

# 5

## 変量効果モデル

### 要 旨

反復測定を伴った研究やクラスターを単位としたランダム化試験 (cluster randomised trials) を解析する際には、変量効果モデルというのは有用な手法である。ここでは観測値はグループに分けられ、グループ内のデータは独立ではない。仮にクラスターのあることを無視してしまうと、求めたい推定値の標準誤差を低く見積もることになる。これらのデータを扱うモデルには、クラスター別モデル (cluster-specific model) と周辺モデル (marginal model) の2種類がある。周辺モデルは当てはめやすく、一般化推定式 (generalised estimating equations ; gee) とよばれる手法でも使われる。

### 5.1 はじめに

これまで記述してきたモデルでは、誤差項はたった一つしかなかった。たとえば重回帰 (multiple regression) はモデル (2.1) に示されるように、予測変数 (predictor) には誤差項  $\epsilon$  が含まれている。ロジスティック回帰での誤差項は二項分布に従い、観測値と予測値とを関連づけている。しかし、そこに誤差項を複数加えることもできる。簡単な例を示せば、個人内で経時的に反復測定されているような場合である。つまり、個人内 (within individuals : 個人で反復測定しても同じ値になるとは限らない) と個人間 (due to individuals : 個人同士は独立とみなす) による確率変動とがある。

別の例を挙げよう。複数の医師がそれぞれ数名の患者を治療する場合

## 5. 変量効果モデル

である。そこには医師の中での変動 (within-doctor variation : 患者にはいろいろな人がいるので) と医師の間での変動 (between-doctor variation : それぞれの医師は異なった治療成績を示しやすいため) がある。この例のように自然に階層ができているため、階層的データ構造 (hierarchical data structures) とよんでいる。つまり、あるデータが別の何かの中に入れ子となっている状況である。このような種類のデータに適したモデルのひとつに変量効果モデル (random effects model) がある。近年ずっとこのタイプのモデルに大きな関心が払われてきており、現在では、その結果を医学文献で普通に目にするようになってきた。本章では、このトピックの重要性について、少しでも読者の関心を引くことができればと思っている。

### 5.2 変量効果に対するモデル

患者ごとに同一の評価指標で測定しているが、それぞれの患者が違う医師から治療を受けているというような、人ごとに異なる群を形成しているランダム化試験を考えてみよう。群  $i$  の中の個人  $j$  に対して連続量の結果  $y_{ij}$  があるとし、以下のように推定する。

$$y_{ij} = \beta_0 + z_i + \beta_1 x_{1ij} + \dots + \beta_p x_{pij} + \varepsilon_{ij} \quad (5.1)$$

このモデルは式 2.1 に極めてよく似ているが、 $z_i$  項が追加されている。

ここで  $z_i$  は  $i$  群の効果を表しており、 $E(z_i) = 0$ 、 $\text{Var}(z_i) = \sigma_B^2$  の確率変数であると仮定する。添字 B は群間 (Between groups) ということを示す。 $x_{kij}$  は  $i$  番目の群における  $j$  番目の個人の  $k$  番目の共変量であり、それに対する回帰係数を  $\beta_k$  とする。

さらに、

$$\text{Var}(\varepsilon_{ij}) = \sigma^2 \text{ と仮定すると、} \text{Var}(y_{ij}) = \sigma^2 + \sigma_B^2 \text{ となる。}$$

このように、観測値の変動は群内と群間の二つの構成因子に分けるこ

とができる。

群内の観測値は相関し、 $i$  と  $j$  が異なるとき、

$$\text{Corr}(y_{ij}y_{ik}) = \rho = \frac{\sigma^2_B}{\sigma^2 + \sigma^2_B}$$

である。この値を級内相関または群内相関 [intra-cluster (group) correlation ; ICC] と呼んでいる。

個人効果  $z_i$  を無視するモデルでは、 $\beta_i$  の推定値の標準誤差はたいへい非常に小さくなることが知られている。その結果、一般に第1種の過誤が大きくなりやすい。特に、すべての群が同じサンプルサイズ ( $m$ ) であれば、推定値の分散は  $(1+(m-1)\rho)$  倍大きくなり、このことはデザイン効果 (design effect ; DE) として知られている。

モデルへの当てはめを行うとき、 $z_i$  と  $\varepsilon_{ij}$  が正規分布していることを仮定しないとイケない場合がある。もちろんいつもそうであるとは限らない。

