

看護学研究の方法と分析
—統計的手法を使う価値と意味—

柳 修平

Implementation and Significance of Statistical Methods in Nursing Research

Shuhei Ryu

姫路大学大学院看護学研究科論究

第2号

平成31年3月1日発行

看護学研究の方法と分析

—統計的手法を使う価値と意味—

柳 修平

Implementation and Significance of Statistical Methods in Nursing Research

Shuhei Ryu

1. はじめに

かつて医療の実践は身体的な機能や症状の改善の結果が判断基準とされ、その成否に社会的な評価が下される時代があり、医療の主体性は確かな直感と強い信念に従い懸命に治療する医師の側にあった。しかし健康を左右する外的条件として最も比重の大きい要素が、人々自身の行動様式であることに認識が高まり、人々の健康を生活概念で捉えることが社会的に重要視され、医療、看護、福祉等の実践はこれを包括的に支援するものとして理解されるようになる。広義の医療はQOL (Quality of Life) を共通概念に据え、社会的機能の喪失の改善に取り組むことが専門職者の責務となり、多職種による協働作業の態勢づくりの整備が図られている。それまでの専門職者の直感はただ暗示するだけであり、信念は実行力を支えるだけである。言い換えれば情熱と使命感を表明することではなく、実践し展開する一連のケアの根拠を客観的・科学的に説明し、チームで共有するのみならず、そのケアを実施するかどうかを患者・クライアントの自由な意思決定により保障できる十分な情報を提供し同意を得ることが不可欠であ

る。

人を対象とする研究では、倫理的妥当性及び科学的合理性を確保するために、倫理審査委員会に研究実施計画書等の審査を依頼し承認を得る。適切な研究デザインと解析手法がエビデンスの質を高めるといわれ、統計的手法の選択が科学的合理性の判断根拠の審査に大きく絡む。統計はエビデンスを得る研究計画に欠かせない道具であり、日頃の実践活動で収集・蓄積するデータの管理と情報への整理は研究を効果的に進めるものである。

雑誌体であった「医中誌」が、いまはWEBからリアルタイムで検索でき、関連論文や参考資料を寸時に取り出すことが可能である。抜き出したデータを後々役立つ情報とするために、分類と統合が可能な紐付けを加えておくことも直ちにできる道具があり、新たな知識や実践の指針を創造する研究活動の源になっている。

2. 情報の管理

今日のような大量な情報が一方的に発信されている情報過剰化社会では、必要な情報を取捨選択しその真偽や精度を判断することは受信者側の責

任に委ねられる。しかも保健医療の現場では、情報を発信する側にその信頼性を保障し確保する責任と義務がある。例えば臨床現場で看護師が誤った情報を発信することは論外であり、情報を受け取る患者側が容易に理解できないことや受けとめに誤解を生じる恐れのあるものは、正しく伝えることができている看護師の責任であることは容易に理解できよう。すなわち受け手が分かってこないのは、相手が理解できなかったことではなく、分かるように伝えられなかったことを研究者の問題意識にすることである。専門職者の一員である研究者は、情報の送受信の両方で、その質に責任をもち確実に伝わるように実行することが要求されている。

研究者が研究成果を世に問うためには、意図を明確にした情報を伝え、受け手に最高の価値を与えることとなる。少なくとも情報の内容、形式を明瞭に整理し発信できていれば大きな誤解は生じない。システム論的にいうと、研究活動はその性格上閉鎖系システムの中で行われるので、情報の収集、生産、伝達というコミュニケーションを円滑にでき、情報処理プロセス全体で関係者に共通の理解を得ることの保障を確保することであり、他の人から共通の理解を得られる道具をもつことは最善の手段である。

3. 統計の意義

統計は実践活動や研究活動の評価を論じる際に共通言語として欠かすことのできない道具である。コンパクトに情報をまとめ、その確かさを客観的に説明し、有意義な情報を提供し、効果的に活用する意味を伝える。ところが実践的意義は専門分野の独自の知識や知見から判断するもので、統計学的に有意なことがらが実践的な意義をもつ

