

第Ⅴ部 Logistic Model

4 多変数化

ポイント

- ダミー変数
- 交互作用の把握

- カテゴリカル・データのダミー変数化
- 序数データのダミー変数化

4.1 はじめに

- 多変数化は変数の種類と数を増やすことです。
- 独立変数が複数あれば、独立変数間の関係も把握できます。
- 種類が増えれば考慮すべき事柄も増えます。

4.2 カテゴリカル・データ

- 分類・所属・種類の違いをあらわす値。
- 属する区分をあらわす値。
- 同じか異なるかの情報のみを持っている。
- 分類の記号として数字を使う場合があるが、この数には量としての意味はなく、大小の概念すらない。
- 名義尺度ともいう。

4.2.1 ダミー変数

- 1 または 0 しか取らない変数をダミー変数という。
- ある条件を満たしていた場合 1 とし、満たしていない場合 0 とすることで数量化を実現。
 - 習慣・経験・行為の有無
 - 特定の集団への所属
 - 特定の属性の有無
- カテゴリカル・データを数量化可能となる。
 - 性別
 - 血液型
 - 曜日

4.2.2 食事制限と体重の関係の例

x_i を食事制限の有無をあらわすダミー変数とする。

- 1: 食事制限あり
- 0: 食事制限なし

一定期間内に一定量の体重減少があった場合を『効果あり』とし $y_i = 1$ 、無かった場合『効果なし』とし $y_i = 0$ とする。

表 1 食事制限と体重減少

x	$n_{y=1}$	$n_{y=0}$
1	13	7
0	9	11

$$p_i = \frac{1}{1 + \exp(-(\alpha + \beta x_i))} \quad (1)$$

$$1 - p_i = \frac{1}{1 + \exp(\alpha + \beta x_i)} \quad (2)$$

$x_i = 1$ のときの p_i

$$p_i = \frac{1}{1 + \exp(-(\alpha + \beta(1)))} \quad (3)$$

$$= \frac{1}{1 + \exp(-(\alpha + \beta))} \quad (4)$$

$x_i = 0$ のときの p_i

$$p_i = \frac{1}{1 + \exp(-(\alpha + \beta(0)))} \quad (5)$$

$$= \frac{1}{1 + \exp(-(\alpha))} \quad (6)$$

$x_i = 1$ のときの $1 - p_i$

$$1 - p_i = \frac{1}{1 + \exp(\alpha + \beta(1))} \quad (7)$$

$$= \frac{1}{1 + \exp(\alpha + \beta)} \quad (8)$$

$x_i = 0$ のときの $1 - p_i$

$$1 - p_i = \frac{1}{1 + \exp(\alpha + \beta(0))} \quad (9)$$

$$= \frac{1}{1 + \exp(\alpha)} \quad (10)$$

尤度関数と対数尤度関数

尤度関数は、

$$L = \left(\frac{1}{1 + \exp(-(\alpha + \beta))} \right)^{13} \left(\frac{1}{1 + \exp(\alpha + \beta)} \right)^7 \left(\frac{1}{1 + \exp(-(\alpha))} \right)^9 \left(\frac{1}{1 + \exp(\alpha)} \right)^{11} \quad (11)$$

である。尤度関数を対数変換すると、対数尤度関数

$$\begin{aligned} \ln L = 13 \ln \left(\frac{1}{1 + \exp(-(\alpha + \beta))} \right) + 7 \ln \left(\frac{1}{1 + \exp(\alpha + \beta)} \right) \\ + 9 \ln \left(\frac{1}{1 + \exp(-(\alpha))} \right) + 11 \ln \left(\frac{1}{1 + \exp(\alpha)} \right) \end{aligned} \quad (12)$$

を得る。

最尤推定値とオッズ比

表 2 最尤推定値とオッズ比

α	-0.200 671
β	0.819 710
$\ln L$	-26.711 709
$\exp(\beta)$	2.269 842

「食事制限をする」と「しない」場合に比べ、オッズは 2.27 倍になる。

問題 V-4-1

x_i を運動の有無をあらわすダミー変数とする。

- 1: 運動あり
- 0: 運動なし

x	$n_{y=1}$	$n_{y=0}$
1	12	8
0	10	10

一定期間内に一定量の体重減少があった場合を『効果あり』とし $y_i = 1$ 、無かった場合『効果なし』とし $y_i = 0$ とする。表の値を用いてロジスティックモデルを当てはめなさい。