モデル (5.1) では切片は  $\beta_0 + z_i$  であり、群 ( $i$ ) により異なる。しかもこれら切片は確率的に変動するので、**変量切片モデル** (random intercepts model) と呼ばれている。これは**多水準モデル** (multi-level models) と呼ばれるモデルの一種である。このモデルでは、階層の各水準に対して、異なる誤差項が想定されており、個々人は群内に入れ子となっている。また、変量効果と固定効果が混ざったモデルのことを**混合モデル** (mixed models) と呼ぶ。さらに、モデル (5.1) を**交換可能モデル** (exchangeable models) と呼ぶ。クラスター内の二つの観測値を交換しても、推定には影響しないからである。交換可能かどうかを見るには、別の方法もある。モデル (5.1) で  $z_i$  を固定してみる。このとき同じ群内の個人  $j$  と  $j'$  に関して、 $y_{ij}$  と  $y_{ij}'$  の相関がどの  $j$  および  $j'$  でも同じということである。

**変量係数モデル** (random coefficient model) というモデルもある。ここでは、傾き  $\beta_i$  は母集団の固定値ではなく確率変数と仮定している。こ

## 5. 変量効果モデル

れは本書で扱うレベルを超えている。

これらのモデルの詳細については、Brown and Prescott<sup>1</sup>を参照されたい。また、反復測定についてはCrowder and Hand<sup>2</sup>やDiggle, Liang, and Zeger<sup>3</sup>を、階層モデルについてはGoldstein<sup>4</sup>を参照されたい。

### 5.3 変量効果と固定効果の違い

いくつかの群を扱うモデルの中に、ある変数を組み入れることを考えよう。それは「喫煙者 (smokers)」か「非喫煙者 (non-smokers)」といった一般的なものであったり、A 医師あるいは B 医師に診てもらった患者というように、かなり特定されたものであったりするだろう。カテゴリ変数を扱う通常の方法では、2章で述べたようなダミー変数を当てはめる。ここでは、特定の群の効果は固定であると仮定している。このため固定効果モデル (fixed-effect model) と呼ばれている。すなわち、母集団パラメータが固定だとして表される。例えば、「喫煙すること (smoking)」により平均的に、ある一定量肺機能が低下することが想定されるだろう。一方で、あなたが喘息で A 医師に診てもらっていれば、それも肺機能に影響を与えるかもしれない。しかし、A 医師がどうこうというのは意味のないことであり、実際、それは単なるノイズにすぎないのである。それに引き換え、喫煙の効果というのは一般的にも興味がある。どこに解析目的があるかによって、固定効果か変量効果かが決まってくる。仮に研究が反復されるなら、そのとき同じ群を用いるだろうか。もしそうではないというなら、変量効果モデルのほうが適当である。もし群が共通なら、群の違いによる影響 (つまり交絡) をダミー変数により取り除くことができる。しかし、これらの群がこの研究にしか用いられておらず、新たな研究には別の群が含まれるなら、そういったアプローチではうまくいかない。このようなときには、通常の「誤差 (error)」項に加えて、個人や群が原因の「誤差 (error)」要因 (つまり変量効果)

を含んだモデルが必要となるのである。

## 5.4 変量効果モデルの利用

### 5.4.1 クラスターを単位としたランダム化試験

クラスターを単位としたランダム化試験とは患者個人を単位とするのではなく、患者のひとまとまりを単位にして、介入群か対照群かをランダムに割付けするものである。そうした群の例は地域であったり、開業医/家庭医 (general or family practice) や学校であったりする。開業医 (GP; general practice) による試験では、開業医やヘルスケアの専門家が積極的に参加するので、ランダム割付けは患者単位ではなく開業医やヘルスケアの専門家単位となる。介入の効果は患者ごとの結果として評価する。

クラスターを単位としたランダム化試験に関する特徴はいろいろあり、統計学的な観点については Cornfield<sup>5</sup> によって初めて議論された。また、Zucker<sup>6</sup> からも参考になる議論をしている。大きな特徴としては、同じ医師に治療された患者同士は、違う医師に治療された患者同士よりもよく似てくるということである。

