

基礎心理学：統計

手元にあるデータは、そのままでは数字や文字にすぎないため、それらを何らかの形でまとめたり、利用しやすい形に加工したりすることが不可欠です。その際、個人の独善的な判断・解釈に陥らないよう、何らかの科学的方法によって進めることが重要となりますが、そのためのひとつのツールに**統計**があります。

もしかすると、公認心理師の活動に統計は必要ないと思われるかもしれませんが、しかし、心理的アセスメントや心理療法は、統計的分析により得られた研究成果をベースにつくられていますし、心理検査の結果の評価にも各種の統計手法が登場します。したがって、それらを適切に用いるためには、**統計**の各手法について、その内容と目的を理解しておくことが必要です。

キーワード

- 代表値と散布度
- 尺度水準
- 相関係数
- 回帰分析
- 仮説検定
- ノンパラメトリック検定
- 因子分析
- 構造方程式モデリング
- マルチレベル分析
- メタ分析

分布の特徴を表す方法

データがどのような値をとるのかを**分布**といますが、その分布の特徴をできるだけ簡単に表すには、まずは視覚的にグラフや表を使って表現するとよいでしょう。データの各値がそれぞれいくつあるかを示す**度数分布**を表にした**度数分布表**、横軸に各値、縦軸に度数を図示した**棒グラフ**や**ヒストグラム**があります^{注1}。

また、数値を使って分布の特徴を表現するものに**代表値**と**散布度**があります。

代表値とは、ある集団のデータの特徴を代表的に表したものです。代表値には**平均値**、**最頻値**、**中央値**があります（表1）。代表値としては平均値がよく使われますが、データに極端な値（外れ値）や分布に大きく偏りがある場合には、平均値が代表値として必ずしも適切ではないため、中央値や最頻値も用いられます^{注2}。

同じ平均値でも、図1のAとBとでは平均値は同じですが、分布の特徴に大きな違いがあります。このような分布の広がり（ばらつき）具

注1：棒グラフは各値に連続性のない質的変数を表すのに使われるため、度数を示す各棒どうしの間隔があいている。一方、ヒストグラムは連続性がある量的変数を示すのに使われるため、各棒どうしがくっついているという違いがある。

注2：日本人の所得金額（2016年）の平均値は約546万円（約9%）だが、最も割合の多い（最頻値）のは「200～300万円未満」の約14%で、中央値は428万円、平均金額以下の人の割合は約61%である。これを図示すると左側に頂点をもつグラフとなり、この例では平均値が日本人の特徴を代表するものとしてはしっくりこない（平均値だと少数の高額所得の人の影響で高めになってしまう）。このように、目的に対して適切かどうかを検討したうえで使い分ける必要がある。

合を示す指標に**散布度**があり、**分散**や**標準偏差**などが使われます(表2)。

通常、この代表値と散布度の値でデータ全体の特徴を表現し、なかでも平均値 (M) と標準偏差 (SD) がよく用いられます。

表1

平均値	すべてのデータの値を足して、データ数で割ったもの (例) 1、2、3、4、5 $\Rightarrow 1+2+3+4+5=15$ 、 $15\div5=3$ \Rightarrow 平均値は3
最頻値	データ全体の中で最も個数の多い値のこと (例) 1、2、2、3、4 \Rightarrow 最頻値は2
中央値	データを小さい方から並べた際に真ん中にくる値のこと (データ数が偶数の場合は真ん中2つの値を足して2で割った値) (例) 1、2、3、4、5 \Rightarrow 中央値は3 1、2、3、4 \Rightarrow 中央値は $(2+3)\div2=2.5$

表2

分散	$\text{分散} = \frac{(\text{値}-\text{平均値})^2\text{の和}}{\text{データ数}}$
標準偏差	分散に $\sqrt{\quad}$ (平方根)をつけたもの $\text{標準偏差} = \sqrt{\text{分散}} = \sqrt{\frac{(\text{値}-\text{平均値})^2\text{の和}}{\text{データ数}}}$