パラメータ推定に関しては何らかのアプリを使うものとする。

解例 V-4-1

α	0.000 000
β	0.405 465
$\ln L$	-27.323 177
$\exp(\beta)$	1.500000

「運動をする」と「しない」場合に比べ、オッズは 1.50 倍になる。

4.3 独立変数の多変数化

$i (1, 2, \dots, n)$ をデータの組を識別する添え字とし、 $j (1, 2, \dots, m)$ を独立変数を識別する添え字とする。

y_i をダミー変数、 x_{ij} を連続量としデータの組を $(y_i, x_{i1}, x_{i2}, \dots, x_{im})$ とする。

$y_i = 1$ の時の確率を p_i とする。

p_i を従属変数、 $x_{i1}, x_{i2}, \dots, x_{im}$ を独立変数、 $\beta_0, \beta_1, \beta_2, \dots, \beta_m$ をパラメータとし、多変数ロジスティックモデルを以下のように定義する。

$$p_i = \frac{1}{1 + \exp(-(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_m x_{im}))} \quad (13)$$

問題 V-4-2

$$p_i = \frac{1}{1 + \exp(-(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}))}$$

とする。 $1 - p_i$ を求めなさい。

解例 V-4-2

$$\begin{aligned}1 - p_i &= 1 - \frac{1}{1 + \exp(-(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}))} \\&= 1 - \frac{1}{1 + \frac{1}{\exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2})}} \\&= 1 - \frac{1}{\left(\frac{1 + \exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2})}{\exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2})} \right)} \\&= 1 - \frac{\exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2})}{1 + \exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2})} \\&= \frac{1}{1 + \exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2})}\end{aligned}$$

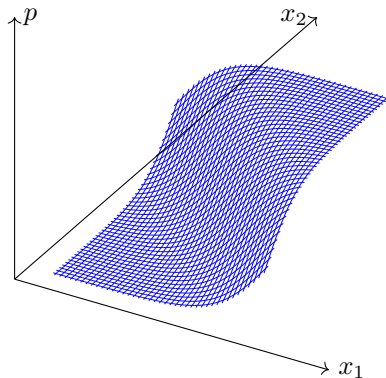
4.3.1 2変数ロジスティックモデル

2変量のロジスティックモデルを次のように定義する。

$$p_i = \frac{1}{1 + \exp(-(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}))} \quad (14)$$

$$1 - p_i = \frac{1}{1 + \exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2})} \quad (15)$$

図1 2変数ロジスティックモデルの形状



4.3.2 x_2 の違いの意味

$x_{i2} = c_1, c_2$ のとき、それぞれのロジスティックモデルを p_{c_1}, p_{c_2} とあらわすと

$$p_{c_1} = \frac{1}{1 + \exp(-(\beta_0 + \beta_1 x_{i1} + \beta_2 c_1))} \quad (16)$$

$$p_{c_2} = \frac{1}{1 + \exp(-(\beta_0 + \beta_1 x_{i1} + \beta_2 c_2))} \quad (17)$$

