

# 単一事例実験計画法における評価手続

— A R モデルの臨床への応用 —

芝 野 松 次 郎

## はじめに

近年臨床家の注目を集めている単一事例実験計画法 (single case experimental design) によってもたらされる結果の分析は、グラフ化されたものの視覚的分析 (印象) に頼ってきた<sup>(注)</sup>。一時期統計学的分析の活用について論じ合われたこともあったが、臨床に携わる者は大方統計学的分析の導入には消極的であった。こうした臨床家の消極的態度には種々な理由があると思われるが、その一つとして考えられるのは、単一事例実験計画法から得られるデータが時間的に依存しているために集団実験計画法で従来使われてきた統計学的分析法がストレートに使えないことであろう。本稿では、さまざまな分析法を概観するとともに、視覚的分析 (eyeballing) を越え時間的に依存するデータの分析に適しており、かつ臨床での活用にも適した統計学的分析法である A R モデルを紹介したい。

## 1. 単一事例実験計画法の普及

個人や集団が抱える問題に対する心理社会的な処遇の過程をモニターし、その効果を評価する方法として単一事例実験計画法は早くから注目されてきた。その理論および手続の発展は、Sidman (1960) や Chassan (1967) に負うところが大きい。単一事例実験計画法は、その名が示すとうり個人あるいは一事例を調査の対象として実施できる実験的調査方法であるために、臨床に携わる

者が日常の実践場面で実験的な調査を行うことを可能にした。単一事例実験計画法は、調査と臨床をより密接につなぐうえで大きな役割を果すものであり、近年米国のソーシャルワークで論じられている新たなソーシャルワーカーの役割、すなわち「調査者であり同時に実践家でもある (researcher-practitioner) ソーシャルワーカー」の概念を発展させるのに寄与していると言える。

集団比較実験計画法とは異なり単一事例実験計画法では、独立変数をいくつかの比較集団 (実験群と統制群) を用いて操作するのではなく、一個人あるいは一つの事例を時の流れの中である特定の実験条件 (一般には処遇) 下に置く時期とそれとは別の実験条件 (一般には処遇しない状態でベースラインと呼ばれる) 下に置く時期を連続して作り出すという仕方です独立変数を操作する。

もっとも基本的な単一事例実験計画法は、処遇を施さないベースライン期とそれに続いて処遇が施されるインターベンション期からなるもので、A B デザインと呼ばれている。この両期間に観察記録された従属変数 (一般に人の問題行動) を両期間について比較することによって、独立変数との間の因果的関係を明らかにしようとするのである。

A B デザインに見られるように、ベースライン期とインターベンション期を通して繰り返し継続して従属変数が観察記録されることによって従属変数の時系列 (time-series) データが生み出されるのが、単一事例実験計画法の特徴である。もっとも基本的な A B デザインでは、ベースライン期からインターベンション期に移る時点で中断され

(注) この計画法は、異なる実験条件への無作為割当がないために実験計画法というよりは、擬似実験計画法 (quasi-experimental design) と呼ぶべきであろう (Cook and Campbell, 1979)。

た時系列データを生み出すので、これを中断時系列 (interrupted time-series) デザインと呼ぶこともある (Cook and Campbell, 1979)。

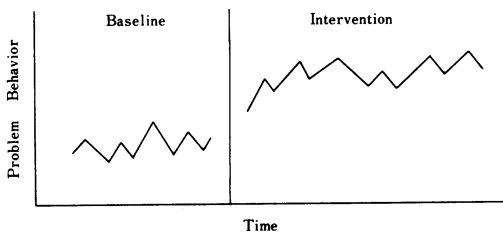
こうした特徴を備えた単一事例実験計画法は、心理社会的な問題を持つ個人 (クライアント) の問題行動が処遇とともに変化するプロセスを時系列データとして提供してくれる。クライアントがソーシャルワーカーの処遇に対してどのように反応し、その問題がどう変化して行くのかは、臨床に携わるものの最大の関心事である。したがって、クライアントの抱える問題行動が変化して行くプロセスを時系列データとして提供してくれる単一事例実験計画法は、まさに臨床家の興味と一致するものであると言える。

一事例のみで実施でき、臨床家の最大の関心事であるプロセスデータを生み出すという二つの利点を備えているために、単一事例実験計画法の活用が臨床の場に普及しつつある。しかし、この実験的調査法から生み出されるデータの分析は、もっとも単純な視覚的印象による評価 (interocular traumatic test) に頼ることが多い。

## 2. 単一事例実験計画における評価手続

単一事例実験計画における評価は、実験的条件の変化にともなって被験者の行動に変化が見られるかどうかを評価することである。図2・1に示

図2・1



すような臨床の場でもっとも使いやすいABデザインの場合には、処遇を施さないベースライン期から何らかの処遇を施すインターベンション期に移るときにクライアントの問題行動に変化が見られるか否かを評価する。

