

「さまざまな多変量解析」

■今後の予定

- ・前回まででクラスター分析は終わり。
- ・次回からしばらく「回帰分析」。
- ・回帰分析は、ほとんどの多変量解析の基礎になっている。使用頻度も高い。
- ・「クラスター分析」+「回帰分析」+「 ? 」(もう1つ)
- ・テキストで扱われている中で代表的な技法は、以下のとおり。
- ・授業で取り上げない技法についても、だいたいどんなものなのかを知っておこう。

	技法の目的	技法の名前	具体例のテーマ	難度
2-3	連続的な変数の原因を説明する	一般線型モデル (GLM)、回帰分析、分散分析	家事分担と不公平感	2
2-4	カテゴリーに分けられた社会事象の原因を調べ予測する	プロビット分析、ロジット分析	社会階層と教育機会	2
2-5	異なる分析レベルの因果を同時に考える	階層線型モデル (HLM)	社会的な不平等と学校	3
2-8	3つ以上の変数の因果関係をモデル化し関係の強さを調べる	パス解析、構造方程式モデル	社会的地位はどのように形成されるか	3
2-9	社会学的概念を測定し、その因果関係をあきらかにする	共分散構造分析	地位達成アスピレーションと社会階層	3
2-10	ある社会現象が生じるまでの時間の長さを予測する	イベントヒストリー分析	勤続と離職	3
3-6	社会的カテゴリー間の結びつきのパターンをあきらかにする	ログリニア・モデル	母娘の家族形成プランの類似性	2
3-7	複数の連続変数間の類似性を検討し要約する	因子分析	権威主義的攻撃とF尺度	2
3-8	複数のカテゴリー変数の類似性を検討する	双対尺度法と数量化Ⅲ類	趣味と文化的慣習行動	2
3-9	類似性にしたがって、分析対象をいくつかの集団に分ける	クラスター分析	社会的地位の非一貫性	2
3-12	複数の変数を重みづけて新たな合成変数を作る	主成分分析	都市度	2

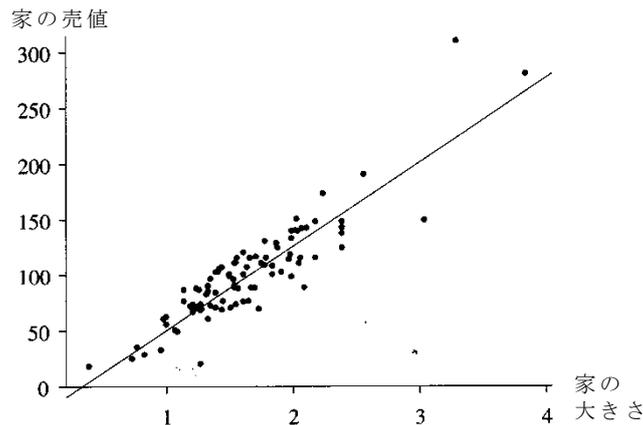
*** 宿 題 ***

- ①テキスト『社会の見方、測り方』の2-3節 (pp.95~112) を読んで、「一般線型モデル」「回帰分析」「分散分析」が、それぞれ何を指しているのか、自分の言葉で整理する。(たぶんわかりにくいですが、間違ってもよいので、まず自分の考えをもつ)
- ②2-3節を読んでいて、わからなかった概念や文に下線を引いておく。
- ③上表の代表的な技法について、教科書を流し読みして、最後の1つの技法として何を取り上げてほしいか、希望を決めてくる (来週、尋ねます)。

■ 回帰分析（一般線型モデル、分散分析と関連）【2-3】

1つの量的変数の値がケースによって違うのがなぜなのかを、別の（複数の）変数で説明しようとする分析技法。たとえば、人によってIT技術に関心が高かったり低かったりするのなぜなのかを、その人の年齢や仕事、趣味などによって説明しようとする。具体的に独立変数の値が1段階違うことで、従属変数の値がいくら上下するのかを、統計的に推定する。