である。 c_1, c_2 は定数なので、その差は、分母第二項冪数内の第三項の差であり、この差は x_1 軸方向の平行移動である。

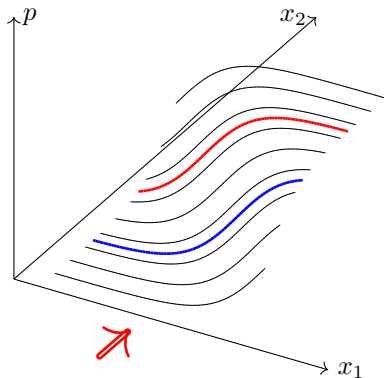
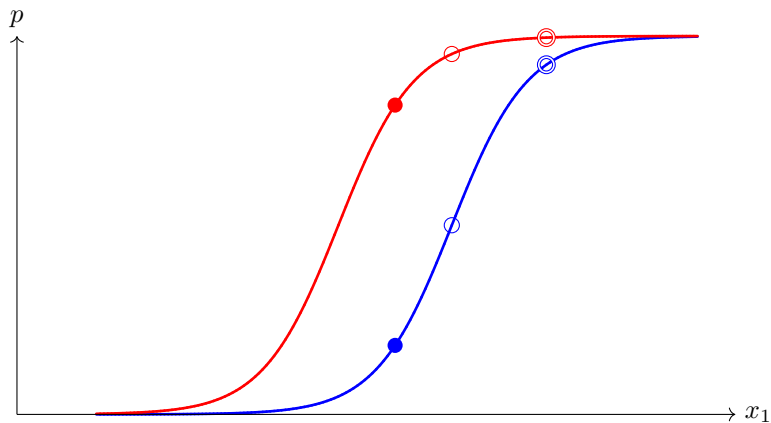
図2 x_2 ごとの p 

図3 $x_1 \times p$ 平面から見た x_2 の差

4.3.3 $x_1 \times p$ 面上から見た x_2

- p_{c_1}, p_{c_2} において β_1 は共通なので、同じ曲線を描いている。そしてロジスティックモデルは狭義単調関数である。
- x_1 軸上における x_2 の違いは、ロジスティックモデルの x_1 軸方向の平行移動として観測される。
- このときの $x_1 \times p$ 面上から見た p_{c_1}, p_{c_2} 上の点をそれぞれ $p_{c_1}(x_{i1})$, $p_{c_2}(x_{i1})$ とあらわす。
- x_1 軸上の任意の点 a において、 $p_{c_1}(a) > p_{c_2}(a)$ ならば、 x_1 軸上のいかなる点 b においても $p_{c_1}(b) > p_{c_2}(b)$ の関係が成り立つ。
- このことは、二つのロジスティックモデル p_{c_1} , p_{c_2} は交わらないことを意味する。

4.3.4 多重共線性

- 同一の独立変数を複数採用した場合。
- 独立変数間に線型変換の関係が有る場合。
- 独立変数間に強い相関がある場合。

これらの場合、パラメータの推定値は不可能または不安定になる。

ロジスティックモデルにおいても多重共線性は当然発生する。

4.3.5 ダミー変数による2変数化

 x_{i1} を食事制限ダミー

- 1: あり
- 0: なし

 x_{i2} を運動ダミー

- 1: あり
- 0: なし

減量の効果があった場合 $y_i = 1$ 、無かった場合 $y_i = 0$

表3 食事制限と運動と減量効果

組合せ	$n_{y=1}$	$n_{y=0}$
$x_1 = 1, x_2 = 1$	9	1
$x_1 = 1, x_2 = 0$	4	6
$x_1 = 0, x_2 = 1$	3	7
$x_1 = 0, x_2 = 0$	6	4

$x_{i1} = 1, x_{i2} = 1$ のとき

$$p_i = \frac{1}{1 + \exp(-(\beta_0 + \beta_1 (1) + \beta_2 (1)))} \quad (18)$$

$$= \frac{1}{1 + \exp(-(\beta_0 + \beta_1 + \beta_2))} \quad (19)$$

 $x_{i1} = 1, x_{i2} = 0$ のとき

$$p_i = \frac{1}{1 + \exp(-(\beta_0 + \beta_1 (1) + \beta_2 (0)))} \quad (20)$$

$$= \frac{1}{1 + \exp(-(\beta_0 + \beta_1))} \quad (21)$$

 $x_{i1} = 0, x_{i2} = 1$ のとき

$$p_i = \frac{1}{1 + \exp(-(\beta_0 + \beta_1 (0) + \beta_2 (1)))} \quad (22)$$

$$= \frac{1}{1 + \exp(-(\beta_0 + \beta_2))} \quad (23)$$

 $x_{i1} = 0, x_{i2} = 0$ のとき

$$p_i = \frac{1}{1 + \exp(-(\beta_0 + \beta_1 (0) + \beta_2 (0)))} \quad (24)$$

$$= \frac{1}{1 + \exp(-(\beta_0))} \quad (25)$$

(14) 式 (15) 式による尤度 l_i は

$$l_i = \left(\frac{1}{1 + \exp(-(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}))} \right)^{y_i} \left(\frac{1}{1 + \exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2})} \right)^{1-y_i} \quad (26)$$

