解 説

経時測定データの解析法

ーその文献とソフトの紹介ー

七 堂 利 幸*

要旨 時間経過とともにデータを取っていく経時測定データは、時点間に系列相関があり、各時点を独立した値と見て解析すると、間違った結果になる。この経時測定モデルは鍼灸研究では、かなりの頻度で使われている。鍼灸研究における統計誤用が頻発しているため、この経時測定データの解析法について調べても、ほとんど誤用といってよい。そのため、正確な解析法の情報を学会会員に伝えるため、経時測定データの解析法の文献とコンピューター・ソフトウェア・パッケージ(統計ソフト)のレビューをした。

はじめに

時間経過とともに変化するデータを解析する方法には、時系列解析法がある。心理学、行動科学や医学では、工学、経済学で用いている時系列解析という言葉より、時点数が少ないという点で経時測定データの解析と呼ばれている¹⁾。

この分野はまだ発展段階で定型的方法と言うものは無い。しかし、鍼灸の試験や実験ではこの経時測定データになる場合が多い。何が問題かというと、同一個体で時間を変えて繰り返し測定すると、その観測値は個々に独立していない、系列相関が見られるのである。だから相関を考慮した解析方法をとる必要がある。よく見られる誤用はその相関を無視した解析法を使っていることである。現場では各時点毎の比較をしたい。そこで例

えばt検定を繰り返し、検定の多重性の問題を犯してしまうのである。これらの対策について和文の教科書が少ない。そのためかこの誤用が多く相変わらず続いている。

例えば、この経時測定データを一元配置分散分析とt検定の繰り返しや、三元配置分散分析や繰り返しのある二元配置分散分析での誤用が相次いで起こっている。

鍼に関係する国内の三大雑誌といわれている、全日本鍼灸学会誌、日本温泉気候物理医学会誌、そして日本東洋医学会誌には、この誤用が継続している。また西洋医学では欧米の論文へ投稿した時、この誤用のため差し戻されるケースがあるという。

米国の製薬企業では繰り返し測定の分散分析が最も頻回用いられ、繰り返し測定の分散分析の応用に関心を持っていると言われている。そして同時に使われるソフトウェア・パッケージは、SASやBMDPが最も多いことが調査されている²⁾。

本論では、特に鍼灸の分野で多用されるケースについて選択し、繰り返し測定データの解析法について、現場の実務家のために論文とソフトウェアのレビューを行いたい。

* Toshiyuki SHICHIDO 全日本鍼灸学会学術部研究委員会、情報・評価班

Key Words : 経時測定データ, 系列相関, 繰り返し測定の分散分析, 共分散分析, 自己回帰モデル, 時系列解析, 生存分析, 成長曲線モデル, ロジスティック回帰モデル, 混合モデル, randomization test, マルコフモデル, 累積法, クロスオーバー法, 対数線形モデル, 多変量分散分析, 多重比較, 輪切り検定

繰返し測定データの解析法の種類

数的(count)データは、そのまま数えられるデータだが、Origasa(1988)によると、そのモデルは、以下のとおりである³⁾。

- 1) 自由度調整付 ANOVA
- 2) 成長曲線モデル
- 3) 変量(Random)そして混合効果モデル
- 4) 2 値データのロジスティック回帰モデル
- 5) 時系列回帰モデル
- 6) 時系列解析
- 7) Stochastic process

また、質的(categorical)データのモデルは、Origasa(1987)⁴⁾によると、

- 1) ロジスティック回帰モデル
- 2) Greenland のモデル
- 3) 対数線型モデル
- 4) Koch の線型モデル
- 5) McCullagh の回帰モデル

があるという。

症状の陰性化を経時的にみた場合、陰性化率を各時点で調べれば、以上のモデルを使わなくとも普通の検定方法、例えばノンパラの群間検定でよい。

どの解析法を使うか、つまり定式化の問題は汎用的な方法はまだない。どのような考え方で解析法を選択していくかは、文献 5), 6) が参考になる。特に後者の繰返し測定データの解析という章では、何を比較したいのか、例えば時点間の比較なのか、群間比較なのか、ベースライン値との比較なのかなどによって解析方法が異なる。研究目的にあった解析方法を十分検討してから実験・試験する必要がある事が強調されている⁶⁾。

邦文の当該データのいろんな手法の実例は、文献 1) が分かりやすい。

1. 数的(count)データの解析

1.1 繰返し測定(多変量)分散分析

混合(mixed)モデルや split plot デザインを、繰返し測定(分散分析(repeated measurement designs))と呼んでいる⁷⁾。時点と治療の交互作用が

無い場合用いられる。これには先に経時的プロット図を書いて経時的な変化がどうなっているか、例えば群間で一定の傾向が認められるかどうか等を調べる必要がある。

欠測は実際の現場では起ると考えるべきで、欠測値がある時は得られているデータで乱塊法が適用できる。繰返しのない 2 元配置分散分析と同じだが、ただ被験者が有意となっても無視することである⁸⁾。時点に有意差があると、どの時点間で差があるかは contrast を使う。SAS での方法は文献 8) にある。

時点と被験者の交互作用は、計算値でも交互作用の図でもその有無が判定できる。

平均自乗の比を評価するのに使われる F テストが如何に影響するか Box (1953) が実際の手続きで示した^{9,10)}。なお初期の議論に触れると、等共分散マトリクスのための Box test と Mauchly 等方性(sphericity)基準(W)の妥当性の論議は、文献 11) にある。すべての対角成分が等しく、すべての非対角成分が等しい同一分散・同一共分散行列、つまり複合対称(compound symmetry)行列を考えると、初期の研究者たちは、この複合対称性が一変量 F 検定の妥当性の必要十分条件だと考えた。繰返し測定データの分散共分散行列がこの形かどうか見るのが、等方性検定といわれる。ただこの検定は臨床試験程度のサンプル数では強力でないが、大抵の繰返し測定データ解析用ソフトに用意されている¹²⁾。

繰返し測定では、時点に関する要因効果を検定するための F 値が大きくなりすぎ有意差が出やすくなるので、自由度を調整する Geisser and Greenhouse(1958)の G-G 法¹³⁾、Huynh and Feldt(1970)の H-F 法が考えられている¹⁴⁾。

G-G 法はあまりに保守的で実際の有意水準が小さくなりすぎる、H-F 法は G-G 法よりゆるやかな修正方法である。どちらを採用するかは、明確な指針がない。一応 2 群比較の場合は G-G 法、多群の場合は H-F 法が薦められている⁶⁾。これらの修正機能があるソフトウェアが一般的になってきている。

繰返し測定(分散分析)のデザインの学習には、文献 15) が分かりやすい。

最も多いと考えられるデザインは、一群の繰返し測定デザイン、繰返し1因子の多群、繰返し2因子の1群であろう。

名義尺度の繰返し測定でノンパラメトリック検定を使うときは、同一被験者が一つの治療前後を見る場合、McNemar 検定、同一被験者で多時点の測定では Cochran Q 検定を用いる。後者は二者択一の順序尺度にも有効である。

① t 検定の乱用

私どもの調査では全日本鍼灸学会誌の7割強にt検定が使われている¹⁷⁾。また図1のように、誤用論文の45%が多重性の問題を起している¹⁸⁾。ここで言うt検定の乱用とは、経時測定データで行なわれている所謂、t検定による輪切り検定のことである。この誤用だけを防ぐようにしても学会誌の統計誤用率は一度に大幅に改善されよう。

輪切り検定には、多重比較の方法を考えるべきである。多重比較は文献19)、その準拠ソフトはYukms(サイエンティスト社)がある。翻訳本で

は、文献20)のDunnettの「多重比較：いつ、どのように行なうか?」が参考になる。教育・心理の統計教科書として文献21)がある。洋書では理論はMiller(1981)が有名で見やすい²²⁾。このMillerの本の概略は佐久間(1970)の解説がある²³⁾。計算の実際にはZar(1984)²⁴⁾が分かりやすく研究者に一冊薦めたい。

多重性の問題点は、サイエンス誌にTurkey(1977)²⁵⁾が警告してから、日本では藤田・椿・佐藤(1986)の臨床評価誌上の論文¹⁶⁾、海外の臨床の場合では、例えばSmithら(1987)によって議論されるようになった²⁶⁾。

ただ、多重比較も無批判に使うべきでない。多重比較の問題点は指摘されており²⁷⁾、多重比較の批判は、Chew V. (1976) (1976) (1980)にあり^{28~30)}、臨床では文献31)に各手法の長所と短所が述べられている。多重比較の普及がなされると今度はその誤用が問題となろう。例えば既に、対比(contrast)を使うべきを、無批判に対応あるデータに多重比較を適用する例が発表されている。測定前の状態

