

第8章 変量効果

グループ化されたデータ→複雑なデータ構造

- ・階層型データ
- ・入れ子型データ
- ・縦断データ
- ・空間データ

同じグループの値に相関がある→独立ではない→値同士が独立であることを仮定する分析には適さない。「変量効果」を考えることによってこのようなデータセットの分析が可能になる。

・固定効果: 定数

・変量効果: 確率変数→この値を知りたいのではなく、この値の分布パターンを知りたい。

例として、数種類の薬の効き方を知りたい→固定効果とみなす。

でも当然患者によって薬の効き方は違うはず。ただし特定の患者に対する効果を知りたいのではなく、全患者に対する効果に関心がある。

変量効果は、特定のサンプルからより広い範囲の集団の性質を推測する上で有用なツールである。

二元配置の分散分析モデルで考える。

$$y_{ijk} = \mu + \tau_i + \nu_j + \epsilon_{ijk}$$

ここで μ : 平均, τ_i : 固定効果, ν_j : 変量効果, ϵ_{ijk} : 変量効果

これらの変量効果はそれぞれ平均が0で独自の分散 (σ^2, σ_v^2) をもつ独立な正規分布に従う。

τ_i 及び σ_v^2 を推定しこれが=0であるという仮説を検定したい。

8.1 推定

固定効果モデルほど単純ではないが、最小二乗法が有用。もっとも単純な一元配置の分散分析モデルで考える。

$$y_{ij} = \mu + \alpha_i + \epsilon_{ij}$$

α : 平均0の変量, ϵ : 平均0の誤差

σ_α^2 : α (変量) の分散成分, σ_ϵ^2 : ϵ (誤差) の分散成分

<分散成分の推定>

$$\rho = \sigma_\alpha^2 / (\sigma_\alpha^2 + \sigma_\epsilon^2) : \text{級内相関係数 (ICC)}$$

水準間で変動がなければ α の分散成分は 0 だから ρ は 0。

水準間の変動が水準内の変動と比較してはるかに大きければ ρ は 1 に近づく。

変動は、 SST (全平方和) = SSE (誤差平方和) + SSA (級間平方和) のように分解して表せる。

a を水準数、 n を水準内の値とすると、 SSE の期待値が $a(n-1)$ と σ_ϵ^2 の積、 SSA の期待値が $a-1$ と $n\sigma_\alpha^2 + \sigma_\epsilon^2$ の積と表せる。このことから推定量である。 $\hat{\sigma}_\epsilon^2$ と $\hat{\sigma}_\alpha^2$ は、 MSE (誤差の平均平方) と MSA (級間の平均平方) を使用して p.154 最下部の式に変形される。

この推定方法の欠点：

1. 推定値が負の値をとることがある (現実にはありえない！)。
2. バランスドデータ (全水準の値の数が等しいデータ) の場合は平方和によって一意的に分解できるが、アンバランスドデータ (全水準の値の数が等しくないデータ) の場合このことが成り立たない。
3. 複雑な代数的計算が必要

これらの解決策として最尤法がある。誤差分布には通常正規性を仮定。

$y = X\beta + \epsilon$: 固定効果のみのモデル

X は n 行 p 列の行列 (データ行列)

β は固定効果を表す長さ p のベクトル (パラメータベクトル)

ϵ は誤差成分を表す長さ n のベクトル (ランダム誤差ベクトル)

$y = X\beta + Z\gamma + \epsilon$: 固定効果 + 変量効果モデル

Z は n 行 q 列の行列

γ は変量効果を表す長さ q のベクトル

ここで、 γ が平均 0、分散 $\sigma^2 D$ の正規分布に従うという仮定をおくと、 y の分散は $Z\gamma$ の分散と ϵ の分散の和として表され、その結果 y は平均 $X\beta$ 、分散 $\sigma^2 (I + ZDZ^T)$ の正規分布に従うとみなせる。

$y \sim N(X\beta, \sigma^2 (I + ZDZ^T))$

D がわかっていたら β は一般化最小二乗法で推定できる。 D を推定する方法としては最尤法がある。

$I + ZDZ^T$ を V とおくと、応答変数及び対数尤度は p.155 最下部及び 156 最上部の式で示される。これを用いて最尤法で β 、 σ^2 、 D を推定したいのだが、...