変化が生じたかどうかの判定で問題となるのは、変化の度合が臨床的に見て有意であるのか、

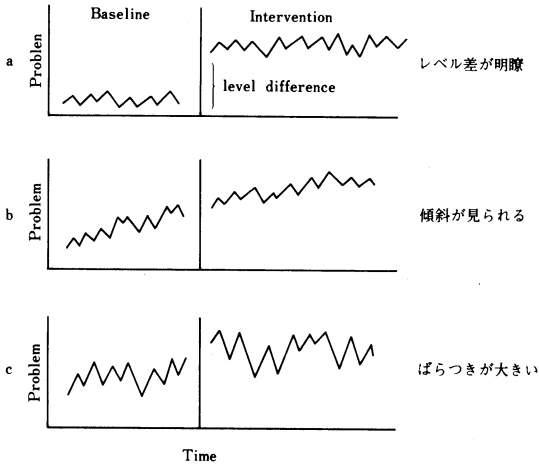
それとも実験的に見て有意であるのかという点である。臨床的有意性と実験的有意性が同じである場合にはまったく問題はないのだが、しばしばこの両者の間に違いが生じる。例えば、肥満に悩む人に対してウェイトコントロールのプログラムを始めたとして、2週間のベースライン期の後にプログラムが実施され、4週間のインターベンション期が持たれた。インターベンション期後半での体重は、ベースライン期よりも5kg少なかった。この変化は実験的には有意な変化であり、ウェイトコントロールのプログラムが効果をあげたと判断できるかも知れない。しかし、臨床的に見ると5kgの体重の減少は、クライアントが目指す健康的な生活の獲得ということを考えれば、微々たる変化であり意味がないと判断されるかも知れない。

このような実験的有意性と臨床的有意性の差は、処遇の目標が明瞭に定められている場合には生じない。体重が60kgまで減少することが目標であるならば、実験的にも臨床的にも目標に到達したかどうかを判断の基準となるのである。

ところが、問題の中には、明瞭で絶対的な目標を設定しにくく、ベースラインとの比較で判断する方がよい場合がたくさんある。例えば、不安のレベルや夫婦間の対話量などは、一定のレベルまで減少あるいは増加すべきであると決めることが困難で、ベースラインと比べて少しでも減少あるいは増加することに目標を置く方がよい場合もある。

このような場合、臨床家の中には、グラフ化された時系列データ (図2・1) に視覚的に明瞭な変化をとらえ得るものでなければならぬとする人たちがいる (Baer, 1977; Parsonson and Baer, 1978)。彼らは明瞭な変化を示す条件として次の二つを挙げている。1) ベースラインが安定していること (stable baseline)。2) ベースライン期とインターベンション期に明瞭な水準差が認められること (clear level difference)。2の条件を満たすためには、1の条件が満たされており、かつインターベンション期の時系列データに傾斜 (slope) やばらつき (variability) がないことが必要とされる。つまり、図2・2のaに示されるような中断された時系列データが、変化を視覚的

図2・2



にもっともとらえやすいものと考えられている。図2・2のbやcでは、傾斜やばらつきのために変化の有無が判断しにくいとされる。

Parsonson と Baer (1978) は、図2・2のaに見られるような理想的な変化は処遇の効果が臨床的に見て十分に強い場合にのみ見られるもので、このような視覚的に明瞭な変化を生み出せない処遇は臨床的に価値がないものであるとした。単一事例実験計画法における評価手続として視覚的分析のみを主張する人々は、基本的には Parsonson と Baer のこのような見解を支持する人たちである。

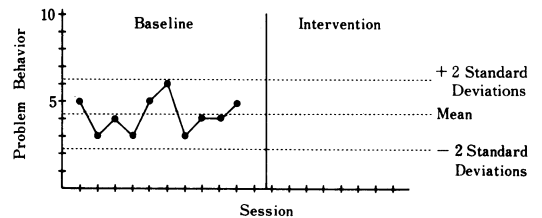
しかし、実践の場では実験的に十分に統制できない要因が多数あり、その結果ベースライン期およびインターベンション期のデータにばらつきが生じやすい (Kazdin, 1976; Gottman and Glass, 1978)。また、同様の理由およびワーカーとの接触や問題行動の観察測定によって一時的にクライエントの行動が不安定になるために (reactivity), 安定したベースラインを得ることが難しい。こうした臨床場面特有の事情によって、効果的な処遇であっても視覚的にとらえ得るほどの変化をもたらし得ないことがしばしばあると考えられる。したがって、「視覚的に明瞭な変化」という基準のみでは、実験的に統制不十分な臨床の場では、一見してわかるような変化を生み出し得ないが、有効な処遇手続としての可能性を秘めた処遇技法を捨てざるようになるかも知れない。ばらつきや不安定さによって視覚的にとらえにくい微妙な変化