全日本鍼灸学会誌の統計誤用調査(1982~1991年)

(誤用論文総数 164, 誤用の重複あり。
解析法・ソフト名無記載のみ 1/33。)

図1 全日本鍼灸学会誌約10年分の統計誤用調査のグラフ表示

と各時点との比較は Dunnett の多重比較が使える。それはそれらに相関がないからである。研究者が探索的に時点毎に対応ある t 検定を行なうのは有用だが、タイプ 1 のエラーを考慮せず、それを公表するのは問題がある。

念を押したいことは、多重比較を使いすぎ、全時点で群間比較をして検定し、星印を一杯付ける必要が本当にあるのか、研究者は今一度考える必要があろう。

② 一元配置分散分析の誤用

過去の全日本鍼灸学会誌における繰返し測定の解析では、自由度調整付 ANOVA (分散分析) で多くは対応できる可能性がある。以下は日本温泉気候物理医学会誌の誤用例をあげ、再計算の結果を示す。

対象は同一方法で、なんと 70 余りの一元配置表の解析を行なっている 88 頁に及ぶ長論文である。特定経穴を刺激した時の脈波波高を経時的測定したものである。行が繰返しデータで、データ間に対応がある。故に、くりかえしのない二元配置で計算して、時間変化をみるという方法でもよいが、行が独立を前提にした一元配置の分散分析は使えない。1979 年掲載のこの論文は当時統計誤用指摘

が少なかったため、この誤用は止む負えないと思う。そこで図示する効果も狙って、図 2～4 のように、原論文の表をグラフに書換えた。これを見ると時間とともに各経穴の脈波波高の変化は変動しており、経穴群間に差があるように見えるものとそうでないものがある。当該論文は健康者 5 名、気管支喘息者 5 名で肺経の原穴(太淵)を刺激したものである。図 2 はそのうち健康者 5 名のデータだと考えられる。原論文では表 9-1-a に相当する。しかし、測定値は 5 名の平均なのかどうか記載がない。この 5 名をどう使ったのか分からない。このような測定条件ですでに問題があるが、それは正しいとしてここでは検定だけを問題にしたい。

結果は、単なる一元配置の分散分析で 1 % 有意だという。しかし、繰返し測定データの分散分析でやり直してみると、 P 値は 0.0001 で自由度を調整すると $P=0.0062$ (H-F 修正済) となり、結果的には同じになる。図 3 の健康者で非経穴での脈波(原論文の表 9-2-a 相当)も 1 % 有意としているが、同様に計算すると P 値は 0.0106 で H-F 修正では 0.1336 となり、有意差は認められない。次に図 4 の患者での経穴での計算(原論文の表 9-3-a 相当)も 1 % 有意だというが、 P 値は 0.0018 で自由度調整をすると、0.0187 (H-F 修正) となり、5 % の

肺経経穴上の脈波波高経時変化

図2 統計誤用例 1 (原論文の表 9-1-a をグラフ表示)

図3 統計誤用例2 (原論文の表 9-2-a をグラフ表示)

図4 統計誤用例3 (原論文の表 9-3-a をグラフ表示)

有意水準に改める必要がある。よって、残り 70 前 変量解析を考える。
後の分散分析結果も再計算の必要がある。

時点と治療の間に交互作用がある時は、次の多

1.2 多変量分散分析(MANOVA)

繰返し測定分散分析で、時点と治療の間に交互作用がある時は、多変量分散分析を考える。しかし、これは正規分布と等分散が前提になる。これを無理な前提と見る人がある。経時測定データは各時点毎に横断的に別々の特性と考えれば多変量データとみなせる。MANOVAは一種のスクリーニングとして行なうのがよいといわれている⁸⁾。それで有意差が出れば、個々の特性のANOVA、共分散分析(ANCOVA)を行ないMANOVAで有意差が出なかった場合は、ある特性で有意差が出ても割り引いて考えるべきであるという。MANOVAによるプロフィール解析も、詳細な検討に先立つスクリーニングの要素が強い。

分散共分散構造が満足されなければ、乱塊法や分割型のANOVAは正確でなくなる。その時MANOVAの適用を考える。このデータ構造の検証に使う測定分布の形、分散はヒストグラムで確認できる。これで外れ値はあるか、平均や標準偏差のパターンは何かを確認する。これらのモデルの選択の手順は例えば後述するソフトウェアBMDPのマニュアルやテクニカル・レポート(BMDP Technical Report No.83)に詳しい。ただMANOVAは、時点数が自由度を越える場合は適用できない。これは良質のソフトではソフトの方でエラー・メッセージとして表示される。

MANOVAの時点数が症例数を越えているとき、時間と群・時間交互作用のテストが不利になる欠点を補って、randomization testを利用したANOVAが提案されている³²⁾。

2群以上の繰返し測定の探索的解析を示した例では、試験と群の交互作用を考慮しながら、多変量の上に1変量アプローチを適用する方法がある³³⁾。

いわゆる輪切り検定が横行しているので、データの独立性、多変量正規分布、そして共分散行列の同質性という前提はあるが、一度に多くの変数の平均値を個々にt検定でやらないで、その多変量版であるホテリングのT²検定がもっと使われてよい。1標本のホテリングのT²検定では、サンプル数が時点数より多く、時点が等間隔である必要があるが、多重性を考慮した信頼区間を構成できる。2標本のホテリングのT²検定ではプロフィール分

析で、平行性の検定・レベルの差の検定・繰返し測定間の平均の差の検定を行なう³⁴⁾。

1.3 共分散分析(ANCOVA)

共分散分析は無作為化試験では精度を高めるため、観察実験では無作為化ができなかったための交絡因子の調整のために使われる。

具体例でいうと、観察実験で血圧が2群で有意差があった。しかし、よく検討すると2群には年齢で大きく食い違いがあった。血圧の差はこの年齢の影響ではないかということで、年齢という交絡因子を調整して2群を比較するのが共分散分析という方法である。

共分散分析は分散分析と回帰分析を合わせた方法であるから、経時測定データにも使えると思うのは当然であろう。この共分散分析を経時測定データに使う工夫をしたのは、Kenward(1987)である³⁵⁾。この具体的計算方法は、文献36)に分かりやすく紹介されている。この方法はプロフィールの比較をするが、繰返し測定のMANOVAのプロフィール分析より効率的で、かつ脱落の対応が容易に可能な方法である。

1.4 自己回帰(AR)モデル

Rosnerら(1985)によって紹介されている自己回帰モデルの方法は³⁷⁾、しばしば経済学モデルで利用されている方法で、SASやBMDP(5V)が利用できる。時点間の相関を表現する線形モデルを使った方法である⁶⁾。

単行本Lindsey(1993)にもこのモデルが薦められている³⁶⁾が、ただ自己回帰モデルは、定常性といって、ある時点での系列値がそれ以前の一定個数の系列の値に依存しており、しかもその依存の仕方が各時点で同じであるという条件を満足しているかどうか前提になる³⁹⁾。この定常性というのは、時系列データでは調べたい特性値に対して観測値が不足しているから、発想を変えて対象とする特性値の数自体を少なくする想定である⁴⁰⁾。

今まで、時系列解析では自己相関を求めるのに少なくとも50、望ましくは100以上の時点数が必要とされてきた⁴¹⁾。しかし、Rosnerら(1985)は、いくら時点数が少なくとも自己回帰モデル(AR)が

適用できるという⁴²⁾。この方法は、過去の影響を誤差に含まず、一つの説明変数とするのでそれに対しても、有意性検定ができる。そして2種データでも適用でき、不完全データ(incomplete data)にも対応でき、複雑な最尤法でなく普通の最小自乗法を採用すると、BMDPなどのソフトが容易に利用できるというわけである。

1.5 時系列解析

時点数が多い場合、時系列解析を使う。しかし、時点数の少ない鍼灸の臨床試験では、使う機会が少ないだろう。

時系列解析は、経済学の分野で発達してかなり便利な方法がある。移動平均は項数が任意なため、前項の自己回帰モデルが使われる。

ただ、理論的に理論モデルを設定し、測定データによって検証するという仮説演繹法に慣れている研究者にとって、時系列解析はとっつきにくい。その理由は、時系列解析の手続きはデータから逆にモデルを同定するという帰納的手続きの為ではないだろうか。だからよけいに自らがデータ解析を経験しないと理解しづらい面がある。ARIMA (autoregressive integrated moving average) は特にその感があると思う。