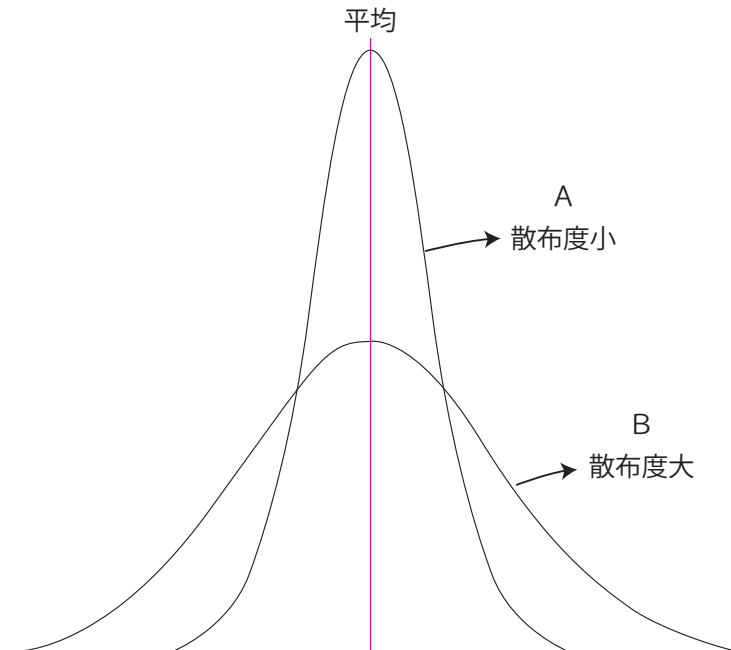


図1

尺度の種類（尺度水準）

データを表すために用いるものさしのことを**尺度**といいます。尺度には**名義尺度**、**順序尺度**、**間隔尺度**、**比率尺度**の4つの水準（**尺度水準**）があります（表3）。このうち、名義尺度と順序尺度のデータを**質的データ**、間隔尺度と比率尺度のデータを**量的データ**といいます^{注3}。

相関と相関係数

相関とは、身長と体重、親子の身長など、2つ（以上）のものどうしの**関連性**を表すものです。相関は2つの間の直線的関係を扱っており、一方（x）が増えるともう一方（y）も増える正比例の関係を**正の相関**、一方（x）が増えるともう一方（y）が減る反比例の関係を**負の相関**、このような関連がみられない場合を**無相関**といいます。関連の様子は**散布図**（図2）という図で示したり、相関の強さを0～±1の数値で表す**相関係数**を使って表現します^{注4}。

表3

質的データ	名義尺度	カテゴリーに分類するもの。数字は分類上つけただけであり、順番に特別な意味はない。（例）1=男／2=女
	順序尺度	カテゴリー分類（名義尺度）に順序や大小関係が加わったもの。それぞれの間隔は等しいとはいえない。（例）服のサイズ:S、M、L／国の大きさ順:1ロシア、2カナダ、3アメリカ
量的データ	間隔尺度	順序尺度（カテゴリー分類+順序性）に間隔性が加わったもの。各数値の間隔が同じ（一定）。足し算・引き算、平均の算出が可能。ただし、原点がない（0が何もない状態を表さない）。（例）気温、テストの点数
	比率尺度	間隔尺度（カテゴリー分類+順序性+間隔性）に比率性（0が何もない状態を表す）が加わったもの。足し算・引き算・掛け算・割り算、平均の算出が可能。（例）身長、体重

注3：扱うデータがどの尺度か、正規分布とみなせるか、各群の分散が等質であるか、またデータ数はどの程度かなどによって使用可能な統計手法が異なる。そのため、分析を進める際にはこの違いを押さえることが重要である。