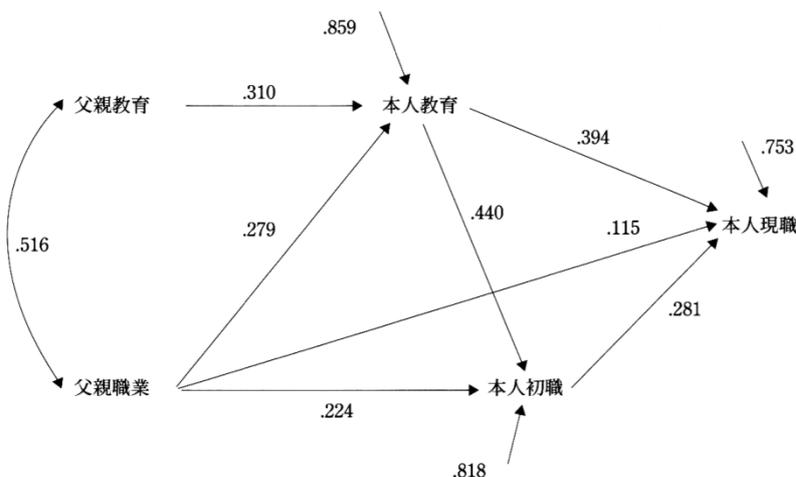
関心の中心になる変数が決まっていて、とにかくその値を左右する原因を知りたい、という考え方は、ごく自然なものなので、回帰分析はもっとも頻繁に活用される技法である。



■ 構造方程式モデル（SEM、共分散構造分析、パス解析）【2-8、2-9】

構造方程式モデルは、関心のある複数の変数間の関係性を1つの図式に整理する分析技法である。関係の有無と方向性を矢印の記号で表わし、その関係の強さを0~1の間の数値で表わす。

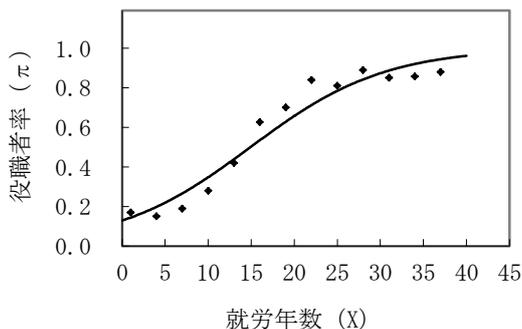
直感的に意味がわかりやすい構造方程式モデルは、とくに同じ視点からいろいろな集団を比較するとき（たとえば、年齢層による違いを比較したり、国・文化による違いを比較したりするとき）に、有効な技法である。



■ロジスティック回帰分析（ロジット分析）【2-4】

ロジスティック回帰分析は、基本的に回帰分析の仲間である。従属変数（説明したい変数）が連続的な量的変数ではなく、カテゴリーを表わす質的変数の場合に用いられる。たとえば、タバコを吸う原因は何なのか、田舎に住む原因は何なのか、といったようなことを問題にする。

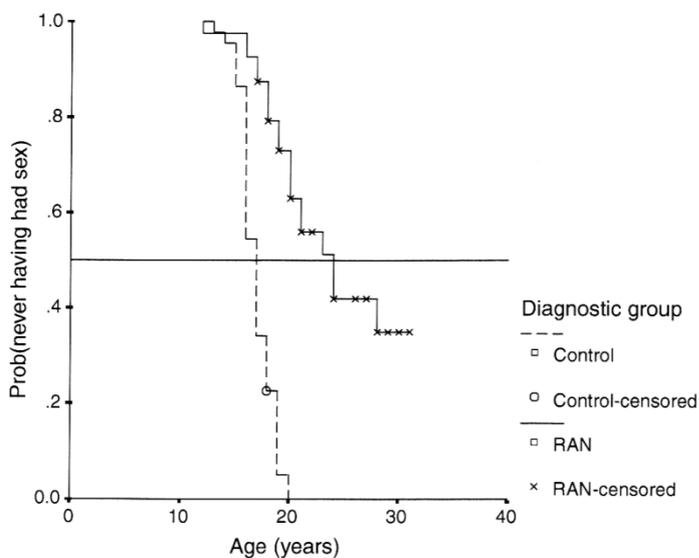
社会調査のデータには、連続的な量的変数では測りにくいものが多いので、ロジスティック回帰分析は、実際に重宝する。



■イベントヒストリー分析（生存分析）【2-10】

イベントヒストリー分析は、時間が経つにつれて「ある状態」に移行する人が徐々に増えていくようなことが想定される現象について、どういう人が早く移行して、どういう人が遅く移行するのかを調べる技法である。たとえば、「就職先が決まる」という状態に移行する時期の早い／遅いについて考えたりすることができる。