この対策としては、できれば信頼性の高い時系列解析専用ソフトウェア(例えばTSPやRATS等)で、研究者自ら解析し、慣れるしかないと思われる。

一事例の研究で時系列解析を使うには、Gottman & Glass(1978)の説明がある⁴³⁾。C統計は⁴⁴⁾、サンプル数が8個で正規分布が保証されるということであるから、時点数が少ない鍼研究に向いている。これは傾向性の検定といわれているもので、サンプル系列が無作為かどうか調べているとも言える。

時点数が多くなれば、単純最小自乗法で求めた回帰方程式の係数の安定性を確認することによって、構造変化があったかどうかみるチョウ(Chow)テストがある⁴⁵⁾。これは計量経済学ではよく使われるが、F検定の応用である。

自己回帰(AR)モデルとこの時系列解析の本は、たくさん出ている。経済時系列の本は門外漢には何が良いのか分からない。しかし各自専門書参照

ではあまりに無責任だと思う。そこで次の二点を挙げたい。コンピューターで出力されるコレログラムの読み方の例がある比較的分かりやすい本ということで文献38)、ARIMAモデルのBox-Jenkins法については、Vandaele(1983)の訳本⁴⁶⁾が出ている。生物統計学としての時系列の本は文献47)参照のこと。

1.6 生存分析

生存分析の歴史や応用については、文献48)を参照されたい。これはSASの解説書であるが参考書として優れている。洋書の計算例の本は、例えばMatthew & Farewell(1988)⁴⁹⁾が定評があり分かりやすく、またもう少し突っ込んだ議論では文献50)も見やすい。

生存分析においては、不可逆のエンドポイントが選ばれるべきという。七堂(1991)はこの方法でシミュレーションしたとき、緩解までの時間をとった⁵¹⁾。しかし、厳密には再発のある緩解や軽快をエンドポイントにするのは適当ではない。では発想を変えて、再発までの時間をとってはどうか。これは不可逆過程である。そこでこの解析方法を使うという発想になる。また、鍼治療は疾患や重症度によっていわゆる「戻り現象」があるとされている。これは鍼の問題ではなく患者側の問題で、再発と言い換えられる。この再発時間データは分布型について十分な情報がない。再発回数や再発期間を扱うとすると、並べ替え検定が役にたつ⁵²⁾。また、並べ替え検定に基づいた再発率比の信頼区間も求められるようになった。

解析対象となる生存時間とは、病気が楽な時間、鎮静期間、失敗は再発や死を意味する⁵³⁾。生存分析を生存時間以外に応用したのは、Petoら(1976)⁵⁴⁾・(1977)⁵⁵⁾によって最初に解説された。

時間データの欠測値に関しては信頼性工学に補填の方法が確立している。もともと信頼性工学の考え方、例えば自動車の耐久性試験を調べたりする考え方を、生体を相手にする医学に応用して発達してきたものが、生存分析である。1960年前後から医学特有の解析法が開発され出した。そして特に癌の臨床試験で使われている。癌の臨床試験のまとまった本は文献56)、邦文では文献57)を参

照のこと。生存分析という言葉で生死以外使えないというイメージは誤解である。

いずれにしても生存分析の特徴は、時間間隔が解析対象となる。

生存曲線の比較には、脱落者が多い場合適用できる Kaplan・マイヤー法 (Kaplan and Meier 1958)⁵⁸⁾ や、一般化 Wilcoxon 検定、ログランク検定などがある。これらは単変量であるが、生存時間分析の多変量解析では生物統計学の標準手法となったセミパラメトリックな解析方法である Cox 回帰 (Cox 1972) がある⁵⁹⁾。これらについては Dawson-Saunders & Trap (1990) の訳本参照のこと³¹⁾。

1.7 成長曲線モデル

時間を説明変数、測定値を応答変数として誤差項の正規性の仮定のもとで主として多項式回帰モデルが当てはめられるのを成長曲線モデル解析法という。脱落や欠測に対する方法も考慮されている。またパラメトリック、ノンパラメトリックな比較法が開発されている²⁰⁾。

Matthews ら (1990) によって、医学データをベースライン値を基準として、ピークと曲線から時間的変動パターンの面積 (the area under the curve: AUC) を計算し、個体毎の平均値の差の検定を行なう実際的な方法が提案された⁶⁰⁾。これは簡便な良い方法であるが、この方法の限界は、変動パターンが単調であれば問題無いが、増加と減少を繰り返す場合は検出力が低下する。

1.8 2 値データのロジスティック回帰モデル

2 値データとは、生・死、有効・無効、成功・失敗などのように 2 値の尺度の応答変数 (outcome variable) を持つ場合、ロジスティック回帰モデルを使う。

ロジスティック回帰モデルの考え方については文献 61), それを含む一般化線形モデルは、文献 34) に詳しい。医師の書いた本としては、高橋 (1995) のものがあり、Cox 回帰も書かれていて現場で苦労した痕跡が見える⁶²⁾。

短期の 2 値データで経時測定データの解析法として、Working likelihood analysis と、general-

ized estimating equation (GEE) 解析が非常に flexible な方法として薦められており、N. Lange et. al. 編の本に、J. Graham, A. J. Petkau (1994) が詳細な例題を載せている (データが添付ディスクに収められている)⁶³⁾。GEE は、ロジスティック・モデルを拡張したものである。

宮原・丹後編 (1995) によると、まだこの方法は理論的にも計算的にも十分整備されていないという⁶⁾。しかし、MANOVA 適用には症例数が足りない、ANOVA では、十分に詳細な欲しい結果が得られない場合は考慮すべき方法であろう。

1.9 混合モデル (クラス分けモデル)

混合モデルについては、Scheffe (1959) に詳しい⁶⁴⁾。治療法を固定し、被験者を変量、時期を母数とすると、このモデルが使える。混合モデルでは、経時測定データの相関構造を系列相関でなく交互作用として捉える。

カテゴリカル・データで、なおかつ繰返し測定データで、しかも分布パターンが、群内でも平行的でなく、個体毎に反応のパターンが異なっている場合の解析に、広津 (1989)⁶⁵⁾、や丹後 (1989)⁶⁶⁾ の方法が注目されている。反応に幾つかのクラス分けが出来るものがあるとするクラス分けモデルである。広津 (1989) の方法はクラス分けした後、分割表解析が行なわれるので、ある程度の自由度が残り、そこでは主観的な判断が必要であるという。丹後 (1989) の方法は、クラス分けと各クラスでの反応曲線のパターンを主観的に決めている。そこにはかなり自由度がある¹⁾。事前に鍼の有効性について、クラスの型を決めるという作業は行なわなければいけないだろう。

被験者と時点の交互作用が無視できれば、今までの分散分析では、枝分かれ配置の分散分析 (繰返し測定の ANOVA と同義) が適用できる。被験者と時点の交互作用を積極的に評価するとそれは、経時変化パターンの個体差が考慮でき、今までの検定法では、群の平均値しか問題にしなかった、あるいは多重比較でも横断的な点の比較であったわけである。このクラス分けの方法は、単一事例 (N of 1) の方法論が強調していることだが、集団のアプローチにない利点としてあげている個体差

を検討する個のアプローチにもなる。

ただ一事例の研究方法について、特に一般化(外的妥当性)の構造で、小牧(1992)が問題点を指摘している⁶⁷⁾。すなわち、一事例で言う個体内変動を見るという行為は群型研究法でいう誤差分散を調べることと同じ、また一事例でいう要因作用を見る行為は、有意差検定の方法と同じではないかという指摘がある。個体型と群型の二つの研究方法は決して対立するものではなく、個体型は群型の範疇にはいるという見解でもある。個体型研究法は、一般的な群型研究法と共通の言葉でないところに誤解が生じやすい。一事例の方法に限らず一般的でない方法を使うときは、それに対する賛否両論を読んで、その長・短所をしっかりと把握して使用しないと独善的になる可能性がある。

丹後(1989)の方法⁶⁶⁾の基本モデルは、原点を通る3次までの低次の多項式を設定し、混合分布モデルのパラメーターの最尤推定値は、E-M アルゴリズムを利用している。E-M アルゴリズムは不完全データに基づく最尤推定に際し比較的簡単な手順で尤度最大化を図る数値上の技法であり、宮川(1987)の説明が分かりやすい⁶⁸⁾。最尤推定法はフィッシャーが確立したもので、生存率などの線形モデルのなじまない非正規分布で推測に用いられる道具である。最尤法は例えば、文献 69) 参照のこと。限定、非限定尤度推定の計算のためアンバランスな繰返し測定で、Newton-Raphson, Fisher scoring, generalized EM の3つの一般的なアルゴリズムについては文献 70) に述べられている。