注4：相関係数は、1に近づくほど相関が強いことを示し、正の相関は+、負の相関は-の符号をつけて表す。大まかな目安として0～±0.2程度が「ほとんど相関がない」、±0.2～±0.4程度が「弱い（低い）相関」、±0.4～±0.7程度が「中程度（かなり）の相関」、±0.7～±1を「強い（高い）相関」と表現される。相関係数には主なものに**ピアソンの積率相関係数**（通常これを単に「相関係数」と呼ぶ）があり、量的データに用いることができる。順序尺度の場合には**順位相関係数**を用いる。なお、質的データの関連は「**連関**」という。相関はあくまで関連性を扱っており、因果関係の有無までは扱っていない。関連性をみるふたつの変数以外のものの影響を受け、見た目のうえで相関関係が示される場合があり、これを**疑似相関**（見かけの相関）という。

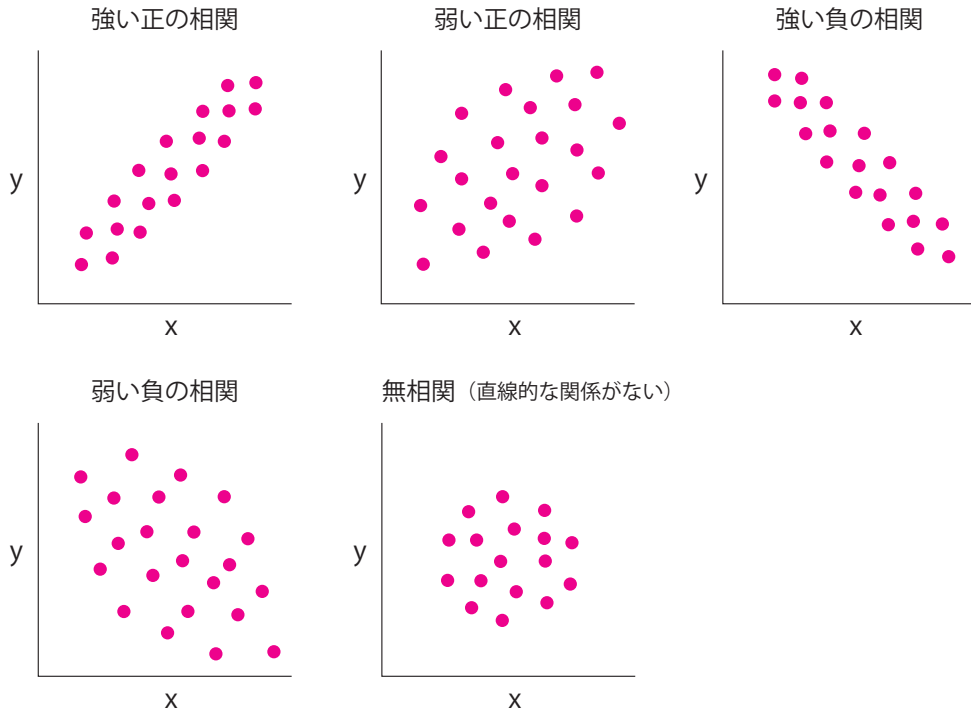


図 2

母集団と標本

心理学では、ある属性（たとえば日本人、大学生、成人など）の心理的特徴やそれに関する法則・理論を導き出そうとします。その際、属性にあてはまる人全員を調べれば、たしかにその属性の特徴ははっきりしますが、それはあまり現実的ではありません。そこで、実際はその中から数人をピックアップし、その人たちの特徴をまずは調べることで、その属性全体の特徴を推測するという方法がとられます。

ここで、属性全体の集団を**母集団**といい、そこから取り出されたものを**標本**といいます。母集団から標本を抜き出す際に偏りがあると、標本の結果も偏ってしまい、標本から母集団の平均値や標準偏差（これらを母数という）を正しく推定できません。そのため、できるだけ偏りなく抜き出すために**無作為抽出法（ランダム・サンプリング）**^{注5}が使われます。

注5：これをたとえると、鍋でつくったみそ汁の味見に似ており、味見をする際に中身をかきまぜたうえで（均質にして）一部をよそい、それを味見することで鍋全体のみそ汁の出来栄の評価や味の加減をするのと同様の考え方になる。この場合、鍋のみそ汁が母集団で、味見のためによそったものが標本となる。このように、標本の特徴から母集団の性質を推測することを**推定**といい、1つの値で推定値を示す場合を**点推定**、一定の幅をもたせた範囲で推定値を示す場合を**区間推定**という。