したがって対数尤度関数 $\ln L$ は

$$\ln L = \sum_{i=1}^n \left(y_i \ln \frac{1}{1 + \exp(-(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}))} + (1 - y_i) \ln \frac{1}{1 + \exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2})} \right) \quad (27)$$

である。

問題 V-4-3

表の値を用いてロジスティックモデルを当てはめ、最尤推定値とオッズを求めなさい。

$y = 1$ となる確率を p_i とあらわし従属変数とする。 x_{ij} を独立変数、 β_j をパラメータとする。当てはめるモデルは

$$p_i = \frac{1}{1 + \exp(-(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}))}$$
$$1 - p_i = \frac{1}{1 + \exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2})}$$

とする。パラメータ推定に関しては何らかのアプリを使うものとする。

組合せ	$n_{y=1}$	$n_{y=0}$
$x_1 = 1, x_2 = 1$	9	1
$x_1 = 1, x_2 = 0$	4	6
$x_1 = 0, x_2 = 1$	3	7
$x_1 = 0, x_2 = 0$	6	4

解例 V-4-3

β_0	-0.414 339
β_1	0.828 676
β_2	0.422 849
$\ln L$	-26.500 735
$\exp(\beta_1)$	2.290 285
$\exp(\beta_2)$	1.526 304

食事制限するとオッズは 2.29 倍になり、
運動するとオッズは 1.53 倍になる。

4.4 交互作用

- 『交互作用』とは特定の組合せの時にのみ観測される特別な効果をいう。
- 『交互作用』がある場合は、個々の独立変数による効果を単純に比較する意味はない。
- 独立変数単独の効果と組合せの効果を分離して比較する必要がある。
- 独立変数単独の効果を『主効果』とよび組合せの効果を『交互作用効果』と呼ぶ。

4.4.1 x_{i1} と x_{i2} の交互作用効果

- x_{i1} と x_{i2} の組み合わせによる y_i の頻度を確認した。
- $x_1 = 1$ のときは $x_2 = 1$ は $x_2 = 0$ より多いが、 $x_1 = 0$ のときでは $x_2 = 1$ は $x_2 = 0$ より少ない。
- このような場合、交互作用が疑われる。

表 4 組合せによる $y = 1$ の頻度

	$x_2 = 1$	$x_2 = 0$
$x_1 = 1$	9	4
$x_1 = 0$	3	6

図4 観測値の分布

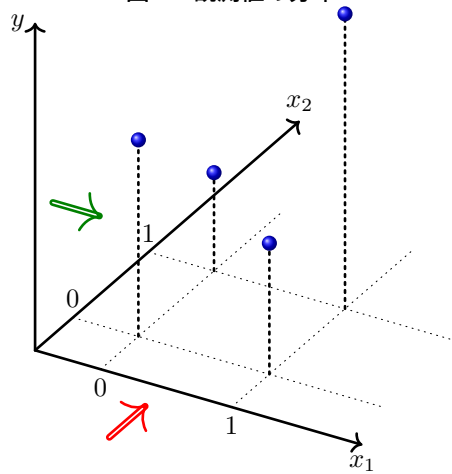
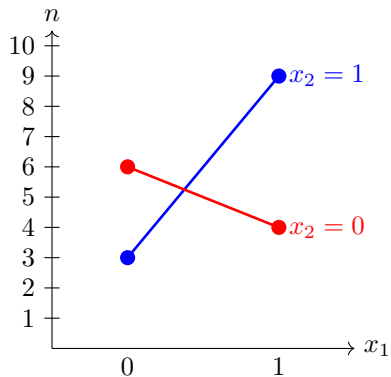
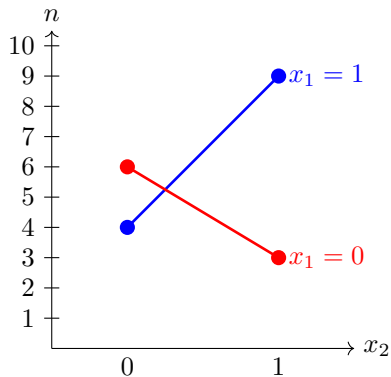