丹後(1989)論文⁶⁶⁾には、広津(1989)の方法⁶⁵⁾との同一データの解析での結果の比較が行なわれ、改善群、不変群、悪化群の3群に分類されるデータに相違した5名について、臨床的に検討すると興味ある結果が見つかるのではと言う。その相違は大きな差ではない。

ただこれらの方法は、モデルの適合度の客観的評価は未だ為されていない。

クラス分けの方法のソフトはまだ無いので、丹後の方法ではE-M アルゴリズムの計算が必要になり、非線形最適化プログラムの例えばソフトウェア ASNOP が簡便だという⁶⁶⁾。

1.10 random effects model

本方法は、Laird and Ware(1982)が紹介している⁷¹⁾。反応プロフィールをモデル化したもので、母数効果モデル(fixed-effects model)では、個体差と誤差を誤差と見なしてしまうが、個人毎のパラメーターの個体差が母平均と各群に共通な分散共分散行列を持つ正規分布に従うという仮定をする random effects model(変量効果モデル)があり、欠測値にも対応できる点はよいが、モデル化に工夫があるので、統計の専門家に相談する必要がある。

1.11 randomization test

単一事例実験デザインでは必ず繰返し測定データとなる。データには自己相関が考えられる。自己相関つまり系列依存性があっても、標本分布に影響しないで、使える検定法として、randomization test がある。計算上の困難を避けるため、t や F 分布の近似が使われている。ただこの方法は、AB デザインで、A から B へまた B から A へ素早く変化するものでなければ、つまり持ち越し効果や鍼に反応する疾患の種類により、それらの変化が少ないと AB 間の差は検出できない⁴¹⁾。

このような単一事例デザインだけでなく、もちろん最初から無作為割付された群間検定でも今後普及するだろう。

検定方法は、近似値を使わなくとも簡単なプログラムでモンテカルロ法による組み替え法(random permutation test あるいは randomization test)で容易に計算できる(中谷 1983)⁷²⁾。これは正規乱数を発生させてデータを群間に無作為に振り分ける。その他の色々なノンパラメトリック法を組み替え法でシミュレーションから有意水準を求める判定 BASIC プログラムは、篠原(1989)にある⁷³⁾。

或はサイエンス社の Yukms Ver.5 でも、wilcoxon の順位和検定を直接確率として計算する non-parametric permutation test のプログラムが追加された。単一事例デザインの Nonparametric randomization tests に関しては、Levin ら(1978)に詳しい⁷⁴⁾。

この検定方法は分布型不明でも使える、直接確率が計算されるので統計数値表は不要、そして一

般の検定は、厳密には患者が無作為抽出サンプルでもないのに、それでも現実に役立つという態度で母集団を想定しているが、この方法にはその必要もなく、論理的なうしろめたさもなくなる。現実データは、ある病院に来る患者を対象に臨床試験や臨床研究を行なっているわけであるから、母集団からの無作為抽出という条件は現実的ではない。

randomization test を使う時は、無作為割付という前提条件だけでよく、無作為抽出という前提条件は不要となる。

randomization テスト・アプローチを群間に応用するときは、各々の被験者のための応答プロフィールのいくつかの要約された機能を前もって選択する必要がある。

具体的には、対数線型モデルのように、randomization test の方法では群間、時間変化、群・時点交互作用まで一度に検討できない。しかし、対数線型モデルでは多変量を推定するために、その分大きいサンプルサイズが必要となる。randomization test は、時系列解析よりはるかに使いやすしいといった利点があり、今後多用されるだろう。

浜田ら(1994)によって、並べ替え検定は、SAS によるプログラム例が掲載されている⁵²⁾。

2. 質的データの為の繰返し測定データ解析のための文献レビュー

カテゴリカルデータの為の繰返し測定データ解析のための100以上の論文レビューがAshbyら(1992)によって行なわれており⁷⁵⁾、本論で触れられなかった繰返し測定に関する論文は、これを参照されたい。

単行本では、文献(47), (76), (77)そして(78)がある。

順序尺度の繰返し測定におけるノンパラメトリック検定は、同一被験者の単一治療の前後における検定にはWilcoxon signed-rank 検定、同一被験者の多時点測定にはFriedman 検定が使われる。Friedman 検定における処理間の差は多重比較が必要で、白旗編(1987)にはScheffe のS法とTukeyによるT法の多重比較が示され⁷⁹⁾、そのソフトも同出版社から用意されている。他にFriedman 検

定後の群間比較はStudent-Newman-Keuls 検定(対の比較)とDunnett 検定(対照群と各群の比較)が使える⁸⁰⁾。

2.1 マルコフモデル

時系列を確率論的に考察するとき、確率過程という。n 回目の事象の生起確率がその直前の、n-1 回目の事象の生起の結果に依存し、それまでの事象の生起の結果には依存しない場合、マルコフ過程という。マルコフ過程の特殊な例をマルコフ連鎖といい、分岐点の状態が等しければ、段階が同一であるか異なっているかにかかわらず、わかれ方の確率の値が一定であるようなマルコフ過程をいう。そして状態の数が有限の場合、有限マルコフ連鎖という。また、一度入れば抜け出せないときを、吸収状態と呼び、少なくとも一つの吸収状態を含む過程を、吸収マルコフ過程という^{81, 82)}。

有限マルコフ連鎖を用いた薬効評価法については、広沢(1979)の報告⁸³⁾がある。

このモデルでの手計算の例は、計算ミスの多い本だが水野(1978)にある⁸⁴⁾。生存研究にマルコフ・モデルを適用した例は、Kay(1986)にある⁸⁵⁾。これは癌のマーカー(α -フェトプロテイン)のレベル変化を時系列で見たものである。Cox(1986)はセミ・マルコフ過程で生活の質(quality of life)を評価する方法について触れている⁸⁶⁾。

一事例の研究にマルコフ連鎖を利用した方法は、Gottman & Norarius(1978)に詳しい⁸⁷⁾。

ソフトはGAUSSの応用ライブラリーMARKOV等の時系列専用ソフトで計算するか、対数線型モデル用ソフトGLIMでも可能。

ランダム・ウォークといって、酔っ払いの行き先を計算する方法がある。ランダム・ウォークモデルを使った経時測定の臨床試験の評価方法は、Hopper(1988)により紹介されている⁸⁸⁾。

2.2 田口の累積法

鍼の臨床評価は時系列で順序尺度になりやすい。この場合、以前から田口(1977)の累積法⁸⁹⁾が提唱されている。これは一様性検定であるが、時系列多重比較については、斧田ら(1986)によって、重み付平均の工夫が提唱されている⁹⁰⁾。データを累積し

ていく、この累積法のアイデアは高く評価されている(広津 1994)⁹¹⁾。

田口の累積法の問題点は、中村(1988)⁹²⁾、Hirotsu(1990)⁹³⁾、Hamada & Wu(1990)⁹⁴⁾にある。それは広津(1994)によると、直交表において多くの要因に有意差が認められた時は自由度の調整、疑似交互作用、要因平方和が大きければ、必然的に誤差は過小評価になるといった点である⁹¹⁾。

田口の累積法は、日本規格協会のソフトウェア、ANOVA Ver.5 で対応できる。

筆者らは 1980 年、慢性肝疾患に対する灸の臨床試験を報告した⁹⁵⁾。15 年前だから、当時パーソナル・コンピューターは容易に手に入らなかった。当然適当なソフトも見あたらなかった。当時は全部電卓で、ノートを何冊も使って夜遅くまで病院内で計算した。だから 1 ヶ月毎、6 ヶ月間の試験期間における複数症状の変化を一度に計算するのは無理があった。そのため私たちのこの報告では、将来再計算可能な形で、原文内の表 19 として全部のデータを掲載した。今ではこのソフト ANOVA Ver.5 があれば、容易に計算できる。

また、線型モデルによる多変量データ(累積ロジット・モデル)として、Grizzle, Starmer and Koch の方法で同じデータを解析できる^{34,96)}。また順序カテゴリーのある対数線形モデルを使うと、時系列データでない順序カテゴリーのあるデータの田口の累積法による解法(増山ら 1974)⁹⁷⁾の代用になる。