をとらえ得る評価方法、言い換えれば、仮説検証のための統計学的方法が必要となる。このような方法によって、有望な処遇手続が無為に捨てられず、検討改善の機会を与えられることになる。これは臨床の向上発展へとつながるものであろう。

統計学的评价手続には、単純な二標準偏差法やそれに類似する簡便な方法から ARIMA (autoregressive integrated moving average) 法と呼ばれる高度な技法まで多数ある。これらの方法については Kazdin (1976) が簡単な紹介をしている。

一般的でもっとも基本的な単一事例実験計画法である AB デザインに用いることのできる統計学的手続は、Gottman と Leiblum (1974), Gingerich と Feyerherm (1978), Jayaratne (1978) などが詳しく紹介している。Gottman と Leiblum の方法は、Shewart (1931) がクオリティコントロールの方法として提唱した手続に基づく。ベースライン期の観察データから平均と標準偏差値を計算し平均の上下に2標準偏差の間隔を取り水平に線を引く (図2・3)。こうしてできたチャー

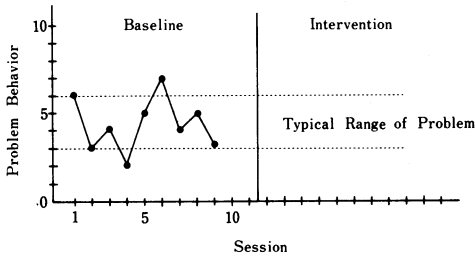
図2・3



トにインターベンション期のデータをプロットしてゆく。プロットingの過程で、観察されたデータ点が2つ連続してこの2本の線よりも上あるいは下にくれば変化が生じたと判断される。この方法では仮説の検証は行われないが、Jayaratne や Gingerich と Feyerherm が紹介した方法では、Bloom と Block の提案に従い二項分布 (binomial distrion) を使って仮説の検証がなされる (Bloom and Block, 1977)。

Jayaratne の相対頻度手続 (relative-frequency procedure) では、ベースライン期の全データのうち中程の2/3が含まれる領域を問題行動の典型域 (typical range) とし、これをグラフ上に2本の水平な線で囲んで示す (図2・4)。次にベースライン期においてこの典型域から望ましい

図2・4

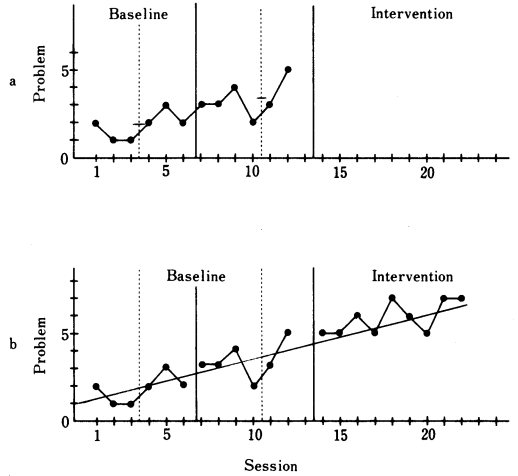


方へ（上あるいは下）はみ出したデータ点の比率を計算する。さらにインターベンション期におけるデータについて典型域からはみ出したデータ点の比率も計算する。もし問題行動に変化が生じたとすれば、ベースライン期のはみ出し率とインターベンション期のはみ出し率は等しくならないはずである。処遇が効果を発揮したとすればインターベンション期のはみ出し率の方が高くなければならない。この仮説をベースライン期とインターベンション期のはみだし率が等しいという帰無仮説について二項分布に基づき検定するのである。

ベースライン期のデータに傾斜が見られる場合には、相対頻度法を使うことは難しい。そこで、GingerichとFeyerhermは、傾斜のあるデータにも使える *celeration line* 手続きを紹介した。*celeration line* は時系列データの傾斜を示す直線であり、中央分割法 (*split middle technique*) と呼ばれる方法で求められる (White, 1974)。中央分割法では、まずベースライン期のデータをその期間の中央で二等分する (図2・5のaのベースライン中央に引かれた縦実線)。さらに二等分された前後の期間をもう一度二等分する (図2・5のaの点線)。そしてベースラインの前半および後半のそれぞれについて平均を算出し、それぞれの期間の縦点線上にする (図2・5のaの点線と交わる水平な短線)。こうして求められた2つの点線上の2つの点を通る直線を引くと *celeration line* が得られる (図2・5のb)。平均のかわりにメジアンを使って *celeration line* を求めることもできる (Kazdin, 1976)。

ベースライン期において *celeration line* よりも上あるいは下にあるデータ点の比率が求められる。*celeration line* をインターベンション期にまで延長し、同様にこの線よりも上あるいは下にあるデータ点の比率が求められる (図2・5のb)。