・多くの変量効果パラメータを含むモデルではパラメータ推定が困難になる。

・分散成分の推定に際しては偏りが生じる (因子あたりの水準数が多くない場合には、因子に関連する分散成分の最尤推定量の偏りが大きくなる)

制限つき最尤法 (REML) を使用することによりこの問題を回避できる。

応答変数と線形に結合する $k^T X=0$ となる行列 k を想定し $\{k^T X \text{ は } N(0, k^T V k) \text{ に従う}\}$ 、最尤法を実行すると不偏推定量が得られる。

<Demonstration>

オペレーターのシフトと紙の明るさに及ぼす影響を知りたい。オペレーター 4 水準にそれぞれ 5 つの観測値がある。

・固定効果モデル: 一元配置分散分析モデル → オペレーターの効果が有意

・変量効果モデル: 線形混合モデル (固定効果: 切片、変量効果: オペレーター、デフォルトで制限つき最尤法を使用)

σ^2 , σ_α^2 , $\hat{\mu}$: 一元配置分散分析と同じ推定値 ← バランスデータだから

制限なし最尤法 (method="ML") でやってみると、変量効果の分散成分が制限つき最尤法で解析したときより (0.0681 → 0.0482) 小さくなった。逆に誤差の分散成分が大きくなった。

8.2 推測

・尤度比検定

$\hat{\beta}_0$, $\hat{\sigma}_0$, $\hat{D}_0$: 帰無仮説でのパラメータの MLE (最尤推定量)

$\hat{\beta}_1$, $\hat{\sigma}_1$, $\hat{D}_1$: 対立仮説でのパラメータの MLE (最尤推定量)

尤度比検定統計量はカイ二乗分布に従う、ただしこのためにはいくつかの仮定一母数空間の条件などを満たす必要がある (文献参照)。

[固定効果の検定]

固定効果のみ異なる入れ子モデルで尤度比検定をする場合、制限つき最尤法を使用してはならない。制限つき最尤法は固定効果を除去して変量効果を推定する方法だから、通常の最尤法を使用すべき。

尤度比検定での p 値は小さくなりがちなので、変数の効果が誇張されがち。正確な p 値を求めるにはブートストラップ法がいいだろう。誤差や変量効果のように正規分布を仮定する場合の方法はパラメトリックブートストラップ法と呼ばれる。

<手順>

ヌルモデルのもとでフィットしたパラメータ推定値を用いてデータを発生させ、尤度比統計量を繰り返して算出し、有意性を評価する。

なお別のアプローチとして、変量効果パラメータを条件づけておいてその後 F 検定あるいは t 検定を行うというものがある。

[変量効果の検定]

多くの場合、変量効果の検定は誤差分散が 0 であるという仮説を含む。通常の自由度を考慮したカイ二乗分布を用いるときは、結果が保守的になる (p 値が大きくなる傾向がある) ため、有意な効果が観察されたならその結果は信頼すべきものだ。 p 値が小さいが有意ではないときは正確な値を求めるためブートストラップ法を用いるという手がある。

[平均平方期待値]

上記以外の仮説検定; 平均平方を用いた分散分析による平方和の分解。時に最尤法よりパワフルだが、正確な導出のためには、それぞれのモデルに対して多量な代数的計算が必要。しかもアンバランスデータには多くの場合使えない。

<Demonstration>

変量効果の検定のために、尤度比検定を用いて制限なし最尤法と制限つき最尤法の結果を比較した。固定効果には興味が無いので平均値を使用。

8.1 のデモでの算出結果

lmod: 分散分析結果 (p.156 下部)

mmmod: 制限つき最尤法での線形混合モデルの結果 (p.157 中部)

smmod: 制限なし最尤法での線形混合モデルの結果 (p.157 下部)

nulmod: 最尤法での線形モデル (ヌルモデル) の結果 (p.159 下部)

smmod と nulmod との尤度比検定結果

→ カイ二乗分布を用いての検定の p 値は 0.1 で 5% 水準で有意ではない。パラメトリックブートストラップでより正確な p 値を求める。尤度比検定統計量が観察値をもとに算出した 2.5684 より大きければ p 値を再び算出しよう。計算の結果、推定された p 値は 0.02 ± 0.004 (平均 $\pm$ 標準誤差) であり、5% 水準で有意である。