田口の累積法を使って、海外へ投稿する時、その一般性を心配する人がいると思う。英文の Taguchi Methods という本が ASI Pr. から発行されているので、それを確認後引用して方法を説明すれば良い。

2.3 判別分析

判別分析を繰返し測定に応用したのは、斧田(1988)である⁹⁸⁾。具体的には、慢性関節リウマチの薬効検定で、杉山ら(1969)⁹⁹⁾、塩川ら(1974)¹⁰⁰⁾(1976)¹⁰¹⁾によって報告されている。

判別分析のソフトは、手軽なところでは、HALBAU Ver.4(現代数学社)であるが、ロジスティック解析の自動変数選択機能といった欲しい機能が

ない場合があるらしく⁶²⁾、始めから SPSS, BMDP を利用した方がよいかもしれない。

2.4 因子分析

経時測定で因子分析? といふかるむきもあるかもしれないが、鍼研究でも時に使われている semantic differential (SD) 法がある。尺度は様々考えられるが結局、被験者×尺度×概念という 3 相データになることが多い。

吉野ら(1995)¹⁰²⁾が 3 相因子分析を経時測定データで解析している。質問紙法で得た順序データから、被験者相・尺度相・時点相の 3 相に分け、すべての時点のすべての被験者をひとまとめにして、平均が 0、分散が 1 になるようデータの基準化を行ない計算する。時点相因子負荷行列で負荷の高くなっている時点を見るとそこで症状の変化が起っているとみる。

村上ら(1978)にこの因子分析運用上の問題点が書かれている^{103,104)}。

ただ因子分析であるので、相対的にしか物がいえないので、全体的な動きを捕まえ仮説を構築するための探索的方法として使用すべきである。

2.5 Randomization model の方法

Landis ら(1988)は、順序スケールの多変量の繰返し測定データに randomization model の方法を適用して、SAS の FREQ と CATMOD で計算できることを示している¹⁰⁵⁾。ただ限界同義性(対称)のテストのためには、一般化マンテル・ヘンツェルテストの運用が考慮されている。

randomization model に関しては、Scheffe (1959)⁶⁴⁾に詳しい。

randomization model では、randomized blocks とラテン方格デザイン等の正確な検定のための randomization test (permutation test) の使い方についても、Scheffe (1959)が書いている⁶⁴⁾。

2.6 対数線型モデル

対数線形モデルの解説には、田中・垂水・脇本編(1990)³⁴⁾がソフトも用意されていて分かりやすい。分割表の対数線形モデルを含む一般化線形モデルの解説書として定評のある本には、Dobson

(1990)⁹⁶⁾がある。

Koch ら (1977)⁷⁾ によって群間、時点間比較をするために考え出された線形モデルの方法で、例題が載せられている。Crowder & Hand (1990)⁷⁶⁾ の 118~124 頁には、Koch ら (1977) の例題を使って、群間および時点間比較のための対数線形モデルについて解説がある。

対数線形モデルの計算には、ソフトウェア SPSS-X の LOGLINEA、SPSS や、HILOGLINEAR、LOGLINEAR が、かえて専用ソフト GLIM や人気の SAS より使いやすいという (松田 1988)¹⁰⁶⁾。

3. クロスオーバー法

これはデザインからいうとラテン方格法の応用であり、繰返し測定の実用でもある。本法を使った鍼の臨床試験では、単に症例数を節約するために (検定の時の症例数を増やすために) この方法を選び、また前後の期間の症例数を単に加算して、群間検定をするという初歩的誤りを犯している例がある。この方法の解析はデザイン、解析が研究者が考えているほど容易でない。ただ、クロスオーバー法の利点は、背景因子の偏りをなくすることができる。

しかし、前期治療の影響が後期に影響するという持ち越し効果 (carry-over effect) の処理が問題となる。そのためクロスオーバー法は、誤用が多く、なるべく使用を控えるようにという意見が強い。

適応疾患であっても、wash-out の期間をどうか、第一治療で完全治癒してしまう場合は困る。理想的には第一治療中止で再発しなければいけない。この再発ということが重要で、クロスオーバー法がむいているのは、糖尿病、高血圧症、アレルギー、てんかん、慢性関節炎、パーキンソン氏病などである。これは患者間より患者内変動の方が小さい場合である。それ以外で強引にクロスオーバー法を行なうと誤った結論に導く恐れがあり無駄である。

鍼で腰痛を治す場合、第一治療で治ってしまう可能性がある。一般に鍼で第一治療中に治ってしまい、その中止で再発する場合は少ないと考えられ鍼の臨床試験では疾患の選択で、その成否が決

まってしまう。そこで WHO の鍼臨床研究の方法論のガイドライン (Guidelines for clinical research on acupuncture 1995) では¹⁰⁷⁾、このクロスオーバー法は鍼研究に向いていないと書かれている。

解析方法は途中で治療を交替する方法であるが、これも経時測定データである。一事例 (N of 1) 研究デザインも見方を変えれば、一事例のクロスオーバー法とも言える。ただ単一事例研究法には、群間比較に比べ一般化 (外部妥当性) に問題がある。

Senn & Hildebrand (1991) は¹⁰⁸⁾、前後期の試験結果の比較を Scheffe (1953) の対比 (contrast) による平均値の比較で対応することを提唱している¹⁰⁹⁾。興味ある方法であり、この方法に対し考えられる批判には彼らは丁寧にその論文で答えている。ただこれは後知恵によって探索するので有意性検定を行なう場合は、Scheffe 多重比較の基準を考慮する必要があるだろう。

2 期のクロスオーバー法の計算の仕方は、Hill & Armitage (1979) に分かりやすく詳しい¹¹⁰⁾。狭心症、関節リウマチなどの慢性疾患の 2 回交差試験には、斧田 (1975)¹¹¹⁾ の解説がある。複雑なクロスオーバー法や他の文献については、Crowder & Hand (1990) にある⁷⁶⁾。

クロスオーバー法は、日本規格協会のソフトウェア、田口の ANOVA Ver.5 で対応できる。また、順序効果を無視するという仮定を置けば、対数線形モデルでも解析できる (田中ら 1990)³⁴⁾。GLIM では、グループ数×プロフィール数という応答データが独立した変数でポイズン分布する場合、使える。

4. 左右対称で複数個あるデータの解析法

鍼治療成績で、ある人の膝を問題にするとき、右膝と左膝を独立と考えて処理する例が見られた。同一人の左右の膝は独立とはいえない。右膝の治療は当然、片足の人を除いて、悪い方をかばい反対の左膝に影響するだろう。だから例えば 10 人の被験者で膝 20 として検定するのは間違いである。このような相関あるデータの解析方法として、混合効果モデルを適用するとよい。

眼科領域では、同一人の左右の眼は独立変数ではない。この評価をするときの方法がすでに、Rosner

(1982)によって提示されている¹¹²⁾。

鍼による変形性膝関節症の患者の治療評価を、左右の足を独立と考慮して検定していた実例がある。これはサンプル数を増やすという目的が考えられ、鍼研究での実例であるが、2期2群比較のクロスオーバー法で元の症例の2倍にして単純な2群比較の検定を行なっているのと同じ発想である。

このほか同一人の手腕や足腿、耳あるいは赤羽氏法のような左右対称点の測定値解析には、同一人の左右の部位は独立していないことを考慮して解析すべきであろう。

左右対称で複数個あるデータの解析には、データの独立性を考え、左右対称対象から得られたデータに相関がある場合は今まで述べてきた混合(mixed)モデルを利用する。

5. コンピューター・ソフトウェア・パッケージ

Lindsey(1993)は、自国(英国)の generalised liner interactive modelling (GLIM) パッケージで主に繰返し測定データを解析している。PC用の GLIM は、英国王室統計協会が提供している低価格のソフトウェアで、業者を通して10万円以下で買える。ただこれは線形モデルに関してある程度熟知している必要があると言われている⁶⁾。Statistical Analysis System (SAS) や Statistical Package for the Social Science (SPSS) でももちろん経時測定データは解析可能である。SAS では、SAS GLIM, SAS ANOVA, SAS CATMOD というプロシージャが使われる。

短期2値データで経時測定データの解析法として、Working likelihood analysis と GEE 解析が非常に flexible な方法として薦められている(N. Lange et. al. Ed.; J. Graham, A. J. Petkau 1994) が、SAS の GLM プロシージャ用に GEE マクロ (John Hopkins GEE SAS macro) が用意されており、e-mail [sz@zegger. sph. jhu. edu] を経由して申し込み手に入れる事が出来る。SAS は日本でのサポート体制も十分に企業・研究施設ではよく使われている。しかし、リース方式でコスト面で個人購入は必ずしも容易でない。SPSS では、SPSS MANOVA が中心となろう。元々日本語の解説書が出ていたが、SPSS 自体の日本語版が出来、