とは限らない。統計解析の結果が有意な情報を提案したとしても、どの程度の危険率で活用するか判断や実践に適用する意思決定は統計で定められるものではない。極めて個別性の高い専門領域の知識に基づき問題解決の手がかりをつかみ繰り返し検証し実証される。専門色が濃い学会の口演や示説は、情報に多少の欠落があっても、発表の場で送り手や受け手の反応をじかに体感でき、質疑によって適切に埋め直して理解することが可能である。ところが幅広い境界領域を含む学会では情報交換において共通の言葉を探すことに苦労する。統計は分析・解析方法が定義されているので、情報伝達的手段に公用語として使える強みがあり、データの背景にある専門的な現象の意義を説明する手助けをしてくれる。統計パッケージや自動解析測定機器が普及した環境では、数式や計算手順を知らなくても実行したい統計課題を操作するだけで結果が得られる。統計は現象を検証する道具として便利であり、自由に使うことが可能な条件を大いに活用すればよい。

統計を用いる研究は、実態を記述することと観察した現象の法則を追求することの2点を基本に据えるとよい。統計処理は、データが発生している環境の記述、解析の対象となる環境の構造、さらに結果が適用される環境への影響の理解と段階的に進んでいく。日々の実践・研究の中で取り組む課題の判断材料を揃え、データを蓄積する環境を創出することが研究活動の第一歩である。

看護系大学院の論文作成の武器に質的研究の手法がさまざまに開発され普及したお蔭で統計的研究を量的研究と呼ぶ人が多い。しかし推論は論理規則にしたがって既知の事柄から未知の事柄について知ることであり、いわゆる量的研究は推論の及ぶ範囲をあらかじめ定めることがあるとしても、統計の論理は観察し分析した環境の量的特性

から予測や影響の結果評価を質的に捉えているものである。研究の質を保証するエビデンスにレベルがあることは周知のことであろう。基本的に統計は道具として上手に使えばよいので、その条件をまず満たすためにはデータの特性を知ることにある。端的に言えば統計手法は数値の計算なので、統計を使う出発点としては目の前で起こる現象を数値で捉えることである。それができると計量・計測・計数により記録することに移れる。統計処理ができるデータが得られない限り統計的手法は役立たない。アウトカムの質の担保がデータにあるのは明白であろう。

4. 統計データの性格

看護研究は観察した事例への印象や疑問がベースにある。研究課題を定める作業の契機となった現象を記録する。研究者が印象に強くとどまった現象を研究によって検証する。それらのプロセスはいわゆる質的研究であれ、量的研究であれまったく異なるものではない。研究者が観察している対象集団にいま起きている現象をどのように記録に留めるかを工夫し、それからデータベース化を試み実施するのも同じである。そこで多少異なるのは統計的研究では対象集団に含まれる個体の見分けに使う目印をあらかじめ与えて作業するところに特徴がある。この操作は標識化というが、

統計では個体 ω について問題とする特性 X を測り、 $X(\omega)$ の値を定めるので数量化という。その定めた値がデータであり、名義（分類）尺度 *nominal scale*、序数（順序）尺度 *ordinal scale*、間隔尺度 *interval scale*、比例（比）尺度 *ratio scale*の4つの尺度のいずれかになる。

データ尺度の性質から名義尺度と序数尺度を定性的（計数、非計量）データ、間隔尺度と比例尺度を定量的（計量）データともいう。

血圧、体温、ADL評価などのデータは数値自体から特性の大小関係が分かるが、性別が男であろうと女であろうとその特性の大きさは比べられない。集団で比較するとき代表値に平均値を定めることができるものとできないものがある。先行研究を検索したところ分散分析や重回帰分析で属性との関係を検討していたので、それで比較したいと熱望してもデータの尺度が比較研究と同一でなければ無理である。すなわち統計解析手法の応用範囲はデータ尺度に限定される。順序に意味があるデータは順位和で比較可能で、平均値に意味があるデータは平均値の差で比較する方法となる。積率相関か順位相関かの判断はデータの尺度に依存する。このようにデータのもつ性格は分析手法と密接している。データを集めた後に統計手法の相談をするのではなく、データを集める前にどのような統計手法で何を説明するかの検討は必