正規分布と標準得点

自然現象や社会現象の多くのものの分布は、平均値を中心に左右対称のつりがね型をした**正規分布**という**確率分布**（各値とそれが生じる確率を示したもの）にあてはまると考えられています。正規分布は、図3のように平均値（ M ）と標準偏差（ SD ）を使って、中心からどの範囲内に全体の何%があてはまるかがわかります。平均が0、標準偏差が1の正規分布を**標準正規分布**といい、母集団における個人の相対的な位置づけがわかるように変換した得点を**標準得点**といいます（図4）。標準得点を用いると、たとえば、ある人のテスト得点が母集団全体の中で高いか低いかなど、どの位置にあるかを把握するのに役立ちます^{注6}。

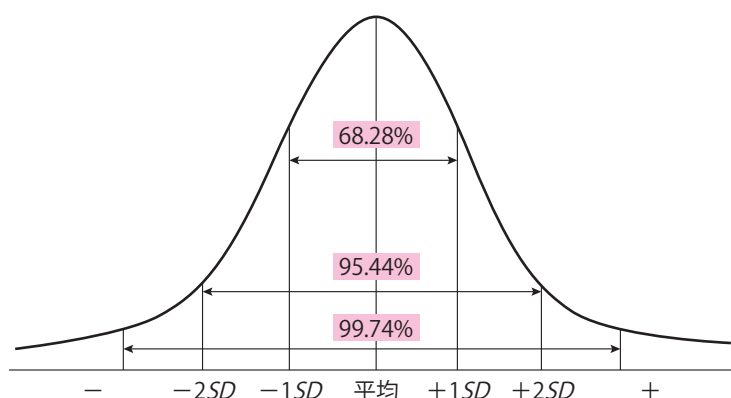


図3

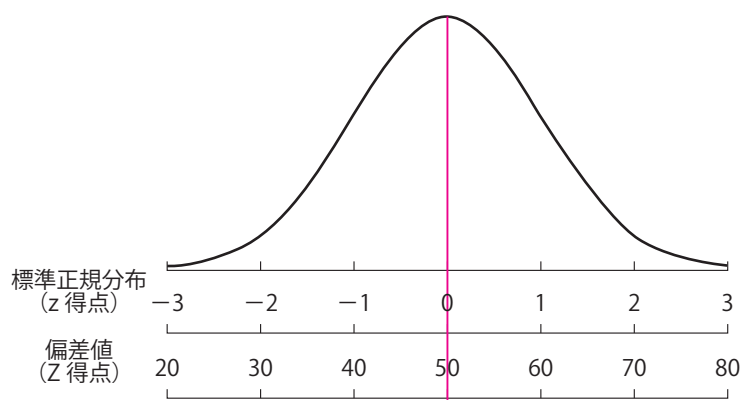


図4

注6：標準得点のうち、平均（ M ）が50、標準偏差（ SD ）が10になるように変換した得点を**Z得点**、また、Z得点が正規分布に近似するように変換した得点を**T得点**といい、これらを**偏差値**ともいう。

仮説検定

見かけ上、データに差がみられた場合に、それが実際の差によることもありますが、差があったとしても大した差ではないこともあるし、たまたま差が出ただけの場合も考えられます。そこで、得られたデータが実際に意味のあるものかどうかを、確率に基づいて統計的に判断するための方法として**仮説検定**が使われます。

仮説検定の考え方は次のとおりです。まず、あるふたつの群^{注7}の平均値に「差がある」という、本来主張したい仮説（=**対立仮説**）を立て、もうひとつ、それとは反対の「差がない」という仮説（=**帰無仮説**）を立てて、帰無仮説に対して検定を進めます。というのも、あくまで標本という一部のデータからの推測であるため、たとえ「差がある」結果が得られたとしても、本当に「差がある」といい切ることはできません。そこで、実際に得られた結果が、まれにしか起こらない程度に大きな「差」が示された場合には、「差がない」とはいいい切れない→意味のある（統計的に**有意な**）「差がある」とみなす、というものです（帰無仮説が棄却され、対立仮説が採択される）。この統計的に有意かどうかの基準となる確率のことを**有意水準**または**危険率**といい、通常5%や1%を用います。