Windows 版のお蔭で益々使いやすくなったという。

使用している Biomedical Computer Program Package (BMDP) は生物統計学用に開発されているだけあって、繰返し測定のための必要な機能は揃っている。BMDP は、PC 用ソフト中、生存分析の症例数は最大規模扱えるという³¹⁾。

BMDP は日本語の解説書も皆無で、ソフトは英語版しかないのが欠点である。BMDP のマニュアル (Dixon ed. 1992) は分かりやすい¹¹³⁾。これは大変良い生物統計学のテキストでもある。日本語訳が待たれる。BMDP は買取場合のコストの低廉さや多機能、そしてソフトのインターフェイスの使いやすさの面を総合的に比較すると、他の追従を許さない面があると思われる。BMDP も Windows 版が出来、より使いやすくなった。

BMDP では、分散分析の V シリーズで、2V, 3V, 4V, 5V, 8V が繰返し測定データの解析に使われる。その中味は 2V は繰返し測定がある場合の母数モデルの一元配置分散・共分散分析、3V は混合モデル、不釣り合い型モデルでの最尤推定法、4V は 2V の多変量への拡張、5V は欠測値のある場合などの unbalanced で共分散構造を指定した繰返し測定の分散分析、時点間の相関を表現する線形モデルはこれを利用する。現場データは欠測値があるのが当たり前と考えると、SAS の GLM もこの機能があるが、SPSS や GLIM にはこの機能がない (Jennrich, Schlucher 1986) のは困る⁶⁷⁾。BMDP 8V は一般混合モデル ANOVA、欠測値のある場合の不釣り合い繰返し測定モデルが揃っている⁶⁾。random effect モデルは BMDP 3V で対応出来る。

ただ、対数線型モデルの解析で言えば、BMDP では 4F だが、SPSS より、あるいはもちろん GLIM ほど込み入った解析は出来ない。多変量(対数線型)モデルを使いたい時の E-M アルゴリズムの計算は BMDPAN が使える。最尤推定で使われる Newton-Raphson 法は、GLIM に備わっている。同じく反復比例あてはめ法 (iterative proportional fitting) は BMDP の P4F で実施できる。

非線型モデルの対応もよくできている。本論で述べてきた必要な計算機能はほぼ揃っている。2値データのロジスティック回帰モデルは、BMDP の

PLR が利用できる。

Mac 用のソフトでは、SuperANOVA が繰返し測定分散分析の機能が揃っている。多変量分散分析や多重比較も可能で、BMDP のような細かい対応は出来ないが、逆に試験デザインをこのソフトに合わせれば、実用上問題無い。また SAS の Mac 版が出たが、200 メガバイトの容量が必要でかつ高コストで個人購入はまだ問題がある。

時系列解析用汎用ソフトウェアとしては、Time Series Processor (TSP) が有名で、和合・伴(1995)による日本語のマニュアル(第2版)が市販されている¹¹⁴⁾。TSP と同じような機能でメニュー選択方式で行なう点が違うだけの MicroTSP(現在 Version 6.6 で McGraw-Hill 発売)があり、Mac や IBM PC 互換機で動く。ソフトウェア GAUSS (Ver. 3) にはマルコフ過程の為の MARKOV ライブラリー等が用意されているという¹¹⁵⁾。また Mac-RATS (Ver. 402) は 10 年以上前に開発されたソフトで、SHAZAM (Ver. 6) 同様多くの人に利用されており、価格は MicroTSP より 2 万円ほど安い。

MicroTSP, RATS, SHAZAM の機能評価の比較は、Cromwell ら(1994)の2文献(一変量、多変量の時系列モデルテスト)にある^{116,117)}。ただ SHAZAM に関して竹村(1994)はダービン・ワトソン比で系列相関をチェックするとき、モンテカルロ法でなく数値計算によっているので、大きな行列の固有値計算が必要で実用上問題があるという¹¹⁸⁾。

従来の統計的仮説検定でなく、情報量統計学というモデルの構成と情報量による評価を行なう領域がある。これは、個々の標本分布論や統計数値表をも排した方法で、想定される最適モデルを赤池情報量基準(Akaike Information Criterion: AIC)で評価する。それには、時系列解析で本邦では Fortran で書かれたものであるが、北川(1993)の提供している廉価なソフトがある¹¹⁹⁾。Fortran で書かれているので MS-Fortran が必要でかつ Laser-Lib 等のレーザープリンタ出力用グラフィックソフトウェアライブラリーが要る。両者の使い勝手の良さは試していないので不明。

BMDP について、BMDP Statistical Software,

Inc. は、1440 Sepulveda Boulevard, Suite 316, Los Angeles, CA 90025; TEL: (310)479-7799, FAX: (310)312-0161

SPSS について、SPSS Inc. は、444N. Michigan Avenue, Chicago, Illinois 60611; TEL: (312)329-2400, FAX: (312)329-3568。日本支社は、TEL: 03-5474-0341。

SAS については SAS インスティテュートジャパン、TEL: 03-3533-6921。「マニュアル販売係」は TEL: 03-3533-3860、規定の注文用紙あり。

GLIM, GAUSS, MicroTSP, RATS, SHAZAM 等は、国内のいくつかの業者が取り扱っているが、取り扱い方法の説明やバージョンアップなどのサポート体制のある所は少ない。

おわりに

統計教育の不備、市販の統計本・ソフトウェアの不備などで統計誤用が頻発している。その中でも鍼研究では、繰返し測定データに関して約半数の誤用が関係している。この分野はまだ研究途上ということもあるが、鍼灸研究では経時測定データは特殊なモデルではなく、ごく普通の研究モデルである。しかし、この分野は研究の発展にもかかわらず、本邦の実験家・実務家の持っている関連情報量は少ない。そして、実験家と統計の理論家との間には溝があるように思う。それ故、現場で仕事をする実務家のためにソフトと文献のレビューを行なった。現場と届いていない情報(文献・ソフト)との溝がこれで少しでも埋まれば幸いである。

経時測定データの解析には、最初は研究デザインの段階で生物統計学者の応援を受けるべきであろう。実務家は、生物統計学者にせめて相談できるレベルに生物統計学の基礎知識を引き上げる努力は必要である。しかし、生物統計学者の数が本邦は極めて少ないという障害があり急な解決の方向にない。現場では、かなり独学を強いられようが、良質の統計ソフトウェアだとそのパワーが肩代わりしてくれよう。

繰返し測定モデルの選択を誤まらないよう、多重性の問題を起さないよう研究目的にあった解析方法の選定は、事前にデザインされるべきである。

検定方法を決定すると同時に、繰返し測定サンプルサイズの設計も忘れてならない。

統計専用ソフトだけでなく、繰返し測定モデルのサンプルサイズの設計可能なソフトを購入するのが実務家には便利である¹²⁰⁾。くれぐれもサンプルサイズが足りないために有意差が出なかったということは避けて欲しい。

ソフトはとにかく「習うより慣れろ」で使うことだと思う。現場の研究者にとっては、理論は後で勉強すれば良い。新村(1993)によると¹²¹⁾、専門家は統計理論を知らないと誤用するというのが、読者がデータに対して十分な知識と常識でもって解釈した結果が誤用であるなら、悪いのはアルゴリズムの方かも知れないと言っている。そしてアルゴリズムを詳しく知っていること、プログラムを組むこと、ソフトを正しく使う事は別のことだという。統計は実用の学問だから使ってはじめて意義が出てくることを、我々は忘れていけない。メカニズムが分からなくとも、船も車も運転できる。医学部や鍼灸師の統計教育システムが完備していない現状では、現場で問題に直面している研究者にとって、臆せず使うというのが、生物統計学を独学で身につけるためには必要な考え方と思われる。