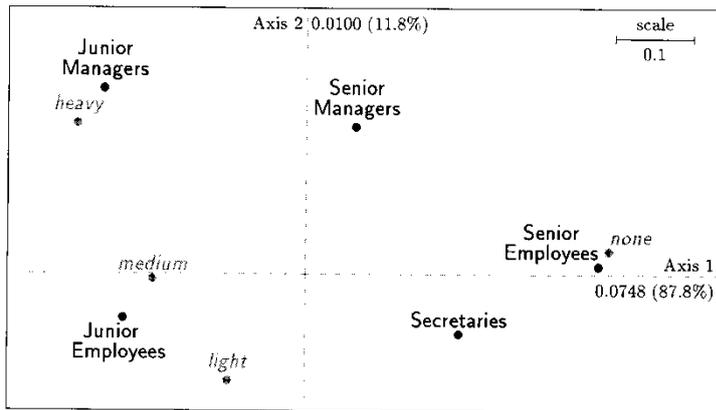
就職や結婚、出産など、社会学でこの種のデータを問題にすることは意外と多い。ロジスティック回帰分析でも似たような分析はできるが、イベントヒストリー分析を用いることの1つの大きな利点は、打ち切りデータ（調査時点ではまだ就職していないので、いつまで就職しない状態なのかが不明なデータ）を含めた分析ができることである。



■対応分析（コレスポネンス分析、相対尺度法、数量化Ⅲ類）【3-8】

対応分析は、質的変数同士の対応関係を1つの図式に整理することができる分析技法である。扱っている変数が連続的な量的変数であれば、散布図を描くなどの方法で、その関係性を直感的に表現できる。これに対して、対応分析では「仕事の種類」と「趣味の種類」のように、数量にできないようなカテゴリー同士の関係（たとえば、管理職の人はゴルフが趣味の人が多、など）を図式にする。

社会学では、やはりこのように1つの数直線上に表現できないデータ（質的変数）を扱うことが多いので、対応分析は意外と便利な場面が多い。



喫煙程度と役職の対応

■ログリニア・モデル（対数線型モデル）【3-6】

ログリニア・モデルは、クロス表で表現される変数間の関係性を整理するための分析技法である。2変数のクロス表であれば、よく見ればその内容を読み間違えることはないが、3変数以上のクロス表では、どの変数とどの変数がどう関係しているのか、読み取るのは難しい。ログリニア・モデルでは、たとえば、「性別」と「仕事の有無」と「結婚」の3変数のクロス表から、何と何の関係しているのかを整理してくれる。

繰り返すが、社会調査のデータでは質的変数が多いので、クロス表で結果を整理する場面が多く見られる。ログリニア・モデルは、クロス表を作る延長線上で多変量解析を進めようとするので、ある意味で自然な分析技法である。

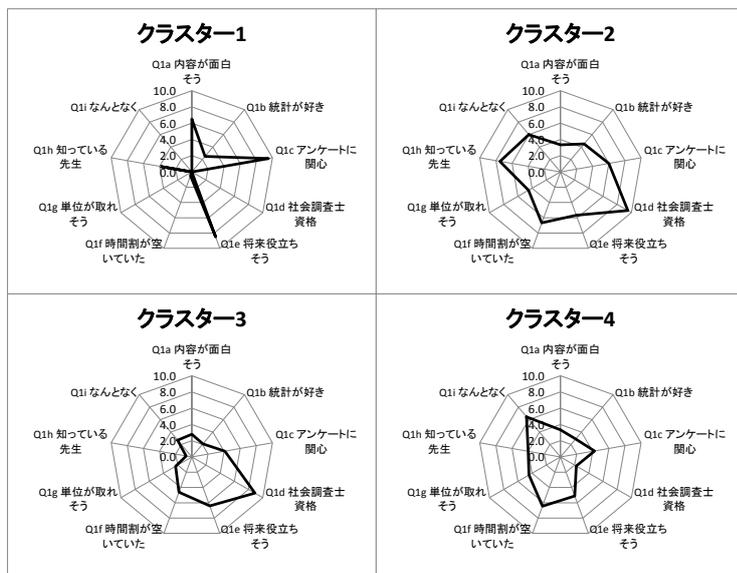
母と娘の類似性
(2×3×3のクロス表)