8.3 変量効果の予測

固定効果モデル (一元配置分散分析モデル) において μ と α の分解のためには、 $\hat{\alpha}$ を十分明らかにしなければならない → 多重比較などの推測が必要。

一方変量効果はパラメータ [定数] ではなくランダム変数で、その期待値は 0。

変量効果の事前分布とその期待値 (=0) があらかじめ分かっているとすると、ベイズ法の視点から $\hat{\alpha}$ は p.161 中下部のように表される。

<Demonstration>

経験的ベイズ推計を使用して、変量効果の予測のために最尤推定量を D (= 変量効果の分散共分散行列)、 V (= $I + ZDZ'$)、 β でおきかえてみる

mmmod で各水準の変量効果を抽出する

mmmod で各水準の固定効果を抽出する

mmmod で各水準の固定効果/変量効果を計算する

この結果、全オペレーターで同じ値 → 予測された変量効果の値は固定効果に完全に比例 → 予測された変量効果の値は過小な推定値かもしれない

$\hat{\mu} = 60.4$ を考慮して、BULPs (最良線形不偏推定量: 線形の不偏推定量の中で最も分散が小さいもの) を予測する。なおこれは固定効果と変量効果との和である。

QQ プロットでチェック → 正規性に問題なし (変量効果モデルは外れ値に敏感 → 分散成分に従属するから → なので、QQ プロットでチェックしたほうがよい)。残差プロットも誤差分散が一定であるかのチェックに使えるので重要。水準数が 4 よりも多ければその効果の正規性と等分散性を検

討すべき。

8.4 変量効果としてのブロック

ブロック: 実験の条件・実験者の判断で設定されるもの・同一時期に行われた行為のまとまり。
ブロックの効果には興味が無いので変量効果として扱う。

<Demonstration>

コーンステーパーリカーを混合したもの (5 種類) を使用して、ペニシリンの製造工程 (A-D の 4 工程) が収率に及ぼす影響を知りたい。

4 処理×5 配合の無作為化完備ブロックデザインによる分散分析の結果、処理に有意差なし、配合に有意差あり。

処理を固定効果、配合を変量効果とした、制限つき最尤法を用いた線形混合モデル (配合は現存する母集団からのサンプルというわけではないから、固定効果にはしないほうがいい) の結果、バランスデザインのため変量効果の残差の分散が分散分析の結果と等しい。固定効果の推定値も同じ。

<2 種類の固定効果の有意性検定 (p.165 上部) >

- ① 処理間の分散分析 (F 検定) → 有意差なし
- ② 処理及び切片を固定効果とした 2 つのモデルの分散分析 (カイ二乗検定) → 有意差なし

固定効果の比較なので②では制限つき最尤法ではなく制限なし最尤法を使用している。

<ブートストラップによる尤度比検定>

カイ二乗分布を前提とした QQ プロット (Fig 8.2) → プロットがずれている → 近似が悪い
ブートストラップで推定された p 値 → 分散分析の値に近い

<変量効果 (ブロックの効果) >

その大きさには直接的な興味は無いわけだが、ブロックの効果をテストしておくことは今後の実験デザインを考える上で有用であろう。

尤度比検定統計量は 0 に近い値が多く、カイ二乗分布の近似はよくない。このようなヌルモデルの分布の近似がうまくいかないケースでは、p 値が分散分析の結果と異なる。0 でない尤度比検定統計量の分布は、端の部分を除き、カイ二乗分布に近くなっている。
そして p 値は 0.05 より小さい → 配合の効果が有意 → 分散分析の値に近い。