標本の分析結果から母集団の特徴を推測しているため、いつでも正しく判断できるとは限りません。帰無仮説が正しいのに誤って帰無仮説を棄却してしまうことを**第一種の誤り**、帰無仮説が間違っているのに棄却できない誤りのことを**第二種の誤り**といい、間違っている帰無仮説を正しく棄却できる確率のことを**検定力**といいます。

仮説検定の種類には、**t検定**、**分散分析**、 **χ^2 検定**などがあります。

t検定

t検定は、たとえば実験群と統制群の比較、事前テストと事後テストの得点の比較など、**2群間の平均値の差**を検定するものです。t検定には、**対応のあるt検定**（2群とも同じ対象者の場合）と**対応のないt検定**（2群で対象者が異なる場合）があります。3群以上の比較の場合は次の分散分析を用います。

注7：たとえば、ある心理療法が不安を減らす効果があることをデータで示すために、それを実施した群と実施しなかった群に不安尺度を実施し、その得点を統計により比較して、実施群の不安の方が低いとみなせれば、その心理療法は不安に効果がある、と考えることができる。

分散分析

3群（水準）以上の差を検定する場合には**分散分析**を用います。なお、分散分析で有意であるということは、比較する群（水準）間のうちのどこかに有意差があるということまでしかわかりません。そのため、具体的にどことどの間の差なのかを**多重比較**というもので検討する必要があります。

また、あるテストの得点を時期（実験前・実験後）と群（実験群・統制群）で比較するように、独立変数が2つ（2要因）の場合の差を検定するためには**2要因分散分析**を用います。ここでは、要因Aの**主効果**と要因Bの**主効果**、**交互作用効果**（要因間の影響の大きさや向きが、要因の水準により異なること）という3つについて検定を行います^{注8}。

χ^2 検定

χ^2 検定は、**名義尺度**のデータに対する検定法であり、**クロス集計表**という表を使って値を示します（表4）。カテゴリーが3つ以上の場合には、どのカテゴリーにおける差なのかを検討するために**残差分析**を行います。すでにわかっている日本人の血液型の割合と、実際に調査した人の血液型の割合を比べる時のように、理論的に出された値と実際に標本から得られた値とのずれを調べる**適合度の検定**や、血液型の割合について男女での違いを調べる時のように、2つのカテゴリーで違いがあるかどうかを調べる**独立性の検定**があります。

パラメトリック検定とノンパラメトリック検定

検定の方法には、大きく分けると**パラメトリック検定**と**ノンパラメト**

表4

	賛成	反対	計
男性	65	35	100
女性	58	42	100
計	123	77	200

注8：2要因分散分析には、**被験者間**（対応のない）**計画**、**被験者内**（対応のある）**計画**、一方が被験者内、もう一方が被験者間の組み合わせである**混合計画**という3つがある。交互作用効果が有意であった場合、要因の水準ごとにほかの要因の影響を検討する**単純主効果の検定**を行う。

リック検定の2種類があります。

パラメトリック検定とは、量的データであり、正規分布とみなせる場合に用いることのできる検定法で、**t検定**や**分散分析**などがあてはまります^{注9}。

一方、ノンパラメトリック検定は、名義尺度や順序尺度の質的データや、量的データであっても正規分布を仮定できない場合やそもそも分布がわからない場合に用いられる検定法です。 **χ^2 検定**や**マン・ホイットニーのU検定**、**ウィルコクソンの符号和順位検定**、**クラスカル・ウォリスのH検定**、**フリードマン検定**などがあり、表5にあるように条件によって使い分けれます。

多変量データ解析

多変量データ解析とは、多くのデータをまとめて分析する方法です。具体的な手法として**重回帰分析**、**因子分析**、**主成分分析**、**クラスター分析**、**判別分析**、**数量化I・II・III・IV類**、**パス解析**、**構造方程式モデリング**、**マルチレベル分析**、**メタ分析**などがあります。