もし、私たちが事前にどの患者がどの医師に治療されたかということを知っていれば、その患者の状態を当てずっぽうに予想するよりはずっとよく予想することができる。すなわち、ある一人の医師に対する観測値は完全に独立ではないということである。いかに相関係数が小さくても、そうした研究のデザインや解析に驚くほどの影響を及ぼす。例えば、級内相関係数 (ICC) が 0.05 (開業医による試験ではよく見られる数値) だとして、このとき一群当たりの患者数は 20 名、治療効果に対する標準誤差の推定値は通常のもので、クラスター化の影響は無視できると仮定すると、正しい推定値よりも約 30% 低い値となる。ということは、

## 5. 変量効果モデル

効果は全くないにもかかわらず有意になる可能性（第1種の過誤）が高くなるということである。

開業医または家庭医を単位としたランダム化試験の利点・欠点については最近議論されている。

### 5.4.2 反復測定

反復測定の研究とは、個々の患者についていくつかの変数が2回以上測定されるような研究である。例として、治療後、3、6、12週、および24週で血圧を測定する降圧薬の臨床試験が挙げられる。個々の患者の血圧に対する反応は、変数 $z_i$ として測定される。

個々の患者は母集団から抽出された標本であり、特定の個人の血圧レベルに興味があるわけではない。

このタイプのデータを解析する簡単な方法の一つは、測定値を要約することである<sup>8,9</sup>。

この方法で、我々は、簡単に個々人に対する測定値の要約（それは多くの場合、平均にすぎないが）を算出し、主要結果変数として解析する。これで個人内変動を消去し、個人間変動についての誤差項だけを考察すればよいことになる。他の要約量としては、時間内に到達した最大値あるいは、直線の傾きが挙げられるだろう。反復測定に対して、データは順番に収集される。この順番にということが重要なことである。モデル(5.1)は順序を考慮した自己回帰モデル (autoregressive model) というモデルに拡張できるが、これは本書で扱う内容のレベルを越えている。

### 5.4.3 標本調査

もう一つのモデルの簡単な利用例として、プライマリケアにおけるうつ病の程度を見いだすことを目的とするような標本調査が挙げられる。医師のランダム標本は患者のランダム標本の中から選ばれる。個々の患者が特定のどの医師によって治療を受け、その効果がどうであったかと

いうことは、それほど興味の対象ではない。もし試験を繰り返すならば、医師ごとに異なるデータセットとなるが、うつ状態になった患者における割合の推定値に関して、医師間のバラツキに非常に興味がある。なぜなら、それは信頼区間に影響するからである。すなわち、我々は全体の推定値において医師間のバラツキを考慮に入れる必要がある。

## 5.5 変量効果モデルの実施

### 5.5.1 クラスタ試験

「*Diabetes from Diagnosis*」というのは、新たに診断された糖尿病患者の治療に対して集中介入するという試験であった<sup>10</sup>。簡単に言えば、41名の医師が参加し、二つの群にランダム割付けされた。すなわち、介入群に21名、比較対照群に20名である。介入群においては、医療関係者に対して、患者集中治療の技術とその根拠について1日半の導入教育を行った。彼らには患者用のパンフレットを配布し、活発に質疑してもらった。もう一方の群には、新たに診断された糖尿病患者の管理に関する英国糖尿病学会ガイドラインを配布しただけであった。ガイドラインには治療1年後のBMI (body mass index)、HbA<sub>1c</sub>のような測定値、さらに患者の満足度を示す多くの評価指標が記載されていた。重要な点は、一人の医師によって全員が治療を受けた場合は、別々の医師からそれぞれが治療を受けたよりも、よく似た治療結果を示す傾向があるということである。さらにその試験は介入を目的としたものであり、異なる医師たちによって実施されたということである。介入効果は1年後に両群間のBMIの差で1.90 kg/m<sup>2</sup> (SE: 0.82) というものであった。クラスタリングを考慮しない場合、標準誤差は0.76であった。これは明らかに効果の有意性を増大させるものである。

## 5. 変量効果モデル

### 5.5.2 反復測定

Doullら<sup>11</sup>は、吸入ステロイド薬の使用前後における喘息児童50名の成長率を観察した。その結果、治療前と治療後0~6週で成長率の差は-0.067 mm/週 (95%信頼区間: -0.12~-0.015)、治療前と治療後19~24週で成長率の差は-0.002 mm/週 (95%信頼区間: -0.054~0.051)であった。吸入コルチコステロイドの成長抑制作用が比較的短期に現れたことが示された。子供を変量効果とした変量効果モデルが適用可能であり、子供1名ごとにいくつかの評価尺度があってもよいとした。このモデルはこの子供たちについてしか当てはまらないのではなく、もっと一般の子供たちにも当てはまることになっている。