図5 x_1 軸から見た x_2 の違い図6 x_2 軸から見た x_1 の違い

4.4.2 交互作用効果を想定した分析

交互作用効果を把握する為に交互作用ダミーを導入する。

- x_{i11}
- 交互作用ダミー
- $x_{i1} = 1$ かつ $x_{i2} = 1 \implies x_{i11} = 1$
- $x_{i1} = 0$ または $x_{i2} = 0 \implies x_{i11} = 0$

4.4.3 交互作用項を含む p_i

$$p_i = \frac{1}{1 + \exp(-(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_{11} x_{i11}))} \quad (28)$$

$$1 - p_i = \frac{1}{1 + \exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_{11} x_{i11})} \quad (29)$$

$$x_{i1} = 1, x_{i2} = 1 \implies x_{i11} = 1$$

$$p_i = \frac{1}{1 + \exp(-(\beta_0 + \beta_1(1) + \beta_2(1) + \beta_{11}(1)))} \quad (30)$$

$$= \frac{1}{1 + \exp(-(\beta_0 + \beta_1 + \beta_2 + \beta_{11}))} \quad (31)$$

$$x_{i1} = 1, x_{i2} = 0 \implies x_{i11} = 0$$

$$p_i = \frac{1}{1 + \exp(-(\beta_0 + \beta_1(1) + \beta_2(0) + \beta_{11}(0)))} \quad (32)$$

$$= \frac{1}{1 + \exp(-(\beta_0 + \beta_1))} \quad (33)$$

$$x_{i1} = 0, x_{i2} = 1 \implies x_{i11} = 0$$

$$p_i = \frac{1}{1 + \exp(-(\beta_0 + \beta_1(0) + \beta_2(1) + \beta_{11}(0)))} \quad (34)$$

$$= \frac{1}{1 + \exp(-(\beta_0 + \beta_2))} \quad (35)$$

$$x_{i1} = 0, x_{i2} = 0 \implies x_{i11} = 0$$

$$p_i = \frac{1}{1 + \exp(-(\beta_0 + \beta_1(0) + \beta_2(0) + \beta_{11}(0)))} \quad (36)$$

$$= \frac{1}{1 + \exp(-(\beta_0))} \quad (37)$$

4.4.4 交互作用項のオッズ比

交互作用項 (x_{i11}) は単独では 1 にはならない。交互作用項 (x_{i11}) が 1 になる時は、 $x_{i1} = 1$ かつ $x_{i2} = 1$ である。 $x_{i1} = 1, x_{i2} = 1, x_{i11} = 1$ のときの $P_{y=1}$ を p_{11} とあらわし、 p_{11} のときのオッズを $odds_{11}$ とあらわす。この表記を用いて交互作用項 (x_{i11}) が 1 になる時のオッズをあらわすと

$$odds_{11} = \left(\frac{p_{11}}{1 - p_{11}} \right) = \exp(\beta_0 + \beta_1 + \beta_2 + \beta_{11}) \quad (38)$$

そしてそれ以外の組合せの時のオッズは

$$odds_{10} = \left(\frac{p_{10}}{1 - p_{10}} \right) = \exp(\beta_0 + \beta_1) \quad (39)$$

$$odds_{01} = \left(\frac{p_{01}}{1 - p_{01}} \right) = \exp(\beta_0 + \beta_2) \quad (40)$$

$$odds_{00} = \left(\frac{p_{00}}{1 - p_{00}} \right) = \exp(\beta_0) \quad (41)$$