表1 データの尺度

量的データ（連続量）	
比例尺度：絶対零点あり	：重さ、長さ、血圧など
間隔尺度：絶対零点なし	：数値の差に意味あり　：摂氏の温度、学力など
質的データ（離散量）	
序数尺度：順序の意味あり	：数値の差に意味なし　：5段階評価、ADLの評価など
名義尺度：順序の意味なし	：性別、診療科名、病名など

須事項である。

人を対象とした研究は研究倫理審査を受けるため、統計的研究デザインを用いるときは分析方法とそれにより得られる主要評価項目や副次評価項目を書くことが必要である。統計的研究は標識化と数量化を明確にすることでデータの収集が可能になる。問題解決を目指すのに有益な解析法はデータの性質に応じた手法に従うので、そのことをプロトコールに記載する。問題となるのはアウトカムの記述にあると推測できるが、これは統計学の役割ではなく研究分野の専門的な学問領域の範疇で苦闘する課題である。

文献検索の作業中にデータのもつ特性と分析手法を紐づけにしておけば、蓄積した情報から統計のイメージをまとめることができる。イメージが豊かであればあるほどデータとコミュニケーションする研究時間を増やすことができる。質的に高い成果が得られそうな見通しをつけ背景要因の関連を十分に考慮した精度と確度の高い解析に進むことが気持ちの上でも容易となる。

5. 統計処理の妙

看護研究は直接ケアの対象となる人を相手にするものと、ケアのための間接的な事象や事物を対象とするものがある。平均値の検定や適合度の検定の式を眺めると、標本数を大きく増やすと数学的に有意差が得られることが分かる。有意差があるという考察で終わることはなく、有意差が生じる要因あるいはそれを制御することがどれだけ有意義な実践活動ができるかが考察のコアである。そこで無視できないことは検証に適切な集団を対象としているか、調べている要素は妥当であるか、説明している特徴が矛盾のない情報であるかなどであり、的外れではない提案を行う前提

条件を理解しておくことである。人間は誤るものであるが、バイアスに対処することができることは知られている。

例として、200人を対象に100人ずつの二つのグループに分け、X方式とY方式の2つの改善介入を行い、ある機能の改善を評価したとしよう。結果は、X方式で39人、Y方式で61人の改善がみられ、改善率はY方式が良い結果にあるとした。出現率の検定は有意にY方式の改善率が高いとなる。ところが対象の特性によって改善に違いがあるのではという思いが気になったとしよう。研究者は特性のデータも取っていたので、表2に示すように、各グループを特性別に改善状況を求め比較することになる。

表2. 対象の特性による判断が必要な事例

	X 方式			Y 方式		
	対象数	改善数	改善率	対象数	改善数	改善率
全体	100	39	39%	100	61	61%
特性1	20	20	100%	50	48	96%
特性2	30	15	50%	30	12	40%
特性3	50	4	8%	20	1	5%

表2の結果から特性別に改善率を計算すると、X方式はどの特性においてもY方式の改善率を上回る結果を示している。先行研究の文献検討で見つからなかった特性であれば大発見であるが、もしも見過ごしていたとすると大変なことである。データの背景にある隠れた要因（交絡要因）を見落とすと大事になる可能性が指摘できる事例であるが、先行研究の検索や対象の理解をしっかりと押さえることがいかに重要であるかの指摘でもある。この特性の例に頻繁にあげられるファクターが年齢因子である。疫学・保健統計学の教科書には年齢調整死亡率として解説が掲載されているが、背景疾患、身体レベル、施設特性などさまざ