本研究は、文部省統計数理研究所共同研究(7-共研A-83)の一部援助を受けた。

参考文献

- 1) 吉村 功, 大橋靖雄編: 毒性試験データの統計解析, 地人書籍, (1992).
- 2) Trautwein C.: 製薬企業の統計部門, 管理者の調査. (1981), Buncher C. R., Jia-Yeong Tsay: 医薬品開発過程における統計的方法, 佐久間ほか訳, MPC, 363~369. (1990)
- 3) Origasa H.: Statistical models for ordinal data in a longitudinal study design, Bulletin of the Biometric Soc. of Japan, 8(2), 1-13, 1987.
- 4) Origasa H.: Longitudinal data analysis using linere models, Bulletin of the Biometric Soc. of Japan, 9(1)11-10, 1988.
- 5) 吉村 功: 経時測定データについて, 第13回応用統計学会シンポジウム予稿. (1991)
- 6) 宮原英夫, 丹後俊郎編: 医学統計学ハンドブック, 朝倉書店, (1995), p.245~260, 619~666.
- 7) Koch G. G., et al.: A general methodology for the analysis of experiments with repeated measurement of categorical data, Biometrics, 33, 133-158, 1977.
- 8) 高橋, 大橋, 芳賀: SASによる実験データの解析, 東大出版, (1989).
- 9) Box G. E. P.: Some theorems on quadratic forms applied in the study of analysis of variance problems, I. Effects of inequality of variance and correlation between errors in the one-way classification, The Annals of Mathematical Statistics, 290-302, 1953.
- 10) Box G. E. P.: Some theorems on quadratic forms applied in the study of analysis of variance problems, II. Effects of inequality of variance and correlation between errors in the two-way classification, The Annals of Mathematical Statistics, 484-498, 1953.
- 11) Huynh H. and Mandeville G. K.: Validity condition in repeated measures designs, Psychological Bulletin, 86(5), 964-973, 1979.
- 12) Karl E. Peace edited: Biopharmaceutical statistics for drug development, Marcel Dekker, Inc., 1988. (医薬統計学, サイエンス社, KR研究会訳, (1992), p.342~347.)
- 13) Geisser S. and Greenhouse S. W.: An extension of Box's results on the use of the F distribution in Multivariate analysis, Annals of Mathematical Statistics, 29, 885-891, 1958.
- 14) Huynh H. and Feldt L. S.: Estimation of the box correction for degress of freedom from sample data in randomized block and split-plot desings, J. Educational Statist., 1, 1, 69-82, 1976.
- 15) Ferguson G. A.: Statistical analysis in psychology and education, McGraw-Hill, 316-357, 1981.
- 16) 藤田利治, 椿 広計, 佐藤倚男: 臨床試験における多重性, 臨床評価 14:477~486. (1986)
- 17) 七堂利幸: 全日本鍼灸学会誌に使われる統計手法, 投稿中. (1995)
- 18) 七堂利幸: 全日本鍼灸学会誌における統計処理の問題点, 全日本鍼灸学会誌 42(2):193~197 (1992)

- 19) 吉村功編：毒性・薬効データの統計解析，サイエンティスト社，(1987)．
- 20) Dunnett C., Goldsmith C.: 多重比較；いつ，どのように行なうか，390～393. (1981), Buncher C. R., Jia-Yeong Tsay：医薬品開発過程における統計的方法，佐久間ほか訳，MPC. (1990)
- 21) 森 敏昭，吉田寿夫：心理学のためのデータ解析テクニカルブック，北大路書房，(1990)，p. 157～175.
- 22) Miller R. G.: Simultaneous statistical inference, Springer-Verlag, 1981.
- 23) 佐久間 昭：多重比較，臨床薬理 1(2)：117～122. (1970)
- 24) Zar J. H.: Biostatistical analysis, Prentice-Hall International Editions, 185-205, 1984.
- 25) Turkey J. W.: Some thoughts on clinical trials, especially problems of multiplicity, Science, 18, 679-684, 1977.
- 26) Smith D. G., et al.: Impact of multiple comparisons in randomized clinical trials, Amer. J. Med., 83, 545-550, 1987.
- 27) 吉村 功：多重比較方式の問題点，統計数理研究所共同研究レポート 18：58～71. (1989)
- 28) Chew V.: Comparing treatment means: A compendium, Hortscience 11(4), August, 348-357, 1976.
- 29) Chew V.: Used and abused of Duncan's multiple range test, Proc. Fla. State Hort. Soc., 89, 251-253, 1976.
- 30) Chew V.: Testing differences among means: Correct interpretation and some alternatives, Hortscience, 15(4), August, 1980.
- 31) Dawson-Saunders B., Trapp R. G.: Basic and clinical Biostatistics, 1990. (医学統計データを読む，森田 監訳，メディカル・サイエンス・インターナショナル，1994)
- 32) Raz, J.: Analysis of repeated measurements using nonparametric smoothers and randomization tests, Biometrics, 45, 851-871, 1989.
- 33) Looney S. W., Stanley W. B.: Exploratory repeated measures analysis for two or more groups, review and update, Amer. Statist. 43, 4, 220-225, 1989.
- 34) 田中 豊，垂水共之，脇本和昌編：パソコン統計解析ハンドブック，V 多変量分散分析・線形モデル編，共立出版，(1990)．
- 35) Kenward M. G.: A methods for comparing profiles of repeated measurements, Appl. Statist., 36(3), 296-308, 1987.
- 36) Everitt B. S., Dunn G.: Applied multivariate data analysis, Edward Arnold, 210-213, 1991.
- 37) Rosner, B.: Statistical methods in ophthalmology; an adjustment for the intraclass correlation between eyes, Biometrics, 38, 105-114, 1982.
- 38) Lindsey J. K.: Model for repeated measurements, Oxford University, 1993.
- 39) 鈴木義一郎：情報量基準による統計解析入門，講談社サイエンティフィク，(1995)，p.138.
- 40) 山本 拓：経済時系列分析，創文社，(1994)．
- 41) パーロー，D.H.，ハーセン，M.：1事例の実験デザイン，1984，高木・佐久間訳，二瓶社，(1989)，p.296.
- 42) 折笠秀樹：臨床試験における最近の統計学的话题，Prog. Med. 8(6)：1354～65. (1988)
- 43) Gottman J. M., Glass G. V.: Analysis of interrupted time-series experiments, (Ed. Kratochwill T. R. Single subject research), Academic Press, 197-235, 1978.
- 44) 桑田 繁：新しい実験計画法としての単一被験者法の紹介(II)，全日本鍼灸学会誌 43(1)：36～43. (1993)
- 45) 吉野，高橋：パソコン計量経済学入門，多賀出版，(1990)，p.113～118.
- 46) Vandaele, W.: Applied time series and Box-Jenkins models, Academic Press, 1983. (邦訳「時系列入門」，蓑谷千凰・廣松 毅訳，多賀出版，1988)
- 47) Diggle P. J.: Time series, Oxford University, 1990.
- 48) 大橋，浜田：生存時間分析，東大出版，(1995)．
- 49) Matthew D. E., Farewell V. T.: Using and understanding medical statistics, 2nd. edition, Karger, 1988.
- 50) Namboodiri K., Suchindran C. M.: Life table and their application, Academic Press, 1987.
- 51) 七堂利幸：腰痛に対する鍼灸と整形外科的治療の比較，全日本鍼灸学会誌 41(2)：212～218. (1991)
- 52) 浜田知久馬ほか：臨床評価における再発事象データの統計解析，薬理と治療 22(10)：21～34. (1994)