である。

独立変数の影響力をあらわすオッズ比において分母に (41) を採用すると

$$\frac{odds_{10}}{odds_{00}} = \frac{\exp(\beta_0 + \beta_1)}{\exp(\beta_0)} = \exp(\beta_1) \quad (42)$$

$$\frac{odds_{01}}{odds_{00}} = \frac{\exp(\beta_0 + \beta_2)}{\exp(\beta_0)} = \exp(\beta_2) \quad (43)$$

$$\frac{odds_{11}}{odds_{00}} = \frac{\exp(\beta_0 + \beta_1 + \beta_2 + \beta_{11})}{\exp(\beta_0)} = \exp(\beta_1 + \beta_2 + \beta_{11}) \quad (44)$$

である。

問題 V-4-4

表 3 を用いてロジスティックモデルを当てはめ、最尤推定値とオッズを求めなさい。
当てはめるモデルは

$$p_i = \frac{1}{1 + \exp(-(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_{11} x_{i11}))}$$

とする。ここで x_{i11} は

$$x_{i1} = 1 \text{ かつ } x_{i2} = 1 \implies x_{i11} = 1$$

$$x_{i1} = 0 \text{ または } x_{i2} = 0 \implies x_{i11} = 0$$

となる交互作用ダミーとする。パラメータ推定に関しては何らかのアプリを使うものとする。

解例 V-4-4

β_0	0.405 465
β_1	-0.810 931
β_2	-1.252 763
β_{11}	3.855 452
$\ln L$	-22.819 706
$\exp(\beta_1)$	0.444 444
$\exp(\beta_2)$	0.285 714
$\exp(\beta_1 + \beta_2 + \beta_{11})$	5.999 989

「食事制限する」とオッズは 0.44 倍になり、「運動する」とオッズは 0.29 倍になる。そして「食事制限」と「運動」を同時に行うとオッズは 6.00 倍になる。

4.5 序数データ

- 並ぶものの中での順序を表す値。
- 順番の中の位置をあらわす。
- 広義では、程度の違いをあらわした区分からなる値を指す。
- 区分の大小・長短・高低・広狭・多少・強弱・上下の情報はある。
- 区分間の差の情報はない。
- 区分の差がどの程度かは分からない。
- 順序尺度ともいう。

4.5.1 従属変数が序数データの場合

従属変数 (y) が序数データ (分類数:3) の場合

1. 破損
2. 変形
3. 正常

加熱時間 (x) を変えながら 100 個ずつ作成し、破損・変形・正常を検査した。

表 5 加熱時間による影響

時間	破損	変形	正常
10	1	0	99
11	3	2	95
12	12	15	73
13	38	35	27
14	73	22	5
15	92	7	1
16	98	2	0

4.5.2 序数データのダミー変数化

序数データを上位から、任意の順序までのデータとそれ以降のデータに分けることで、序数データはダミー変数に変換することができる。今回の例であれば、破損・変形・正常の序列であることは明らかである。これを、破損しているか否か、破損・変形しているか否かでそれぞれダミー変数に置き換える。つまり従属変数を2つのダミー変数 (y_1, y_2) で表現すると、

y_1 : 破損なら 1、変形・正常なら 0

y_2 : 破損・変形なら 1、正常なら 0

ここでは従属変数が序数データである場合を例としているが、独立変数が序数データである場合も同様の処理により、ダミー変数化することが可能である。

表 6 y_1 によるデータ

x	$n_{y_1=1}$	$n_{y_1=0}$
10	1	99
11	3	97
12	12	88
13	38	62
14	73	27
15	92	8
16	98	2

表 7 y_2 によるデータ

x	$n_{y_2=1}$	$n_{y_2=0}$
10	1	99
11	5	95
12	27	73
13	73	27
14	95	5
15	99	1
16	100	0