図2・5



もし処遇の結果変化が生じたとすれば、ベースライン期の比率とインターベンション期の比率は異なるはずである。これを先述の相対頻度法の場合と同じく二項分布を使って検定するのである。傾斜を直線でフィットする方法にはこの他に、回帰直線 (*regression line*) を用いることもできる。

GottmanとLeiblumのShewartチャートは、ベースライン期の観察データの数が10以下でも使えるが、GingerichとFeyerhermが紹介した *celeration line* 法では、ベースライン期に10以上のデータ点がないとフィットは不正確になるとされる。

以上の諸方法は、単なる視覚的評価を越えかつ臨床場面でも手軽に用いることのできる評価方法である。基本的なABデザインに用いることのできるやや複雑な統計学的評価手続きとしては、tテストや分散分析 (ANOVA) を用いる方法が紹介されている (Kazdin, 1976; Jayaratne, 1978)。

これらの方法は、ベースライン期のデータとインターベンション期のデータをそれぞれ一まとまりの集団データとして扱い、それぞれの期間の平均値や分散値を比較検定するものである。tテストをベースライン期とインターベンション期の平均値の比較に用い両期間の水準の差を検定したり、両期における傾斜率 (回帰線の傾き) の比較に用いて両期間の傾斜の差を検定したりする。

Gentile, RodenそしてKlein (1972) は、ANOVA法をABデザインから得られた中断時系列データの分析に用いた。彼らはベースライン

期およびインターベンション期の中でのデータの分散率が withingroup の分散傾向を示し、両期間の間の分散率が between-group の分散傾向を示すと考えた。これらを用いてベースライン期とインターベンション期の違いを F 分布を用いて検定している。

これまで紹介してきた種々の統計学的評価方法には、それらを時系列データの分析に応用することを慎重にせざるを得ない致命的な問題がある。二項分布を使う方法にせよ ANOVA 法にせよ、それらを行う際の基本的な前提条件は、ベースライン期とインターベンション期が統計的に独立しているばかりではなく、個々の観察されたデータ点がやはり統計的に独立していることが必要である。ところが時系列データは、時間の経過とともに繰り返し観察されたデータ点より構成されているのであるから、個々のデータ点が互いに独立しているとは言い難い。事実、臨床場で得られた時系列データに統計的な依存性 (serial dependency) が見られることが報告されている

(Jones, Vaught and Weinrott, 1977; Hartman, et al., 1980; Marsh and Shibano, 1984)。したがって、データの独立性を基礎としている統計学的な評価手続を時系列データの分析に用いることは、統計的に言って不健全 (not robust) と言えよう。

そこで、時間的に依存したデータの評価に用いることのできる統計学的な手続が必要となる。集団を用いる実験計画法で得られた時系列データの場合には、Bock (1963) が提唱した MANOVA (Multivariate Analysis of Variance) という多変量解析の手法がもっとも優れた方法の一つであると思われる。しかし、単一事例実験計画法から得られた時系列データの場合には、ARIMA モデルを基礎とした高度な方法がふさわしいとされる。この方法では、時系列データを構成している deterministic component と stochastic component を区別することから始まる。deterministic component とは、予測できない誤差 (error) に依存していないすべてのパラメーターで、処遇がもたらす効果はこの要素の一つである。一方、stochastic component は観察されない誤差で時系列データの予測を難しくする要素である。この stochastic component は、規則的な部分とそうで

ない全く無作為な誤差の部分 (white noise) から成り立っている。規則的な部分は時間的に依存する要素であり、white noise は統計学的に独立したランダムな要素である。ARIMA モデルのような高度な時系列分析の目的は、stochastic component 中の時間的に依存している規則的な要素を明らかにし、時系列データを white noise にまで還元することである。このようにして初めて、データの deterministic component — 処遇効果 — を正しく評価することができると考えられる (McCain and McCleary, 1979)。

このような高度な時系列分析法は、Box と Jenkins (1970) そして Box と Tiao (1975) の研究に基づく。A B デザインによって臨床場で得られた中断時系列データの分析への応用は、Glass, Willson と Gottman (1975), Gottman と Glass (1978), McCleary と Hay (1980), Gottman (1981) そして Horne, Yang と Ware (1982) などによって詳しく紹介されている。

しかし、こうした高度な時系列分析法には、1) stochastic component を明らかにするのに多数のデータ点が必要である (臨床では、長期に渡って一つの事例から多くのデータを得ることが難しい場合が多い。), 2) 数学的な理論が難解である、そして 3) パラメーターの推定が煩雑であるなどの理由で、臨床家が実践の場で手軽に使えないという問題がある。