## 5.6 群ごとに通常の最小2乗法を用いること

Cornfield<sup>5</sup>は、「ランダム割付けどおりに解析すべきだ (analyse as you randomize)」と述べている。ランダム割付けは群単位でなされ各群の主要評価項目の平均値という「要約指標 (summary measures)」を計算したりする。

単純化するために介入群を1、対照群を0とするダミー変数 $\delta_i$ だけにして、共変量のないモデルを考えるとこうなる。

$$\bar{y}_i = \mu + \tau\delta_i + \bar{\varepsilon}_i \quad (5.2)$$

ここで $\bar{y}_i$ は、 $i$ 群の $n_i$ 人の測定値の平均値である。

$$\text{Var}(\bar{y}_i) = \sigma_b^2 + \frac{\sigma^2}{n_i}$$

式(5.2)は独立な誤差を含んだ単純なモデルであり、 $n_i$ が等しければ誤差も一定である。 $\tau$ の最小2乗推定値にはバイアスがかかっておらず、その標準誤差は治療効果から独立した誤差項にだけ依存する。

つまり、群ごとの単純な解析とは以下のようなものである。もし $n_i$ が群ご



とに共通かほぼ同じなら，群ごとの平均値について2標本 $t$ 検定を適用すればよい．これは5.4.2節で述べた要約尺度の手法である．しかし，もし $\sigma^2$ がゼロ（群内のすべての値が同じ）であれば群の大きさに関係なく意味がない．

群ごとというアプローチには多くの問題点がある．主要な問題点の一つは，いかにして個人の共変量を考慮するかということである．個人の共変量は群内の平均値を使うということでは不十分である．なぜなら個人ごとに共通なことはないからである．要約した値に関して変量効果モデルを用いるとなぜよいかというと，個人のさまざまな共変量を考慮できるからである．そこでは，群ごとに症例数が異なってもよいのである．

## 5.7 コンピュータ解析

### 5.7.1 尤度と一般化推定式

現在多くのコンピュータ・パッケージは，変量効果モデル(5.1)であるが推定方法はいくらか違っている．尤度法ではそれぞれの $z_i$ に分布（通常は正規分布）を仮定し，次に $z_i$ の下で $y_{ij}$ が観測される確率を計算する． $z_i$ の分布を使うため，考え得る $z_i$ すべてについて期待確率が得られる．これには難しい積分の知識が必要である．この方法では群あるいはクラスターごとに回帰係数を算出するため，クラスター別モデル(cluster-specific model)とっている．もっと単純な方法は一般化推定式(generalised estimating equations; gee)であり，そこでは変量効果の正規性を要求していない．本質的には群ごとに結果変数の平均値を用いているだけである．さらにサンドイッチ推定値(robust sandwich estimator)という安定した推定値を用い，群内相関を考慮して標準誤差を求めている．geeはいわゆる周辺モデル(marginal model)の方法である．

一般的なデータ表は右端は余白 (margin) になっており、その部分に平均値が書かれていることから周辺モデルと呼んでいる。gee の厄介なことといえば、群内相関を伴う群ごとの平均値を使う。なぜなら、群は主要な解析項目だからである。クラスター数が 20 を超えないと gee 信頼性は落ちるので、できれば 40 くらいあるのが望ましい。

この方法は、誤差が二項誤差の場合に拡張することができる。すなわち、変量効果ロジスティック回帰である。この場合、最尤法のコンピュータパッケージは真利なく、ほとんどのパッケージは gee しか使えない。

Murray<sup>12</sup> は SAS を用いたプログラムを提供している。STATA<sup>13</sup> は横断時系列モデル (cross-sectional time-series models) と呼ばれるモデルプログラムを統合しているが、本質的には繰り返しデータやクラスター別のデータである。このタイプのモデルは、極めて自然にベイズ流アプローチ ([付録 4] 参照) に一般化される。これは、この本の目的を超えているし、詳細については Turner, Omar and Tompson<sup>14</sup> により述べられている。

### 5.7.2 コンピュータアウトプットの解釈

表 5.1 では、Kinmonth ら<sup>10</sup> によるデータの一部が示されている。これは 10 名の医師により治療された患者の 1 年後の BMI である。