- 53) George S. L. and Desu M. M.: Planning the size and duration of a clinical trial studying the time to some critical event, *J. Chron. Dis.* 27, 15-24, 1974.
- 54) Peto R. et al.: Design and analysis of randomized clinical trials requiring prolonged observed of each patient, I introduction and design, *Br. J. Cancer*, 34, 585-612, 1976.
- 55) Peto R. et al.: Design and analysis of randomized clinical trials requiring prolonged observed of each patient, II analysis and example, 35, 1-39, 1977.
- 56) Buyse, M. E., Staquet, M. J., Sylvester, R. J.: *Cancer clinical trials*, Oxford University Press, 1984.
- 57) 浜島信之: 無作為割付臨床試験, 癌と化学療法社, (1993).
- 58) Kaplan E. L., Meier P.: Nonparametric estimation from incomplete observation, *J. Am. Stat. Assoc.* 53, 457-481, 1958.
- 59) Cox D. R.: Regression models and life-table (with discussion), *J. Royal. Statis. Socie., B*, 34, 187-220, 1972.
- 60) Matthew J. N. S., et al.: Analysis of serial measurements in medical research, *Br. Med. J.*, 300, 230-5, 1990.
- 61) コックス: 二値データの解析, 1970. (朝倉書店, 後藤・畠中・田崎訳, 1983)
- 62) 高橋善弥太: 医者のためのロジステック・Cox 回帰入門, 日本医学館, (1995).
- 63) Graham J., Petkau J.: Spatial association learning in hummingbirds, (Lange N., et al. ed., *Case studies in biometry*, Wiley, Chap. 9, 159-187), 1994.
- 64) Scheffe, H.: *The analysis of variance*, Wiley, 1959.
- 65) 広津千尋: 経過時点測定データのためのモデルとその応用, *品質* 19:172~179. (1989)
- 66) 丹後俊郎: 臨床試験における経時的測定データの解析のための混合分布モデル, *応用統計学* 18(3):143~161. (1989)
- 67) 小牧純爾: 変動と一般性, *心理学評論* 35(2): 133~155. (1992)
- 68) 宮川雅巳: EM アルゴリズムとその周辺, *応用統計学* 16(1):1~21. (1987)
- 69) 東大教養部: 自然科学の統計学, 東大出版, (1992).
- 70) Jennrich R. I., Schlucher M. D.: Unbalanced repeated-measures models with structured covariance matrices, *Biometrics*, 42, 805-820, 1986.
- 71) Laird N. M., Ware J. H.: Random-effects models for longitudinal data, *Biometrics*, 38, 963-974, 1982.
- 72) 中谷和夫: 組換え法による平均値の差の検定, *Psychological Laboratory Network (PLANET) #2*:13~18. (1983)
- 73) 篠原弘章: 行動科学の BASIC 第5巻「ノンパラメトリック法」, ナカニシヤ出版, (1989).
- 74) Levin J. R. et al.: Nonparametric randomization tests, (Ed. Kratochwill T. R., *Single subject research*), Academic Press, 167-196, 1978.
- 75) Ashby M., et al.: An annotated bibliography of methods for analysis correlated categorical data, *Statistics in Medicine*, 11, 67-99, 1992.
- 76) Crowder M. J. & Hand D. J.: *Analysis of repeated measures*, Chapman & Hall, 1990.
- 77) Lindsey J. K.: *Models for repeated measurements*, Oxford University Press, 1993.
- 78) Diggle P. J., et al.: *Analysis of longitudinal data*, Oxford University, 1994.
- 79) 白旗慎吾編: パソコン統計解析ハンドブックⅣ・ノンパラメトリック編, 共立出版, (1987).
- 80) Glantz S. A.: *Primer of biostatistics*, 3rd ed., McGraw-Hill, 358-371, 1992.
- 81) 安田, 海野: マルコフ連鎖, *社会統計学*, 丸善, (1984), p.274~296.
- 82) 渡辺隆一: マルコフ・チェーン, 共立出版, (1979).
- 83) 広沢毅一: マルコフ連鎖を用いた薬効評価法, *医学のあゆみ* 110(13):952~956. (1979)
- 84) 水野哲夫: *臨床統計学*, 医学書院, (1978).
- 85) Kay R.: A markov model for analysing cancer markers and disease states in survival studies, *Biometrics*, 42, 855-865, 1986.
- 86) Cox D. R.: Some remarks on semi-marcov process in medical statistics, In *semi marcov models, theory and application* (ed. J. Jassen), New York; Plenum, 1986.

- 87) Gottman J. M., Norarius C.: Sequential analysis of observational data using markov chains, (Ed. Kratochwill T. R. Single subject research), Academic Press, 237-285, 1978.
- 88) Hopper J. L., Young G. P.: A random walk model for evaluating clinical trials involving serial observations, *Statistics in medicine*, 7, 581-590, 1988.
- 89) 田口玄一: 実験計画法(下), 丸善, (1977).
- 90) 斧田大公望ほか: 時系列多重比較について, *臨床薬理* 17(1):253~254. (1986)
- 91) 広津千尋: 新しい品質管理のための統計的方法, 数理統計学の理論と応用, 東大出版, (1994), p. 250~252.
- 92) 中村恒夫: データ解析の実際, 標準化と品質管理 41(10):83~93. (1988)
- 93) Hirotsu C.: Discussion on the critical look at accumulation analysis by Hamada and Wu, *Technometrics*, 32, 133-136, 1990.
- 94) Hamada, M. and Wu, C. F. J.: A critical look at accumulation analysis and related methods, *Technometrics*, 32, 119-130, 1990.
- 95) 七堂利幸: 慢性肝疾患に対する灸効果, *日鍼灸誌* 29(3):56~73. (1980)
- 96) Dobson A. J.: An introduction to generalized liner models, 1990. (田中ら訳, 統計モデル入門, 共立出版, 1993)
- 97) 増山, 奥野, 田口, 竹内, 広津: 臨床試験の実験計画(実験計画法), 東大出版, (1974), p. 109~156.
- 98) 斧田大公望: 医薬統計解析マニュアル, MRC, (1988).
- 99) 杉山 尚ほか: 慢性関節リウマチにおける Aza-propazone の薬効検定, *リウマチ* 9(3):169~183. (1959)
- 100) 塩川優一ほか: 慢性関節リウマチに対するニフレム酸の効果, *リウマチ* 16(2):199~210. (1976)
- 101) 塩川優一ほか: 慢性関節リウマチに対する 19583 RP(ケトプロフェン)の薬効検定, *リウマチ* 14(3):279~286. (1974)
- 102) 吉野祥一ほか: 抗うつ薬塩酸トラゾドン投与によるうつ病症状の経時変化, 第 23 回日本行動計量学会抄録集 252~255. (1995)
- 103) 村上 隆ほか: 3 相因子分析の適用上の諸問題, *名古屋大学教育学部紀要(教育心理学科)* 25: 19~39. (1978)
- 104) 村上 隆ほか: 因子分析の記述と準 3 相因子分析, *名古屋大学教育学部紀要(教育心理学科)* 1~16. (1979)
- 105) Landis J. R., Davis C. S., Koch G. G.: Some general methods for the analysis of categorical data in longitudinal studies, *Statistics in Medicine*, 7, 109-137, 1988.
- 106) 松田紀之: 質的情報の多変量解析, 朝倉書店, (1988), p.204.
- 107) WHO: Guidelines for clinical research acupuncture, WHO Regional Publication, Western Pacific Series, No.15, 1995.
- 108) Senn S. J., Hildebrand H.: Crossover trials, designs of freedom, the carryover problem and its dual, *Statist. Med.*, 10, 1361-1374, 1991.
- 109) Scheffe, H.: A methods for judging all contrasts in the analysis of variance, *Biometrika*, 40, 87-104, 1953.
- 110) Hill M. and Armitage P.: Two-period crossover clinical trial, *Br. J. Clin. Pharmac.*, 8, 7-20, 1979.
- 111) 斧田大公望: 慢性疾患における 2 回交差の適用, *臨床薬理* 7(1):99~102. (1976)
- 112) Rosner B. et al.: The use of an autoregressive model for the analysis of longitudinal data in epidemiologic studies, *Statist. Med.*, 4, 457-467, 1985.
- 113) Dixon W. J. ed.: BMDP Statistical software manual, vol.1, 2, University of California Press, 1992.
- 114) 和合 肇・伴 金美: TSP による経済データの分析(第 2 版), 東大出版, (1995).
- 115) G.S. マグラ: 計量経済分析の方法, マグロウヒル, (1992), p.415~417.
- 116) Cromwell J. B., Labys W. C., Terraza M.: Univariate tests for time series models, Sage University Paper, 76-78, 1994.
- 117) Cromwell J. B., Labys W. C., Terraza M.: Multivariate tests for time series models, Sage University Paper, 84-86, 1994.
- 118) 竹村彰通: 多変量解析と計算機, (竹内 啓・竹村彰通編, 数理統計学の理論と応用), 東大出版会, (1994), 127~147.
- 119) 北川源四郎: 時系列解析プログラミング, 岩波書店, (1993).

- 120) 七堂利幸：サンプルサイズの求め方に関する文献とソフト，全日本鍼灸学会誌投稿中
- 121) 新村秀一：意思決定支援システムの鍵，講談社，(1993), p.291.
(〒543 大阪市天王寺区上本町 6-3-31-1111)

Methods of Analysis Sequential Measurement Data an introduction to literature and software

Toshiyuki Shichido

Information and Evaluation Group (IEG)

Research Committee, Scientific Department

Japan Society of Acupuncture (JSA)

Abstract

There is a serial correlation to sequential measurement data, where data is taken over a period of time. Analysis of each of these points in times as an independent unit with its own value will lead to inaccurate results. The sequential measurement model is used quite frequency in acupuncture research. Because misuse of statistics occurs quite commonly in acupuncture research, almost all cases we studied of analysis of this sequential measurement model were in error. Because of this, we undertook to review literature on sequential measurement data and computer software packages (statistical software), in order to report on proper methods of analysis to the society members.