### 3. A R モデルによる評価手続

以上概観した種々の評価手続とそれらの持つ問題点を踏まえて、A B デザインから得られた中断時系列データの分析に用いることのできる統計学的評価手続を簡単に紹介する。ここで紹介する方法は、Roberts 等 (1981) や Gottman (1981) により紹介された方法に基づく。基本的には多元回帰法 (multiple regression) で、パラメーターの推定には一般最小自乗法 (ordinary least squares = OLS) を用いるものである。回帰は一般によく知られた統計の方法であり臨床家にも理解しやすい方法である。

前にも述べたように、単一事例実験計画から得られる時系列データは個々のデータ点とその前後

のデータ点と統計学的に独立していないと考えられる。このように時間的あるいは系列的に依存しているデータに OLS 回帰を行うと、回帰からの残差にも系列的依存 (serial dependency) という規則性が見られる。これはパラメーターの推定自体には何らバイアスをもたらさないが、パラメーターの推定値についてその有意性を検定する際にバイアスをもたらす。

ここで紹介する方法は、時系列データのこうした系列依存の規則性を分析し取り除くことによって、回帰を健全に行う方法である。Ostrom (1978) や Roberts 等 (1981) によると残差に系列依存が見られない場合には、バイアスは非常に少ないとされている。

この方法では、次の2つのステップによって中断時系列データに潜んでいる処遇によってもたらされた効果を評価する。1. 時系列データの規則的な stochastic component を明らかにする。2. deterministic component としての処遇効果を検定する。

### Stochastic Component の把握

まずベースライン期の時系列データのみを用いて stochastic component 中の規則性を明らかにする<sup>(注)</sup>。この規則性には2つの過程がある。moving average 過程 (MA) と呼ばれるものがその1つである。これは、ある時点で観察されたデータ点はそれより以前に観察されたいくつかのデータ点の重みづけされた和によって影響されているという規則性で、時系列データのある種の変換によって生じる思いがけない影響を明らかにした Slutsky の考えに基づいている。もう一つの過程は autoregressive 過程 (AR) と呼ばれるもので、moving average 過程よりは理解しやすい。ある時点で観察されたデータ点は、それより以前のデータ点に依存しており、したがってそれより以前のデータを知ることによってある時点のデータが予測可能となるという規則性である。

Box と Jenkins (1970) などの高度な時系列分

析法では、このような2つの過程を同時にとらえ、時系列の stochastic component を明らかにしようとする。しかし、ここで紹介する方法では、次のような理由から AR 過程だけで stochastic component をとらえる。1) MA 過程は、AR 過程によって置き換えることができる。MA と AR の両過程を使用する方がより簡潔に stochastic component をとらえ得るが、臨床から得られたデータは、簡単な AR 過程で説明できる場合が多い (Anderson, 1971 ; Marsh and Shibano, 1984)。2) AR 過程におけるパラメーターの推定に単純な最小自乗法による直線回帰法 (linear regression) が使える。MA と AR の両過程を用いるものは、モデル作りやパラメーターの推定に複雑なコンピュータープログラムと熟練した技術が必要であることを考慮すると、AR 過程のみの方が臨床での使用に適していると考えられる (Gottman, 1981)。

時系列に潜む AR 過程は、自己相関 (autocorrelation) と偏自己相関 (partial autocorrelation) の構成を探ることによって明らかにされる。自己相関は、時系列の中である時点の観察値とそれと一定の時間隔たった時点の観察値との間の関係である。例えば毎日観察されたデータであれば、1日の隔たりがあるデータ点の間の関係、すなわち昨日の観察値と今日の観察値、今日の観察値と明日の観察値との間に見られる関係である。この時間的隔たりはタイムラグと呼ばれる。

時系列を  $x_t$  とするタイムラグが1つの場合の相関係数は

$$r_1 = \frac{\sum_{t=1}^{N-1} (x_t - \bar{x})(x_{t+1} - \bar{x})}{\sum_{t=1}^N (x_t - \bar{x})^2}$$

で求めることができる。N は時系列のデータ点の数、 $\bar{x}$  は時系列の平均を示す。このタイムラグが1つの自己相関係数は第1自己相関係数と呼ばれ、タイムラグが2つの場合は第2自己相関係数

(注) ここでは時系列データの安定性 (stationarity) が前提となっている。安定した時系列はその平均と分散が時間とともに大きく変わることがなく、さらに自己共分散 (autocovariance) も時系列のどの部分を取ってもほぼ同じである。したがって、第1ステップとしては、時系列データの安定性がまず調べられる。non-stationary 時系列でも、何らかの変換によって stationarity が得られるものもあり、そうした時系列の性質は homogeneous non-stationarity と呼ばれる。

と呼ばれる。これを一般化して  $k$  ユニットのタイムラグがある第  $k$  自己相関係数は下の等式で求められる。

$$r_k = \frac{\sum_{i=1}^{N-k} (x_i - \bar{x})(x_{i+k} - \bar{x})}{\sum_{i=1}^N (x_i - \bar{x})^2}$$