解析結果は表 5.2 に示されている。表 5.2 (i) は、クラスターを定めない、一般的な回帰を当てはめた結果である。治療効果の推定値が 0.42、標準誤差は 1.90 である。以前に述べたように、クラスターを無視したため標準誤差は非常に小さくなっている。表 5.2 (ii) では、最尤モデルを当てはめた結果が示されている。治療効果の推定値は 0.39、標準誤差は 2.47 である。クラスターリングを考慮しないモデルに比べて標準誤差が増大している。さらに、クラスター間相関係数  $\rho$  の推定値は 0.85 と計算してくるし、その群間標準偏差 ( $\sigma_u$ ) や群内標準偏差 ( $\sigma_e$ ) も計算してくる。このアウトプットでは、変量効果として正規分布ある

表 5.1 BMIに関するデータ

対象	BMI(kg/m <sup>2</sup> )	治療法	プラクティス
1	26.2	1	1
2	27.1	1	1
3	25.0	1	2
4	28.3	1	2
5	30.5	1	3
6	28.8	1	4
7	31.0	1	4
8	32.1	1	4
9	28.2	1	5
10	30.9	1	5
11	37.0	0	6
12	38.1	0	6
13	22.1	0	7
14	23.0	0	7
15	23.2	0	8
16	25.7	0	8
17	27.8	0	9
18	28.0	0	9
19	28.0	0	10
20	31.0	0	10

いはガウス分布（ドイツの数学者で最初に正規分布を記述した C. F. Gauss に因む）を仮定した。表 5.2 (iii) では一般化推定式の手法を用いている。同様に治療効果の推定値として 0.39 となっていて、標準誤差は 2.61 となっている。このモデルの基本的仮定は、群内の個人データは交換可能 (exchangeable) ということである。gee による推定値は、表 5.2 (iv) で群ごとの平均を用いた方法と比較するとよい。ここで得られた推定値は 0.41 であり、標準誤差はやや大きく 2.74 であった。

この方法を使うと推定値が異なり、群間バラツキも大きくなりがちである。この例では、治療効果の推定値はそれぞれの方法と極めてよく似ているが、標準誤差は微妙に異なっている。

## 5. 変量効果モデル

表 5.2 表 5.1 のデータを回帰モデルに当てはめたコンピュータアウトプット

( i ) クラスタリングを考慮しない回帰

Source	SS	df	MS	Number of obs	=	20
Model	.881999519	1	.881999519	F(1, 18)	=	0.05
Residual	326.397937	18	18.1332187	Prob > F	=	0.8279
Total	327.279937	19	17.2252598	R-squared	=	0.0027
				Adj R-squared	=	-0.0527
				Root MSE	=	4.2583

  

bmi	Coef.	Std. Err.	t	P>  t	[95% Conf. Interval]
treat	.4199999	1.904375	0.221	0.828	-3.580943 4.420943
_cons	28.39	1.346596	21.083	0.000	25.56091 31.21909

( ii ) 最尤変量効果モデル

Fitting constant-only model :

Iteration 0: log likelihood = -51.281055

Iteration 4: log likelihood = -49.644194

Fitting full model :

Iteration 0: log likelihood = -51.269825

Iteration 4: log likelihood = -49.631613

Random-effects ML regression Number of obs = 20

Group variable (i): group Number of groups = 10

Random effects u\_i~Gaussin Obs per group: min = 1

avg = 2.0

max = 3

LR chi 2(1) = 0.03

Log likelihood=-49.631613 Prob > chi 2 = 0.8740

bmi	Coef.	Std. Err.	z	P>  z	[95% Conf. Interval]
treat	.3916501	2.467672	0.159	0.874	-4.444899 5.228199
_cons	28.39	1.740882	16.308	0.000	24.97793 31.80207
/sigma_u	3.737395	.9099706	4.107	0.000	1.953885 5.520904
/sigma_e	1.539626	.3430004	4.489	0.000	.8673579 2.211895
rho	8.54917	.0848639			.6286473 .9630508

Likelihood ratio test of sigma\_u=0: chi 2(1) = 13.34 Prob > chi 2 = 0.0003

## (iii) 一般化推定方程式

Iteration 1: tolerance = .01850836  
 Iteration 2: tolerance = .0000438  
 Iteration 3: tolerance = 1.029 e-07