■ クラスタ分析【3-9】 ※終わっている

クラスタ分析は、複数の量的変数から、よく似た回答傾向の人々をグループ（クラスタ）にして、全体としてどのような種類の人々に分類できるのかを探索する分析技法である。たとえば、各科目への好き嫌いのデータから、中学生を分類して、どのような好き嫌いパターンが存在するのかを整理することができる。

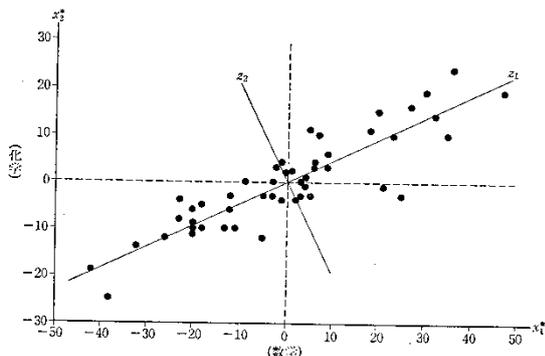
人々を分類したいという気持ちはごく自然なものであり、特別な分析技法を使わなくても直感的に分類作業を進めることもできる。しかし、クラスタ分析を用いれば、分類過程を客観化できるだけでなく、自分では思いも付かなかった分類方法が示される可能性がある。



■ 主成分分析【3-12】

主成分分析は、複数の変数から1つの合成得点を算出するための分析技法である。たとえば、5教科の好き嫌いをそれぞれ5点満点で尋ねて、合計得点（25点満点）を算出すれば、総合的な「勉強好き」得点を算出できるかもしれない。しかし、単純な合計では、科目によって好きな生徒が多い/少ないといったことが考慮されていないし、そもそも本当に1つの得点に合成できるくらい一貫性のあるものなのかも検討されていない。

複数の変数から総合的な得点を出したい場面は非常に多い。主成分分析を使えば、単純に合計するよりも適切な方法で総合得点を合成できることが、魅力である。



■ 因子分析【3-7】

因子分析は、多くの量的変数（意識項目であることが多い）が少数の見えない要素に規定されていると想定して、その見えない要素（潜在因子）を導出する分析技法である。たとえば、結婚の価値について、50個くらいの質問をしたデータから、その背後で結婚の価値を決めている要素は「愛情」「経済状態」「子どものほしさ」の3つであることをはじき出したりできる。

とくに人々の意識については、直接的に1つの質問で測りにくいような事柄が多いので、たくさんの質問の結果から、根本的に大事な要素を導き出そうとする因子分析は重宝される。

パターン行列

	因子	
	1	2
q4mnmga f 男性の幸せは結婚だ	.993	-.265
q4wnmga b 女性の幸せは結婚だ	.901	-.117
q4mghapp k 結婚する方が幸せ	.670	-.029
q4nocomg h 子どもをもつ必要はない	-.463	.051
q4wwhphh j 妻は夫の仕事を助けるべき	.371	.366
q4wwhhx e 夫は仕事・妻は家庭	.334	.569
q4wwjbia a 妻は仕事をすべきでない	.190	.562
q4jbmmfm c 働く母親もよい親子関係を築ける	.176	-.437
q4jbmmcc g 幼児の母は働くべきでない	.207	.384
q4mncodk d 男性も家事をすべき	-.005	-.371
q4wnjb2l i 女性の自立には仕事が必要	.216	-.351

因子抽出法: 最尤法

回転法: Kaiser の正規化を伴うプロマックス法

a. 3 回の反復で回転が収束しました。

■ マルチレベル分析（階層線型モデル、HLM）【2-5】

マルチレベル分析は、データの構造が複層的になっている際に、それぞれの層で適切に回帰分析をおこなうための分析技法である。たとえば、ある県の中からまずいくつかの中学校を抽出し、さらにその学校の中で何人かの生徒を抽出したようなデータを考える。このデータで、生徒の生活態度がよい/悪いことの原因を回帰分析で調べる場合、生徒個人の要因もあるが、学校単位で生活態度の改善に取り組んでいるといった、学校の要因もある。このような複層的なデータをマルチレベル分析は、適切に扱うことができる。

マルチレベル分析は、統計的に適切な分析技法で望ましい、という面もあるが、それ以上に、ふつうの回帰分析では考えられない複数の層の原因についていろいろなヒントが得られる魅力がある。