重回帰分析

重回帰分析とは、たとえば、抱えるストレスとコーピングの数からストレス反応の得点を予測するように、あるもの複数 (x_1, x_2, \dots) から、あるひとつのもの (y) をどのくらい予測・説明できるかを分析する方

表5

	パラメトリックな手法	ノンパラメトリックな手法
対応のない2群	対応のないt検定	マン・ホイットニーのU検定
対応のある2群	対応のあるt検定	ウィルコクソンの符号和順位検定
独立した多標本	被験者間分散分析	クラスカル・ウォリスのH検定
関連した多標本	被験者内分散分析	フリードマン検定
相関（連関）	ピアソンの積率相関係数	ケンドールの順位相関係数 スピアマンの順位相関係数

注9：パラメトリック検定の前提条件としては、正規性以外にも等分散性やデータの独立性などがある。データの独立性とは、独立変数以外の剰余変数が従属変数に対して、群・条件によって偏った影響を与えていない状態をさす。

注10：複数の変数（=値が変化するもの）の因果関係を考える場合に、原因となる方を**独立変数**（説明変数、予測変数）、結果や予測される方のことを**従属変数**（目的変数、基準変数）という。回帰分析で結果が示されたとしても、実際に変数どうしの因果関係の有無については、あらかじめ実験などによって検討する必要がある。

法です^{注10}。x と y のように一対一の関係は**単回帰**といいます。

複数の変数間の関係を、(重) **回帰式**という数式で表し、**決定係数** (寄与率) という指標を使って、回帰式に対するデータのあてはまりのよさ (どのくらい説明できるか) を評価します。各 x から y への影響の程度は**偏回帰係数**という数値で表します。

なお、重回帰分析は x、y ともに量的データの場合に用いるもので、x と y のどちらかに質的データを含む場合には、表6のように各手法を使い分けます。

因子分析と主成分分析

因子分析とは、質問紙の各項目や変数に潜在的に影響を与えている**共通因子**というものを想定し、**共通因子**と各変数との影響の大きさ (**因子負荷量**という数値で表す) によりデータを集約する方法です。また、**主成分分析**とは、主成分という各変数を合成した新しい変数をつくることによりデータを集約する方法です (図5)。

表 6

	独立変数・説明変数	従属変数・目的変数
重回帰分析	量的データ	量的データ
数量化Ⅰ類	質的データ	量的データ
判別分析	量的データ	質的データ
数量化Ⅱ類	質的データ	質的データ

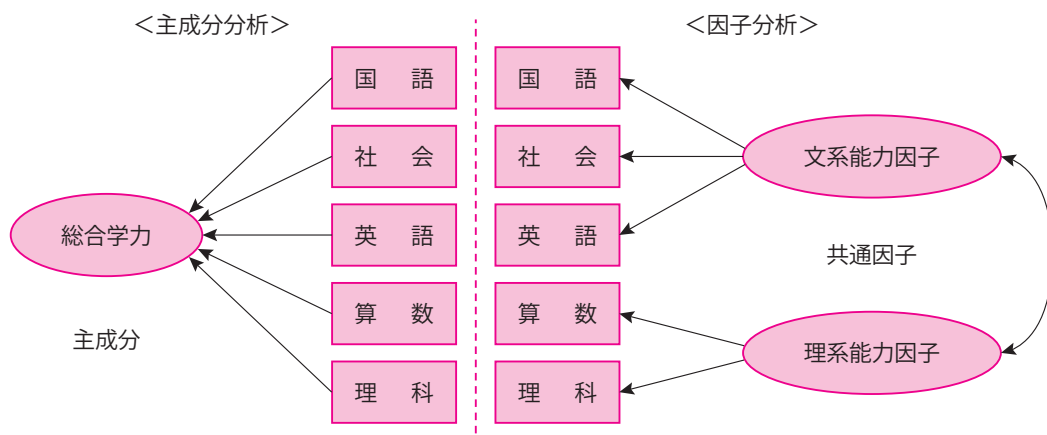


図 5

クラスター分析

クラスター分析とは、類似した回答傾向を示す対象者どうしを、その類似度によってクラスター（群）としてまとめるための方法です。対象者どうしの類似度を距離で表す**デンドログラム**（樹状図：図6）というグラフが作成され、対象者がある特性によってグループ分けする際に用いられます。比較的少ない対象者に用いる**階層クラスター分析**と多くの対象者の場合に用いる**非階層クラスター分析**があります。