GEE population-averaged model	Number of obs	=	20
Group variable : group	Number of groups	=	10
Link : identity	Obs per group: min	=	1
Family : Gaussian	avg	=	2.0
Correlation : exchangeable	max	=	3
	LR chi 2(1)	=	0.02
Scale parameter: 18.1336	Prob > chi 2	=	0.8803

bmi	Coef.	Std. Err.	z	P>  z	[95% Conf. Interval]
treat	.3937789	2.613977	0.151	0.880	-4.729523 5.51708
_cons	28.39	1.844712	15.390	0.000	24.77443 32.00557

## (iv) 群の平均値に関する回帰

Between regression (regression on group means)	Number of obs	=	20
Group variable (i) : group	Number of groups	=	10
R-sq: within =	Obs per group:	=	1
between =0.0027	min	=	2.3
overall =0.0027	avg	=	3
	max	=	0.02
sd(u <sub>i</sub> + avg(e <sub>i</sub> ))=4.34548	F(1,8)	=	0.02
	Prob > F	=	0.8860

bmi	Coef.	Std. Err.	t	P>  t	[95% Conf. Interval]
treat	.4066669	2.748323	0.148	0.886	-5.930976 6.74431
_cons	28.39	1.943358	14.609	0.000	23.90861 32.87139

## 5.8 モデルの確認

変量効果モデルに対する仮定の多くは、第2章や第3章で記載されている線形モデルやロジスティックモデルの仮定とよく似ている。主な違いは変量効果の項に関する仮定である。正式な確認方法の説明は本書の

## 5. 変量効果モデル

目的からずれているが、最尤法では変量効果項は正規分布だと仮定している。クラスターごとの症例数がほとんど同じであれば、ヒストグラムでクラスターの平均値を求めればよい。クラスターごとに症例数がひどく違っていけばむずかしい。クラスター単位のランダム化試験では、クラスターごとの患者数が治療法ごとに大きく異なることが肝心である。介入が運動療法である場合、例えば運動の効果が誘因となり、運動療法群が増えてしまい、症例数がうまく合わないことがあるかもしれない。

### 5.9 変量効果モデル解析結果を報告するには

- 患者数だけでなく、群 (groups) の数を報告すること
- クラスター単位のランダム化試験において、治療群ごとに共変量の平均値を報告すること。それにより、読者はその試験で群が均衡していたか (at a group level) どうか分かる。
- クラスター別モデルか周辺モデルのどちらを用いたか、そしてどのようにして選択したかを記述すること。
- 変量効果の分布に基づく仮定をどのように確かめたかを示すこと

### 5.10 変量効果モデル解析結果の読み方

- 解析の主要な単位は何になっているか？ 統計解析はそれを反映しているか？ 一人の個人について測定が繰り返されるということと、異なる個人に2回目の測定を行っているということと同じではないし、統計解析はそれぞれの状況に応じて異なるべきである。
- クラスター単位のランダム化試験であれば、適切なモデルが使われたか？
- gee を用いたなら、結果を判断するために十分な数の群があったか？

## [よく質問されること (FAQ)]

## 変量効果モデルはどのようなとき使えばよいのか？

変量効果モデルの使い方については議論がある。反復測定データやクラスター単位のランダム化試験は、選択肢の一つだ。しかし、他の分野では使うべきかについては、はっきりしていない。例えば表 3.3 で述べた 4 つの試験結果を併合するのはメタアナリシス (meta-analysis) の例である。仮想的なケースコントロール研究の母集団から、これらの 4 つの研究が抽出されたと見なされるので、この場合は変量効果モデルを用いてもよいだろう。

しかし、これはいつもそうというわけではない。したがって、経験のある統計家に相談し助言をしてもらったほうが良いだろう。

## ■文献

- 1 Brown H, Prescott R. *Applied Mixed Models in Medicine*. Chichester: John Wiley, 1999.
- 2 Crowder M, Hand DJ. *Analysis of Repeated Measures*. London: Chapman and Hall, 1999.
- 3 Diggle PJ, Liang K-Y, Zeger S. *Analysis of Longitudinal Data*. Oxford: Oxford Science Publications, 1994.
- 4 Goldstein H. *Multi-level Models, 2nd edn*. London: Arnold, 1996.
- 5 Cornfield J. Randomization by group: a formal analysis. *Am J Epidemiol* 1978; 108: 100-2.
- 6 Zucker DM, Lakatos E, Webber LS *et al*. Statistical Design of the Child and Adolescent Trial for Cardiovascular Health (CATCH): implications of cluster randomization. *Control Clin Trials* 1995; 16: 96-118.
- 7 Campbell MJ. Cluster randomised trials in general (family) practice. *Stat Methods Med Res* 2000; 9: 81-94.
- 8 Matthews JNS, Altman DG, Campbell MJ, Royston JP. Analysis of serial measurements in medical research. *BMJ* 1990; 300: 230-5.