まなファクターが交絡要因となっていることは日頃の研究・実践活動の印象に少なからずあるはずだが、統計的処理の方法を検討するときにそこに意識がないと、データとして見落とすことも起きる。

比較により浮かびあげる事例をもう一つ考えてみよう。2つの設問をし、設問1では数年前の健康教育の参加経験を、設問2では調査時点での健康活動への参加意識を調べたものとしよう。100名に聞いた結果Yes - Noがそれぞれ50名ずつとなった。参加経験も参加意識もそれぞれ半々という結果からどちらの設問も意味を見いだせず、ここで終わりにすることもあるかもしれない。ところが回答パターンにまったく何の構造がないとは限らない。2つの変数を組み合わせて眺めることによって印象が大きく変わる可能性がある。

表3. 組合せて比較するとみえてくる事例

	例1 Q2			例2 Q2		
Q1	Yes	No	計	Yes	No	計
Yes	45	5	50	12	38	50
No	5	45	50	38	12	50
計	50	50	100	50	50	100

表3に2つのパターンを示した。いずれもそれぞれの設問の回答は半々であるが、例1の結果が得られると、過去の参加経験がいまの参加意識に大きくつながっている可能性も考えられる。反対に例2のようであれば、過去の経験が現在の参加意識を阻害しているのではないかの検証を求められ可能性がある。単純集計では分からないことが、クロス集計で見えることもあるのは、対象を1変数のみで評価するのではなく、2変数から観察することで情報量が倍増した結果である。さらに変数を増やし多変量解析に手を伸ばしたくなる理由

になる。データ解析はデータとコミュニケーションを図るだけのことである。コミュニケーションの取り方にいろいろな工夫があるように、自由に工夫ができる便利な道具として統計手法が揃えられるとよい。

6. データとベクトル

前述の2事例は質的データの基本統計量をクロス集計表で示したもののだが、量的データの場合は散布図を描くと何らかの関連が見つかることがある。

量的データの2変数の関係で線形相関を統計量として表すものを相関係数correlation coefficientといい、n個のデータが $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ で、それぞれの平均を \bar{x} , \bar{y} と表すと、相関係数rは、

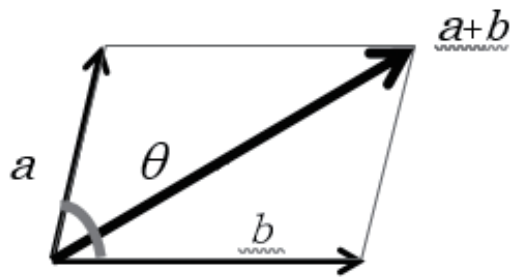
$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

で表せられる。分子、分母ともnで割ることによって、共分散とそれぞれの標準偏差で

$$r = \frac{\frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n}}{\sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}} \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n}}} = \frac{s_{xy}}{s_x s_y}$$

と表せる。S_{xy}が共分散を示す。相関係数はデータの散らばり方を考える基本となる偏差の式で構成され、積率相関係数ともいう。

量的データは個体の大きさと集団の代表値を比較した偏差（個体の値-平均値）が統計的な操作をする際の基本にあり、偏差データを大きさと特性（向き）を持たせたベクトルのイメージで理解するとよい。



$$(a, b) = |a||b| \cos \theta$$

図1. 2変数とベクトルの性質

ここで量的データの偏差をベクトル量で考えると、

$$\vec{X} = \begin{bmatrix} x_1 - \bar{x} \\ x_2 - \bar{x} \\ \vdots \\ x_n - \bar{x} \end{bmatrix} \quad \vec{Y} = \begin{bmatrix} y_1 - \bar{y} \\ y_2 - \bar{y} \\ \vdots \\ y_n - \bar{y} \end{bmatrix}$$