パス解析

パス解析とは、重回帰分析や共分散構造分析を用いて、変数間の因果の方向や大きさなどを分析するものです。**パス・ダイアグラム**（パス図：図5）という図を使ってそれらを表現します。変数は□で囲み、因果の方向は→で、相関関係は両矢印で示し、その大きさの程度を数値で表します。

構造方程式モデリング（共分散構造分析）

構造方程式モデリング（共分散構造分析）は、変数間の関係をモデル化して、それらの因果の方向や大きさを**パス・ダイアグラム**により図示し、仮説モデルに対する観測データのあてはまりを**適合度**という指標を使って評価する方法です。直接観測できない変数（**潜在変数**：楕円で表す）を導入することが可能であり、重回帰分析やパス解析、因子分析を合わせた方法といえます。

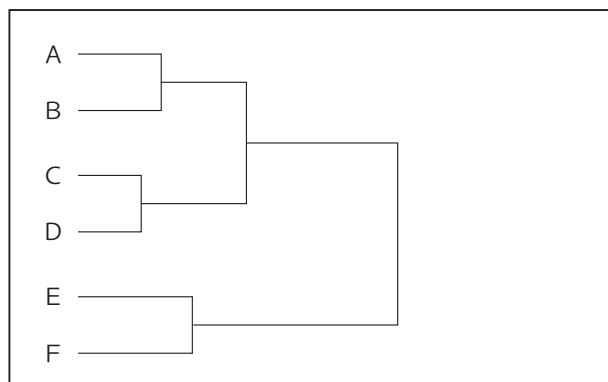


図6

マルチレベル分析

たとえば、複数の病院の患者を研究対象者とした場合のように、分析しようとするデータの中に、複数の集団を含んでいる場合には、所属する**集団内の類似性（相関）**による影響を考慮する必要があります。このように、**階層性のあるデータ**に対する分析法として**マルチレベル分析**があります。

メタ分析

たとえば、複数の研究によって、うつ病患者に対する認知行動療法の効果研究が蓄積されている場合に、それらの結果をまとめて捉える方法があれば、有効性に対するより妥当な判断を行うことができますでしょう。

メタ分析は、このように、ある研究テーマに関する複数の研究結果を統合し、評価するための統計的方法です。

テスト理論

テスト理論は、テストを開発するための理論のことであり、主なものに**古典的テスト理論**と**項目反応理論**があります。

従来多くのテストは古典的テスト理論に基づいて作成されていますが、そのテスト結果は、受験者の特性・能力だけでなく、受験者集団やテストの難易度などの影響を受けてしまう欠点があります。

一方、項目反応理論を用いることで、各項目の難易度・識別力を求めることができ、そこから受験者の特性・能力を捉えることもできます。

まとめ

統計は、得られたデータをもとに、研究目的に沿った結果を導くために使われるものです。そのデータの特徴は、平均値などの代表値や分散、標準偏差などの散布度で表します。複数の変数について、その関係の強さをみるものに相関・相関係数があり、変数どうしの因果関係の予測や説明を行うものに重回帰分析があります。平均値を比較する際には t 検定や分散分析が使われ、名義尺度データの分析には χ^2 検定が使われます。多くのデータを圧縮・整理する方法として、因子分析や主成分分析、クラスター分析などの多変量データ解析があり、近年では構造方程式モデリング（共分散構造分析）もよく用いられます。階層性のあるデータの分析法にはマルチレベル分析があります。複数の研究結果の統合的評

価にはメタ分析が使われます。いずれの手法もその適用に一定の制約がありますので、データの条件に合わせて適切な手法を選択することが重要です。