自己相関係数をその順位に従って第 1 自己相関係数から第  $k$  自己相関係数まで並べると、所与の時系列における自己相関係数の特徴 (autocorrelation function = ACF) が明らかとなる。

自己相関係数は、それよりも順位の低い自己相関の影響を受けている。例えば、第 2 自己相関係数であれば、第 1 自己相関の影響がその中に含まれている。そこで、より順位の低い自己相関の影響を取り除いて自己相関係数を算出することができる。これが偏自己相関係数である。これを自己相関係数と同様に順位に従って並べると、偏自己相関係数の特徴 (partial autocorrelation function = PACF) が明らかとなる。

このようにして得られた ACF および PACF は、AR 過程を適切にとらえるための重要な手掛りを提供してくれる。

所与の時系列が 1 ユニットのタイムラグから  $p$  ユニットのタイムラグを持つ変数を独立変数として含む AR 過程で十分に説明できる場合には、この時系列は AR( $p$ ) と表示され、下のような等式で表わすことができる。

$$x_t = \sum_{j=1}^p a_j x_{t-j} + e_t$$

$a_j$  は autoregression 係数で、 $e_t$  は残差である。このような時系列の ACF および PACF はきわ立った特徴を示す。ACF は順位の低い自己相関係数が大きく、順位が高くなるにつれて速やかに小さくなり 0 に接近する。一方、PACF は、第  $p$  偏自己相関係数までが比較的大きく、それより順位の高い偏自己相関係数は 0 に近い値となる。

つまり、ACF を見ることによって、所与の時系列が AR 過程で適切にとらえ得るかどうか判断され、さらに PACF を見ることによって、AR モデルの中にタイムラグを持つ独立変数がいくつ必要かが判断される。

こうして時系列の AR 過程をとらえるモデル

が決定されると、自己回帰を実行することによってこのモデルが的確であるかどうか検討される。

OLS によって autoregressive パラメーター ( $a_i$ ) が推定され、その有意性が検定される。例えば、もし AR 過程が 1 ユニットのタイムラグを持つ変数を独立変数としたモデルでとらえ得る場合、つまり AR (1) である場合には、その過程は、

$$x_t = ax_{t-1} + e_t$$

となる。このモデルの autoregressive パラメーター  $a$  の推定値を  $\hat{a}$  とすると、 $\hat{a}$  の有意性は  $e_t$  が正規分布している場合には、 $t = \hat{a}/\sqrt{\text{Var}[\hat{a}]}$  によって  $t$  検定される。 $\sqrt{\text{Var}[\hat{a}]}$  は標準誤差である。 $\hat{a}$  の有意性検定は、AR モデルのフィットの良さを判断する重要な手掛りである。

次に、自己回帰の際得られる多元相関係数  $R$  およびその自乗値  $R^2$  も AR モデルのフィットの良さを示す指標となる。殊に  $R^2$  は、自己回帰による時系列の説明力を示している。

こうして AR モデルのフィットの良さを調べるだけでは AR モデルの的確性を十分に検討したことにはならない。直線回帰の前提となる基本的条件が満たされているかどうかを知る必要がある。基本的条件は、

- 1) 従属変数と独立変数の関係が直線であること。
- 2) 独立変数と誤差の項 (残差の時系列  $e_t$ ) が統計学的に独立していること。
- 3) 誤差の項の平均値が 0 であること。すなわち  $E(e_t) = 0$ 。
- 4) 誤差の項の分散が一定であること。すなわち  $E(e_t^2)$  が一定であること。
- 5) 誤差の項が自己相関していないこと。すなわち  $E(e_t e_{t-m}) = 0$ 。ただし  $m \neq 0$ 。
- 6) 誤差の項が正規分布していること。

これらの条件が満たされているかどうかは回帰による残差の系列  $e_t$  を吟味することによって判断される。2~6 の条件は残差系列の平均や分散を調べると同時に、残差系列の ACF を調べることによってその適合性が検討される。つまり自己相関係数が順位の低いものから高いものまで標準

誤差の範囲にあることが確かめられる。このためには Box と Jenkins (1970) が提唱した Q-statistic が便利である。Q-statistic は下のように定義される。

$$Q = N \sum_{j=1}^k r_j^2$$

$N$  は時系列を構成するデータ点の数、 $k$

は自己相関の最高順位である。この Q-statistic は自由度  $k - p$  でカイ自乗分布する。

1 の条件である独立変数と従属変数の直線的関係は、残差と回帰から得られた従属変数の推測値を散布図に表わし視覚的に検査することによって確かめられる。

こうした諸検査によってモデル的確性が確かめられ、時系列の stochastic component が明らかにされると第 2 のステップへと移ることになる。

### 処遇効果の検定

時系列の規則的な stochastic component が明らかになると、deterministic component としての処遇モデルが組み込まれ、その効果が検定される。処遇をモデルするために、ベースライン期には処遇が施されずインターベンション期に処遇が施されるという AB デザインの実験的操作に基づき、処遇以前を 0 そして処遇後を 1 とするダミー変数が用いられる。