と表記し、ベクトルの内積で示すと、

$$(\vec{X}, \vec{Y}) = \cos \theta [|\vec{X}|][|\vec{Y}|] \quad \text{から} \quad \cos \theta = \frac{(\vec{X}, \vec{Y})}{|\vec{X}||\vec{Y}|}$$

ここで、

$$\begin{aligned} (\vec{X}, \vec{Y}) &= \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) && \text{: ベクトル成分の内積} \\ |\vec{X}| &= \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} && \text{: ベクトルXの大きさ(スカラー量)} \\ |\vec{Y}| &= \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2} && \text{: ベクトルYの大きさ(スカラー量)} \end{aligned}$$

の関係式から2つのベクトル間のなす角の余弦 ($\cos \theta$) は相関係数 r と一致する。2つのベクトルはなす角が0のときは、同じ方向で共通のものを捉えているベクトルとみなせ、相関係数は完全相関を示す1 ($= \cos 0$) となる。また直交しているときは無相関の0 ($= \cos 90$) を示し、図1の内角 (θ) が小さくなるほど相関係数が大きくなることを確認してみよう。

データは数値化することで統計的処理ができることと、データがベクトルであるイメージできると統計を理解するのに役立つ。非計量型データである順序尺度はリッカートスケールとして定量

的に扱うことは知られている。名義尺度は単に区別し分類するためだけに用いられる数値を与えるものであるが、好みXの分類を {1. 甘いもの 2. 辛いもの 3. しょっぱいもの 4. その他} で回答を得た場合、1から4をあてはめることができるが、甘いものを {1, 0, 0}, 辛いものを {0, 1, 0}, しょっぱいものを {0, 0, 1}, その他を {0, 0, 0} と与えることができる。重回帰分析等でダミー変数として扱うが、名義尺度をベクトルに変換していると理解するとよい。複数の変数のベクトルを合成し別のある変数を説明するイメージがわくと重回帰分析や多重ロジスティック分析も理解しやすい。既知の変数から未知の潜在変数や予測変数を知ることにつながるイメージを掴むことができれば、あとは道具を使って検証するだけとなる。

7. おわりに

統計は集団での個々の要素がもつ数値の分布や、その分布の特徴を示す数値の総体を示すもので、データの総体になんらかの構造や傾向があれば、適切な分析方法でみつかるものである。そのためには種々の数量をまとめて計算するという数学的処理が必要となるが、その道具も揃っている。最も重要で不可欠なことは専門領域の視点から統計処理の前後でデータとしっかりコミュニケーションをとることである。データはそれが発生した現場で取得した素材であり、まとめた結果が研究に意味がある事実なのかそこに帰ることで考察する。解析手法で加工できない素材では処理に困る。常に出発点で研究目的に応じたデータ収集ができるかを点検することを位置づけておくことが統計的研究を開始する要点になる。

統計処理の道具に最も使われるSPSSは1970年

代には普及していたが、大型計算機センターを行き来しスペルエラーや手順のミスのたびパンチカードにデータとプログラムを打ち込む作業があった。1980年代半ばになるとパーソナルコンピュータが研究室に普及し電話回線で操作が可能になるが莫大な時間ロスを要求する作業があった。デジタルdigitalを道具として自由に扱える環境を享受している現在は、統計も現象をまとめ検証するための道具として自由に使い、データとコミュニケーション、データをコーディネートする習慣を身につけるとよい。必要条件は統計なんてこわくないという直感と信念であるかもしれない。

文 献

- 柳修平：情報検索の世界，縣俊彦編，上手な情報検索のためのPubMed活用マニュアル，改訂第2版，東京，南江堂，2-29
- 柳修平：統計なんてこわくない，第1版(第3刷)，金芳堂，京都，2004
- 柳修平：はじめてのレポート・卒論，第1版，金芳堂，京都，2003
- 柳修平：統計，縣俊彦編，EBMのための臨床疫学，改訂第2版，東京，中外医学社，101-126，2017

