

心理学方法論におけるフィッシャー統計学の受容

R.A.フィッシャーにより開発された実験計画法は、それ以前の大量にサンプルを必要とする方法と比べて、より少数のサンプルでなおかつ複数の独立変数の相互効果までを調べる事を可能にした。実験計画法や分散分析などを含む、彼が発展させた統計的方法論は、1925年に公にした「研究者のための統計的方法」(Statistical Methods For working researchers) 1935年の「実験計画法」(The Design of Experiments)でまとめられ、その後の自然・社会科学研究者に大きな影響を与えた。(Salsburg2002)しかし心理学の分野では、彼の方法論を積極的に支持する研究者も居たものの(Thouless Dawson など)、予想されるような広範な適用はやや遅れた。その経緯の事情を L.Cronbach のテキストを中心に見ていきたい。

実験計画法とは R.フィッシャーがロザムステッド農事試験場において観測値の収集解析のために開発した研究方法論である。農事試験場では日当たりや水はけの違い、土地の肥沃度によって試験結果に系統的な差異が現れる。これらの差異と実験的処置(品種や肥料の種類を変更するなど)による差異を区別するため、現代的統計学の出現以前から、処置のブロックを桂馬飛び状に系統的に配置するなどの方法が知られてきた。

20世紀初頭に全盛を迎えた K・ピアソンを中心とする現代的統計学の生物測定学派の立場では、このような農事試験場の条件に由来する系統的な誤差は不可避なものであり、サンプルの測定値を増やす事で除去するべきものであった。しかし単純に測定値を増やす方法では標本平均に有意な差を見出すために大量のサンプルが必要となった。これに代わる方法としてフィッシャーが提案したのは、各処置をブロック化してその内部では他の条件が変わらないようにする(局所管理ブロック化)、何度も繰り返す(反復)、ブロックの配置順をランダム化する(ランダム化)という手法だった。こうすれば変異の要因ごとに変動を分離できるため、比較的少ないサンプル数の標本から誤差の分散を推定できる。この手法は実験計画法と呼ばれ、1935年に出版されたフィッシャーの教科書「実験計画法」によって、様々な分野に広がった。

しかし心理学においてはメンタルテスト以外の研究分野で、実験計画法はおろか現代統計学的手法が有力な研究手段になる事は無かった。学習や知覚を研究する実験心理学で

は、環境を極度に一貧弱なものにしたり（ラットの）脳の一部を破壊したり、あるいはただ素朴に測定個体数を増やしたり、あるいはその逆にスキナーのように個体数を1にする事で個体間の誤差を無視できるものにしようとしていた。（Cronbach1957）教育心理学者L・クロンバックは1957年の心理学会会長就任講演で、心理学の研究手法について「一つは実験心理学であり、もう一つは相関心理学である。ダシールはこの二つが合流すると楽観的に予測したが、それはまだ完了していない。」と述べている。

心理学で現代的統計学的研究手法の特徴であるコントロール群を用いた実験が行われるようになったのは20世紀に入ってからである。ソーンダイクやウッドワースによって学習効果の転移、すなわちラテン語を勉強する事が一般的な知能の鍛錬となるかどうかを調査され、否定的な結論が出た。これは当時の教育カリキュラムの改訂に大きく影響し、その後スピアマンやサーストンはメンタルテストに因子分析の手法を導入し、教育心理学は現代的統計学の要素を持つようになった。当時フィッシャーの統計学的研究手法を受容した心理学者は主にこれらの教育心理学者たちである。1940年代にフィッシャーの統計学について心理学的解説書を書いたのはみな教育心理学の研究者であった。（Dehue2001）

では何ゆえ他の心理学の研究分野ではフィッシャーの統計学的手法の受容が遅れたのだろうか？まず1930年代40年代は理論発展の時代であり、比較的素朴な実験でも事が足りた事が挙げられる。例えばスキナーのオペラント条件付けの実験で使用するラットの籠は通常1つであり、これはフィッシャーの実験計画法で典型的な 2^n 種類の精密標本を必要とする実験よりもずっと手軽である。ラットのレバーを押す行動と報酬や罰のスケジュールの関係はただ単純にレバーを押す頻度（回数/時間）を用いて測定された。実験個体数を1にして、レバー押しなどのように反復が容易な測定値を用い、処置前と処置後の期間の比較に実験目標を設定すれば個体間の誤差を気にせずに心理学の実験が行えたのである。（Skinner1979）もう一つの理由は実験計画法に不可欠なランダム化の困難さである。例えば第二次世界大戦中の米軍で、戦意高揚映画の効果を調べるために実験計画法を用いた研究がデザインされたがうまく機能しなかった。それは映画を見るために特別に呼ばれたというだけで兵士は当局の意図になんらかの疑念を抱いてしまう事が原因だった。（Hovland1949 Dehue2001）これは農事試験場には存在しない種類の困難である。

クロンバック（1957）によると、メンタルテスト理論家を中心とする相関心理学と、それ以外の実験心理学の研究は、お互いを無視したまま1世代以上が過ぎてしまったが、1950年代後半までに融合の潮流が見られるようになってきた。相関心理学から実験心理学への歩みよりは、テスト基準の構成概念妥当性の提唱（Meehl1955）となってあらわれた。これはテスト結果がただ単に誤差が少なく偏りが無いというだけでなく、試験者の意図した値を本当に測定しているということを保証するための基準である。テストの得点を実験

心理学の理論に埋め込んで予測を行い、理論から当然テスト得点が高くなると予測される場合にテスト得点が高くなる場合、そのテストの解釈の妥当性は保証されるとする提唱である。この妥当性を満たす事でメンタルテストと実験心理学はより哲学的に洗練された協調関係を持つことができ、得体の知れない相関関係を持ち出す研究が減るのではないかとクロンバックは期待している。反対に実験心理学が相関心理学から得ることができる一番大きなものは、クロンバックによると、世界を多変量的に把握するやり方である。これまで実験家の直観と経験に頼っていた最小の冗長性で変数を選び最小の投資で重要な結果を出すというプロセスは現代的な統計学を用いると組織的に行えるようになる。多種多様な環境因子はストレスという生理学的因子と結びつける事ができると、ストレス学説の創始者セリエの研究と自らの研究を引用して、クロンバックは論じている。

またランダム化に伴う心理学研究の方法的困難についても、1960年代以降研究が進んだ。これも軍隊における研究に端を発しており、ドナルド・キャンベルの”**Experimental and Quasi-Experimental Designs for Research**” (1966) ではランダム化が限定された状況での実験計画法の使用法について、特に社会科学における様々なケースについて論じている。(Dehue2001)

参考文献

- “History of Experimentation in Psychology” T.Dehue (2001)
- “The two disciplines of scientific psychology.” Cronbach, Lee J. (1957)
- .“統計学を拓いた異才たち”デイヴィット・サルツブルグ(2002)
- “R.A.フィッシャーの統計理論” 芝村良 (2004)