8.5 分割区画

主たる区画が複数の下位の区画に分割され、1 つの区画は因子のひとつの水準であり、他の因子の水準は下位区画によって変化する。完全無作為化を行わない実験デザインである。

<分割区画とブロックとの違い>

ブロック: 実験デザインに盛り込むべき特性

分割区画: 因子の割り当てに際し制限を課すもの。変化速度の異なる複数の因子がある場合な

どに設定される。

<Demonstration>

作物 (2 種類) の収量に対する灌漑方法 (4 種類) の影響を、複数の圃場 (8 箇所) での実験により検討している。

- ・各圃場には 1 種類の灌漑方法を割り当て。
- ・各圃場は 2 分し、それぞれに別の作物を栽培。
- ・灌漑方法は無作為に圃場に割り当てられている。

[註; すなわち作物の種類及び灌漑方法の各水準に圃場の水準のいくつかは割り当てられている実験デザイン。作物の種類と灌漑方法は交差している。]

収量の平均値・灌漑方法と作物の種類及びそれらの交互作用を固定効果、圃場と作物の種類の交互作用を変量効果として扱う (灌漑方法と圃場との交互作用は、交差していないため計算できない)。

このモデル (p.168 上部) と交互作用の変量効果の項を除いたモデル (p.168) を比較→尤度は同じ→フィールドの変動と誤差変動が分離できない→分離するためには作物あたりの観察値を増やすことが必要。

解析の結果、分散成分が大きいのは圃場で。固定効果は交互作用も含め有意ではない。残差プロットを見ると残差の分散は概ね一定だが QQ 正規プロットを見ると両端が尾状に逸脱している。

変量効果を無視した解析結果を p.170 上部の分散分析表に示す。しかしこの結果は、圃場の無作為化の制限及びこのことにより生じた変動を考慮していないので正しくない。

8.6 入れ子の効果

入れ子: ある因子の水準が完全に他の因子の水準に含まれること

- ・業種別の労働者の能力の測定
- 各労働者が同じ場所のみで働く: 労働者は場所の入れ子
- 各労働者が複数の異なる場所で働く: 労働者と場所は交差している

<Demonstration>

作業員あるいは作業場所によって作業結果に違いが生じるかを知りたい。

均質化された乾燥卵粉の脂肪量を、8つの作業室における2人ずつの作業員に2種類のサンプルを2個ずつ測定した結果を解析する。なお作業員にはこれらは全て均一化されているという事実は伝えない。

- ・作業室ごとに固有の作業員: 作業員は作業室の入れ子
 - ・サンプルの種類 (G と H という区分) は、作業員ごとに異なるもの: サンプルは作業員の入れ子
- [註; 作業室—作業員—サンプルという二重の入れ子を有するデザイン]

実験の目的とサンプリングの無作為性から考えてこれらを変量効果として扱う

解析結果 (p.171 上～中部) を見ると、分散成分の大きさはどれも似たようなもの。脂肪含有率の測定結果のバラツキは、実験室・作業員・サンプルの分散や測定誤差に起因する。

分散成分が小さい変数を除去してモデルを単純化して解析すると (p.171 下部)、サンプルの分散成分が他の因子 [作業者と誤差] に吸収された。

正確な p 値を知るためにブートストラップ→p 値は 0.09→サンプルのバラツキは無視できる。(なお作業者のバラツキを同様にチェックしてみると、これは有意になる)

8.7 交差の効果

“入れ子”でないのを”交差”という。

- ・完全交差: ひとつの因子の全水準が他の因子の全水準と対応

- ・交差が不完全→ラテン方格デザインなど (ただし入れ子ではない)

交差がある場合には入れ子のデザインは使えない。

<Demonstration>

磨耗試験結果 (ラテン方格デザイン) に及ぼす、原材料 (A,B,C,D の 4 種類)・サンプルの位置 (4 箇所)・処理工程 (4 工程) の影響を知りたい。

固定効果の分析結果では原材料・サンプルの位置・処理工程のすべての因子が有意 (p.173 中部)。サンプルの位置と処理工程を変量効果とした分析結果でも原材料は有意 (p.173 下部)。ただし、原材料の自由度が変わる (この結果ではさほど重大な問題ではないが)。

