を行うことが必要とされる。

相関関係を用いた因果関係の推定法には Blalock による偏相関係数に基づく推定法、すなわち3つの変数間に

(a) $X \rightarrow Y \rightarrow Z$, (b) $Z \rightarrow Y \rightarrow X$, (c) $X \rightarrow Y \leftarrow Z$ のうちの1つのモデルが成立する場合、Yを固定したXとZの偏相関係数がほぼゼロになることを用いる方法がある。さらに、偏相関が必ずしもゼロにならない場合は、説明変数間にいくつかの因果モデルを想定して、変数間の連鎖(パス)の強さを重回帰分析における偏回帰係数の大きさとして推定するパス解析の方法があることは周知であろう。さらに最近、変数間の相関係数の大きさのみの情報を用いて、説明変数間の連鎖モデルを構成する方法

(Takagi(1980) Behaviormetrika Vol.8) 等が開発されている。いずれの方法にも一長一短があるが、これらの方法に時間の因子を加えることによって、さらに強力な因果関係の推定法が開発されてこよう。なお、前節で述べた変数選択の方法も因果関係は推定の方法と密接な関わりをもってくる。

(例4) 山梨県の市町村の資料によりパス解析した結果を図2に示す。肝ガンの原因として、日本住血吸虫症の影響が強く出ていることがわかる。(文献(5)参照)

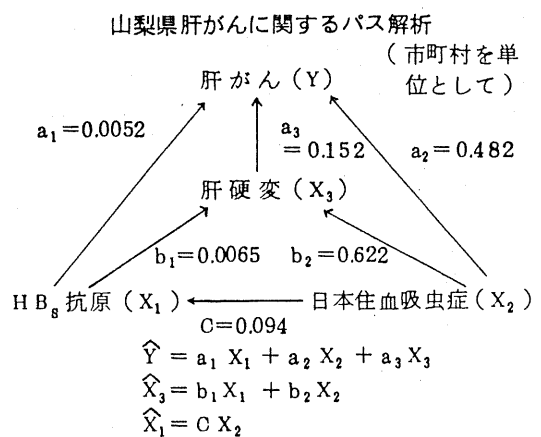


図2 山梨県肝がんに関するパス解析

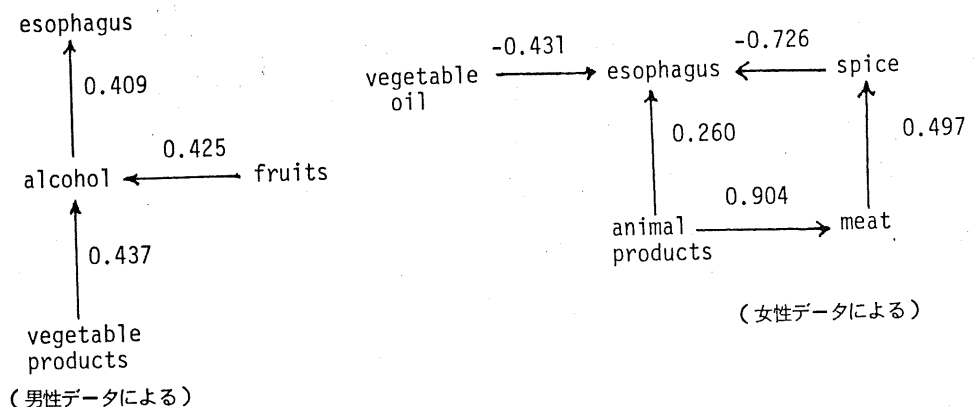


図3 食道がんに関する因果関係の分析

(例5) 前記のTakagiの方法にしたがって世界24ヶ国のデータを用い、食道癌の risk factor について因果関係の分析を行ったものを図3に示す。男子の場合はアルコールが直接的なリスクとなり、女子の場合は動物製品 (animal product) 植物油、香料 (spice) が複合的なリスク予防となっている (香料と動物製品はパス係数がマイナスとなっているために、ともに少量摂取がリスクとなることを示唆するものである)。

今後の研究課題としては、必ずしも線形でない種々の形式をとりうる相関構造をいかに把握するか、そして、それらをいかにして多変量解析と結びつけるか、さらには、時間因子を考慮したうえでの相関関係と因果関係の関連などを計量医学における視点から分析していきたい。

班員 古川俊之

- (1) 古川俊之：病態構造の解析，「臨床化学の進歩」(田村善蔵 編) pp189～206，学会出版センター，東京，1980

病態観測値相互のデータ構造を明らかにするために応用される統計学的方法，とくに多変量解析の手法を解説し，さらに診断過程との関係に逆問題解の概念を導入して，新しい理論体系の可能性を示した。またリスク関数，時系列確率関数，マルコフモデルなどの応用事例を紹介，解説を加えた。

- (2) 古川俊之：老化因子の解析，「日本医師会医学講座」(日本医師会 編) pp1～11 金原出版，東京，1980

老化学説に共通したモデルとして，生命維持のために本質的な何らかの形質が加齢とともに損耗してゆくという概念に注目し，年齢階層別死亡率曲線のシミュレーション，死亡発生のワイブル分布特性，さらに生物学的年齢の推定の手法を解説した。また高齢者の健康調査にもとづく特性，長寿のための安全因子について最近の知見を紹介した。

- (3) 古川俊之：集団検診におけるデータ処理，臨床検査，24(2):157，1980

集団検診データの解釈と予防医学への応用をはかるための方法を系統的に論じた。検診データの収集，記録，解析ならびに追跡調査による事後評価を，スウェーデンの大規模検診を例として考察し，また検査の感度，安定性の概念の意義，多変量解析の効果について述べた。

- (4) 古川俊之：脳卒中死亡のワイブル分布，総合臨床，29(1):69，1980

特定の疾患によって死に至る場合の余命がワイブル分布に従うことを示し，脳卒中，急性心筋梗塞は初期故障型，ネフロ-ゼ症候群は磨耗故障型と異なる死亡モードをとることを証明した。また脳卒中および急性心筋梗塞では発症後4週目で，死亡リスクの減少が共通して起ることを見出した。

文部省科学研究費補助金 総合研究（A）研究報告

研究代表者

東京大学医学部

古川俊之

池田 田 研 二
稲井 甲 通 紘
上 上 弘 敏
岡 松 光 明
開 島 成 治
梶 原 文 允
久 谷 武 彦
児 保 士
野 玉 三
林 瀬 明
三 宅 章 夫
宮 原 英
柳 井 晴 夫

阿 部 裕
木 村 一
吉 利 和

昭和 56 年 3 月