もっとも簡単な AR (1) 過程で stochastic component を表わし得る場合を例にとると、処遇モデルを組み込んだ時系列  $x_t$  は次のようになる。

$$x_t = ax_{t-1} + bI_t + e_t$$

$I_t$  は上述の処遇を示すダミー変数で、 $b$  が処遇効果 (回帰係数) を示す。これは、パラメーターの推定に OLS を用いる一般的な多元回帰である。

この時系列がベースライン期に  $n_1$  のデータ点を含みインターベンション期に  $n_2$  のデータ点を含むとし、データ点は全部で  $N (= n_1 + n_2)$  あるとする。これらのデータは次のように表わすこと

ができる (Gottman, 1981)。

$$\begin{aligned} x_2 &= ax_1 + e_1 \\ x_3 &= ax_2 + e_2 \\ &\vdots \\ x_{n_1} &= ax_{n_1-1} + e_{n_1} \end{aligned}$$

$$\begin{aligned} x_{n_1+1} &= ax_{n_1} + e_{n_1+1} + b \\ x_{n_1+2} &= ax_{n_1+1} + e_{n_1+2} + b \\ &\vdots \\ x_N &= ax_{N-1} + e_N + b \end{aligned}$$

このように処遇の効果  $b$  は徐々に蓄積され  $b/1 - a$  の水準へ到達することがわかる。つまり、処遇の効果は処遇が導入された時点で突然水準差を引き起こすものではなく、徐々に積み重ねられ一定の水準に達することがわかる。臨床的には、処遇効果が唐突な水準差を生み出すとするよりも、ここに示すモデルの方が適していると考えられよう。

処遇効果のパラメーター  $b$  は  $N$  が十分大きければ  $t$  分布すると考えられ、これに基づき多元回帰から得られた  $b$  の推定値  $\hat{b}$  の有意性が検定される<sup>(注)</sup>。

ここでは AR (1) 過程について説明したが、これを AR (p) 過程へと一般化することができる。これについては Gottman (1981) を参照されたい。

第 2 ステップは、deterministic component を組み込んだモデルが適切なものであるかどうかを、第 1 ステップの場合と同様に、モデルのフィットの良さと回帰の基本的条件が満たされているかどうかの検定により確かめて締めくくられる。

### おわりに

単一事例実験計画法の基本型である AB デザインから得られる中断時系列データの分析方法を概観し、臨床の場での活用にふさわしいと思われる統計学的評価手続を紹介した。AR 過程によって

(注) 有意性の検定は  $b = 0$  を帰無仮説とし  $t = \hat{b} / \sqrt{\text{Var}[\hat{b}]}$  で行われる。 $\sqrt{\text{Var}[\hat{b}]} = Se \cdot \sqrt{\sum_{i=1}^{N-1} x_i^2 / [n_2 \sum_{i=1}^{N-1} Y_i^2 - (\sum_{i=1}^{N-1} Y_i)^2]}$  で  $Se$  は残差から得られる標準誤差。自由度は  $N - 3$  である。



時系列の stochastic component をとらえることを基礎とし処遇効果を健全に評価するこの方法は、パラメーターの推定が容易であり臨床に適していると思われる。

しかし、この方法にはいくつかの限界と問題点がある。まず、処遇効果を明らかにするためのモデルが、効果が徐々に蓄積され一定の恒久水準に到達するという仮定に基づいている点である。処遇効果にはこれ以外に種々の形がある。そうした処遇効果はこの方法では評価し難く、さらに複雑な評価方法が必要となる。

前述したように、処遇効果のパラメーターはデータ点の数が多ければ  $t$  分布すると考えられるので、推定値の有意性は  $t$  検定によって確かめることができる。しかし、臨床データの場合には、多くのデータ点を得ることが難しい。データ点がどの程度あればよいのかは今後モンテカルロ法を使用して調べられる必要があろう (Gottman, 1981)。

また、臨床の場面では種々の理由から観察が定期的に継続して行えないために、臨床データにはデータの欠損がしばしば見られる。このような欠損データの分析に及ぼす影響は、まだ十分に知られていない。Rankin と Marsh (1985) はこの点について興味深い調査を行い、欠損データが全体の 20% 以下であれば影響は少ないと結論している。今後さらに詳細な調査が待たれる。