変量効果を考慮した分析は、ブロック因子のバラツキの推定量を組み込める点で有利。

8.8 マルチレベルモデル

この言葉は階層構造を有するデータに対するモデルに使用される。社会科学分野で多い。

<Demonstration>

中学校に入ってから成績は以下の因子のうち何に影響されているのか？

出身小学校・学級・性別・社会階級 (父親の状態)・個人・知能テスト結果・英語テスト結果・数学テスト結果 (テストは年 1 回で合計 3 回—ただし生徒によっては 1~2 回—受けている)

まず、最後 (3 回目) の数学テストの結果を応答変数として分析。予測変数として性別、社会階級、初年度の知能テスト結果 (入学時の能力)。これ以前の数学テストのスコアは無視。

初回の知能テストと 3 回目の数学テストの結果には正の相関がありそう (図 8.4 左)。数学テストの結果は社会階級による違いがあるが、その上限の差は小さい (図 8.4 右)。

重回帰の結果 (p.175 中部) を見ると、性別は 3 回目の数学テストの結果との関係が薄い。性別を予測変数から外した重回帰の結果 (p.175 下部) 及び解釈の単純化のため有意となった交互作用項も外した重回帰の結果 (p.176 上部) を見ると、初年度の知能テストとの関係が強い！

ところで、この解析の前提条件であった「データセットの 953 人の生徒の観測値は独立である」という仮定は妥当なのか？

小学校内での類似性がありそう→そうだとすると独立を仮定した場合より個人に帰する情報量が少なくなる→すなわち生徒の観測値が独立であるという仮定のもとでは、誇張された解析結果が生じる→さらに小学校内及び小学校間でのバラツキを考慮していない。

データのグループ性を考慮しつつ、個人の情報を使用した解析を！

- ・固定効果：性別、社会階級、初年度の知能テスト結果
- ・変量効果：学校及びその入れ子である学級

初回の解析で重要な因子ではなかった性別を外してモデルを単純化 (p.177 中～下部) した。さらに社会階級と比較できるようにするために知能テスト結果をセンタリング (平均を 0 に) した。

3 回目の数学テストの結果は入学時の知能テストと強く関係しており、社会階級が下がるにつれて下がる傾向がある。しかし社会階級 9 のパラメータの推定値からみて、社会階級は完全に順位に沿った関係を示していない。学校と学級のバラツキは小さいので、個人に帰属するバラツキが大きい。

残差はおおむね正規性を示している (左) が、バラツキは予測値が大きくなるほど小さくなっている (図 8.5 右)。→この効果を除去したい。

変量効果の正規性のチェック (図 8.6) →学校も学級もおおむね正規性を示している

・**adjscores**: 50 の小学校の変量効果→生徒の社会階級や入学した生徒の質に応じて調整された学校のランク [註; 生徒の社会階級や質などにより調整された学校間の (機能の) 差を示す]。

・**rawscores**: 3 回目の数学テストの結果を小学校を予測変数として線形回帰した結果の、小学校の係数をセンタリングした値 [註; 生徒の社会階級や質などによる調整を行っていない学校間の (機能の) 差を示す]。

<図 8.7 の解釈>

学校 15, 9: 学校の機能は高そうに見えるが、生徒の質を考慮するとそれほどでもない。

学校 30: 学校の機能はそれほどでもなさそうに見えるが、生徒の質を考慮すると結構高い。

【合成効果】

固定効果には最下位の階層 (生徒個人に帰するもの) を使用してきたが、学校や学級という上位の階層の因子が数学テストの予測に重要である可能性もある。個人という階層の情報を使用して (上位の階層と組み合わせて) 予測変数の効果を考える→これを合成効果という

入学時の能力は、その後の成績の伸びしろにもきいてくるのではないかな？→学校ごとの知能テスト結果の予測値をモデルに取り込んで解析 (p.180) →この変数は有意ではなかった→他の共変量を使用した変量効果との関係も検討すべき。