## 5. 変量効果モデル

- 9 Campbell MJ, Machin D. *Medical Statistics: a commonsense approach, 3rd edn.* Chichester: John Wiley, 1999.
- 10 Kinmonth AL, Woodcock A, Griffin S, Spiegel N, Campbell MJ. Randomised controlled trial of patient centred care of diabetes in general practice: impact on current wellbeing and future disease risk. *BMJ* 1998; **317**: 1202-8.
- 11 Doull IJM, Campbell MJ, Holgate ST. Duration of growth suppressive effects of regular inhaled corticosteroids *Arch Dis Child* 1998; **78**: 172-3.
- 12 Murray DM. *Design and Analysis of Community Trials.* Oxford: Oxford University Press, 1998.
- 13 STATA Corp. STATA Statistical Software. 6.0 1999 Reference Manual. College Station, TX: STATA Corporation 1999, Section 23.11.
- 14 Turner RM, Omar RZ, Thompson SG. Bayesian methods of analysis for cluster randomised trials with binary outcome data. *Statistics in Medicine* (in press), 2001.



## 6

# その他のモデル

### 要 旨

本章では、医学研究でよく用いるその他のモデルを取り上げる。それらは、ポアソン回帰 (poisson regressin)、順序回帰 (ordinal regression)、時系列回帰 (time-series regression)、の三つである。ポアソン回帰というのは、結果変数が頻度 (count) であるときに使える。順序回帰は、結果変数が順位または順序であるときに使える。時系列回帰は、結果変数が連続変数のときに主に使われるが、経時的に取られた予測変数でないといけない。

### 6.1. ポアソン回帰モデル

ポアソン回帰はロジスティック回帰モデルを拡張したものである。すなわち個人のイベントリスクは小さいが、たくさんの人達がいるため、グループ内のイベント数は相当ある。ある個人におけるイベントの有無は直接知らなくてもよいが、人年 (person-years) といった追跡期間あたりでのイベント数は必要である。このことはリスク状態にあった (at risk) 時間の長さとしても知られている。この用法は疫学でよく使われる。特にコホート研究の解析ではよく使われる。詳しくは McNeil<sup>1</sup> を参照のこと。

#### 6.1.1 モデル

ポアソンモデルでの結果変数とは、グループ内におけるイベントの頻度である。通常はある期間あたりで測定する。例えば、石綿の曝露を

## 6. その他のモデル

受けたグループでの20年にわたる死亡数がそうである。第1章の用語でいうと、これは量的離散変数 (discrete quantitative variable) である。ここでの主要な共変量とは、そのグループがその研究にあった時間の長さである。被験者は、それぞれ異なる期間の研究下にあったリスク期間 (at risk period といわれる) かもしれない。だから被験者ごとに曝露時間  $e_i$  を示す時間を記録するのである。ロジスティック回帰ではイベント  $\pi_i$  の確率についてモデル化した。ここでは、時間  $e_i$  で区切った分割期間  $i$  上でイベント  $E_i$  の期待数、つまりそういう率  $\lambda_i$  についてモデル化する。ロジスティック変換の代わりに単純な対数変換を用いる。

このモデルは

$$\log(\lambda_i) = \log(E_i/e_i) = \beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip} \quad (6.1)$$

であり、次の様に書き換えてもよい。

$$E_i = \exp\{\log(e_i) + \beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip}\} \quad (6.2)$$

ここでの仮定は、イベントのリスクは  $e_i$  に従って上がることである。したがって、モデル (6.2) の、 $\log(e_i)$  に関する係数は1と固定する。これは基点 (offset) として知られており、その回帰係数が1という単位として固定した特別なタイプの独立変数である。

ロジスティック回帰で注意することは、観測頻度 (observed counts) がモデルには現れないことである。同じようにポアソン回帰では、観測頻度はポアソン分布の期待頻度とリンクされる ([付録2] 参照)。したがって、観測頻度  $y_i$  はパラメータ  $E_i = \lambda_{ei}$  のポアソン分布すると仮定する。