### 参 考 文 献

- Anderson, T. W. *The Statistical Analysis of Time-series*. New York: Wiley, 1971
- Baer, D. M. "Perhaps It Would Be Better Not to Know Everything." *Journal of Applied Behavior Analysis*, 1977, 10, 167-172.
- Bloom, M. and Block, S. R. "Evaluating One's Own Effectiveness." *Social Work*, 1977, 22, 130-136.
- Bock, R. D. "Multivariate Analysis of Variance of Repeated Measures." In C. W. Harris (Ed.), *Problems in Measuring Change*. Madison, Wis.: Univ. of Wisconsin Press, 1963.
- Box, G. E. P., and Jenkins, G. M. *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day, 1970.
- Box, G. E. P., and Tiao, G. C. "Intervention Analysis with Applications to Economic and Environmental Problems." *Journal of the American Statistical Association*, 1975, 70, 70-79.
- Chassan, J. B. *Research Design in Clinical Psychology and Psychiatry*. New York: Appleton-Century-Crofts, 1967.
- Cook, T. D. and Campbell, D. T. *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Chicago: Rand McNally, 1979.
- Gentile, J. R., Roden, A. H., and Klein, R. D. "An Analysis-of-Variance Model for the Intrasubject Replication Design." *Journal of Applied Behavior Analysis*, 1972, 5, 193-198.
- Gingerich, W. J. and Feyerherm, W. H. *Assessing The Significance of Client Change*. Unpublished paper, School of Social Welfare, Univ. of Wisconsin-Milwaukee, 1978.
- Glass, G. V., Willson, V. L., and Gottman, J. M. *Design and Analysis of Time Series Experiments*. Boulder: Colorado Associated University Press, 1975.
- Cottman, J. M. *Time-Series Analysis: A Comprehensive Introduction for Social Scientists*. Cambridge: Harvard University Press, 1981.
- Gottman, J. M., and Glass, G. V. "Analysis of Interrupted Time Series Experiments." In T. R. Kratochwill (Ed.), *Single Subject Research: Strategies for Evaluating Change*. New York: Academic Press, 1978.
- Gottman, J. M. and Leiblum, S. R. *How to Do Psychotherapy and How to Evaluate it*. New York: Holt, Rinehart and Winston, 1974.
- Hartmann, D. P., Gottman, J. M., Jones, R. R., Gardner, W., Kazdin, A. E., and Vaught, R. S. "Interrupted Time-Series Analysis and Its Application to Behavioral Data." *Journal of Applied Behavior Analysis*, 1980, 13, 543-559.
- Horne, G. P., Yang, M. C. K., and Ware, W. B. "Time Series Analysis for Single-Subject Designs." *Psychological Bulletin*, 1982, 91, 178-189.
- Jayarathne, S. "Analytic Procedures for Single-Subject Designs." *Journal of Social Work Research and Abstracts*, 1978, 14, 30-40.
- Jones, R. R., Vaught, R. S., and Weinrott, M. "Time-Series Analysis in Operant Research." *Journal of Applied Behavior Analysis*, 1977, 10, 151-166.
- Kazdin, A. E. "Statistical Analysis for Single-Case Ex-

- perimental Designs." In M. Hersen and D. H. Barlow (Eds.), *Single Case Experimental Designs : Strategies for Studying Behavior Change*. Oxford : Pergamon Press, 1976.
- Marsh, J. C. and Shibano, M. "Issues in the Statistical Analysis of Clinical Time-Series Data." *Journal of Social Work Research and Abstracts*, 1984, 20, 7-12.
- McCain, L. J., and McCleary, R. "The Statistical Analysis of the Simple Interrupted Time-Series Quasi-Experiment." In T. D. Cook and D. T. Campbell (Eds.), *Quasi-Experimentation : Design and Analysis Issues for Field Settings*. Chicago : Rand McNally, 1979.
- McCleary, R., and Hay, R. A., Jr. *Applied Time Series Analysis for Social Sciences*. Beverly Hills : Sge, 1980.
- Ostrom, C. W. *Time Series Analysis : Regression Techniques*. Beverly Hills : Sage, 1978.
- Parsonson, B. S. and Baer, D. M. "The analysis and presentation of graphic data". In T. R. Kratochwill (Ed.), *Single subject research : Strategies for evaluating change*. New York : Academic Press, 1978, 101-166.
- Rankin, E. D. and Marsh, J. C. "Effects of Missing Data on the Statistical Analysis of Clinical Time Series." *Journal of Social Work Research and Abstracts*, 1985, 21, 13-16.
- Roberts, H. Ling, R., and Bateman, G. "Conversational statistics." Unpublished version. Graduate School of Business, The University of Chicago, 1981.
- Shewart, W. A. *The Economic Control of the Quality of Manufactured Product*. New York : Macmillan, 1931.
- Sidman, M. *Tactics of Scientific Research : Evaluating Experimental Data in Psychology*. New York : Basic Books, 1960.
- White, O. R. "The 'Split middle'-a 'Quickie' Method of Trend Estimation." Experimental Education Unit, Child Development and Mental Retardation Center, Univ. of Washington, 1974.