リスク状態の人年の代わりとして、死亡数の予測値を用いることができる。それは外部データを利用する。例えば、イングランドとウエールズでの年齢・性別の死亡数を用いて、それぞれのグループでの死亡数が予測できる。このようにして標準死亡比 (the standardised mortality ratio ; SMR) をモデル化することができる。詳しくは、Breslow and Day<sup>2</sup> を参照のこと。

### 6.1.2 モデルからわかること

独立変数が単純な二値、つまり0または1であるコホート研究を考えよう。ここで、0とはハザードに曝されない、であり、1は曝された、とする。また、従属変数は病死の有無であり、本研究に入っている期間も含める。このときモデルから推定される係数 $b$ は、曝露された人での病気の発生率と曝露しなかった人でのそれとの比の対数である。したがって、 $\exp(b)$ は発生率の比〔incidence rate ratio (irr)〕または相対リスク (relative risk) の推定値になる。独立変数が連続変数であれば、その回帰係数は独立変数の単位増分あたりの $y$ 変数の%変化 (percentage change) になる。

### 6.1.3 コンピュータ出力の解釈

データのレイアウトは、第3章のグループ別ロジスティック回帰 (the grouped logistic regression) と全く同じである。モデルの当てはめは最尤法で行う。

表6.1は古典的なコホート研究であり、英国の男性医師<sup>3</sup>を対象に冠

表6.1 英国の男性医師の冠動脈死

死亡数(D)	人一年	喫煙者	研究開始時点		(D-E) ÷ √E
			での年齢グループ	期待値(E)	
32	52407	1	35~44	27.2	0.92
2	18790	0	35~44	6.8	-1.85
104	43248	1	45~54	98.9	0.52
12	10673	0	45~54	17.1	-1.24
206	28612	1	55~64	205.3	0.05
28	5712	0	55~64	28.7	-0.14
186	12663	1	65~74	187.2	-0.09
28	2585	0	65~74	26.8	0.23
102	5317	1	75~84	111.5	-0.89
31	1462	0	75~84	21.5	2.05

## 6. その他のモデル

表 6.2 表 6.1 のデータに対するポアソン回帰分析の結果

Iteration 0 :	log likelihood=	-33.823578				
Iteration 1 :	log likelihood=	-33.60073				
Iteration 2 :	log likelihood=	-33.600412				
Iteration 3 :	log likelihood=	-33.600412				
Poisson regression			Number of obs	=		10
			LR chi 2(5)	=		922.93
			Prob>chi 2	=		0.0000
Log likelihood =		-33.600412	Pseudo R2	=		0.9321
y	IRR	Std. Err.	Z	P> z	[95% Conf. Interval]	
Smoker	1.425664	.1530791	3.303	0.001	1.155102	1.7596
Age 45~54	4.41056	.860515	7.606	0.000	3.008995	6.464962
Age 55~64	13.83849	2.542506	14.301	0.000	9.65384	19.83707
Age 65~74	28.51656	5.269837	18.130	0.000	19.85162	40.96364
Age 75~84 pyrs (exposure)	40.45104	7.77548	19.249	0.000	27.75314	58.95862

動脈死 (coronary death) と喫煙の関係をみたものである。Breslow and Day<sup>2</sup> と McNeil<sup>1</sup> が引用している。

ここで問題となるのは、年齢調整済の喫煙による死亡リスクはいくらかということである。ここでの従属変数は、年齢・喫煙グループごとの死亡数である。喫煙グループは原因変数、年齢グループは交絡変数、そして人年は基点変数である。解析結果は表 6.2 に示した。

5つの年齢グループは、ベースラインを 35~44 歳として、4つのダミー変数で当てはめた。このモデルでは喫煙者における冠動脈死の相対リスクは、各年齢グループで一定だと仮定している。非喫煙者と比べた喫煙者の推定相対リスクは 1.43 であり、95% 信頼区間は 1.16~1.76 となり高度に有意 (P=0.001) である。このように男性の英国人医師においては、喫煙により冠動脈死が 40% 上昇すると考えられる。出力内の LR chi 2 (尤度比  $\chi^2$  統計量) はこのモデルの全体の検定結果であり、高度に有意である。しかし、有意になったのは年齢のカテゴリーに強く依