4.5.3 累積ロジスティックモデル

y_1, y_2 を従属変数とする二つのロジスティックモデルを累積ロジスティックモデルという。累積ロジスティックモデルではそれぞれのモデルを個別に推定する。

$$p_{(y_{i1}=1)} = \frac{1}{1 + \exp(-(\alpha_1 + \beta_1 x_i))} \quad (45)$$

$$p_{(y_{i2}=1)} = \frac{1}{1 + \exp(-(\alpha_2 + \beta_2 x_i))} \quad (46)$$

累積ロジスティックモデルはそれぞれのモデルを個別に推定するため通常 $\beta_1 \neq \beta_2$ である。

問題 V-4-5

1. 破損
2. 変形
3. 正常

とする。加熱時間 (x) を変えながら 100 個ずつ作成し、破損・変形・正常を検査した結果が表である。表の値を用いて累積ロジスティックモデルを当てはめ、最尤推定値とオッズ比を求めなさい。パラメータ推定に関しては何らかのアプリを使うものとする。

x	破損	変形	正常
10	1	0	99
11	3	2	95
12	12	15	73
13	38	35	27
14	73	22	5
15	92	7	1
16	98	2	0

解例 V-4-5

$$p_{(y_{i1}=1)} = \frac{1}{1 + \exp(-(\alpha_1 + \beta_1 x_i))}$$

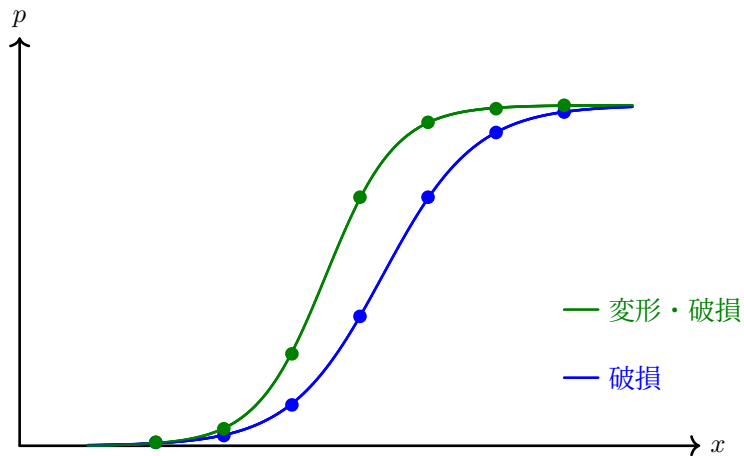
$$p_{(y_{i2}=1)} = \frac{1}{1 + \exp(-(\alpha_2 + \beta_2 x_i))}$$

とする。

加熱時間 1 単位ごとに、「変形・破損する」オッズは 4.34 倍になり「破損する」オッズは 7.00 倍になる。

α_1	-19.579 782
β_1	1.468 446
$\ln L$	-218.233 483
$\exp(\beta_1)$	4.342 480
α_2	-24.324 450
β_2	1.946 129
$\ln L$	-167.743 588
$\exp(\beta_2)$	7.001 533

図7 累積ロジスティックモデルの予測値



4.5.4 順序ロジスティックモデル

- 累積ロジスティックモデルにおいて、 $\beta_1 = \beta_2 = \beta$ と仮定するモデルを順序ロジスティックモデルという。
- 順序ロジスティックモデルでは二つのモデルを同時に推定する。
- 順序ロジスティックにおける (47) と (48) の違いは α_1 の有無である。

$$p(y_{i1}=1) = \frac{1}{1 + \exp(-(\alpha + \beta x_i))} \quad (47)$$

$$p(y_{i2}=1) = \frac{1}{1 + \exp(-(\alpha + \alpha_1 + \beta x_i))} \quad (48)$$

問題 V-4-6

1. 破損
2. 変形
3. 正常

とする。加熱時間 (x) を変えながら 100 個ずつ作成し、破損・変形・正常を検査した結果が表である。表の値を用いて順序ロジスティックモデルを当てはめ、最尤推定値とオッズ比を求めなさい。パラメータ推定に関しては何らかのアプリを使うものとする。

x	破損	変形	正常
10	1	0	99
11	3	2	95
12	12	15	73
13	38	35	27
14	73	22	5
15	92	7	1
16	98	2	0

解例 V-4-6

$$p_{(y_{i1}=1)} = \frac{1}{1 + \exp(-(\alpha + \beta x_i))}$$

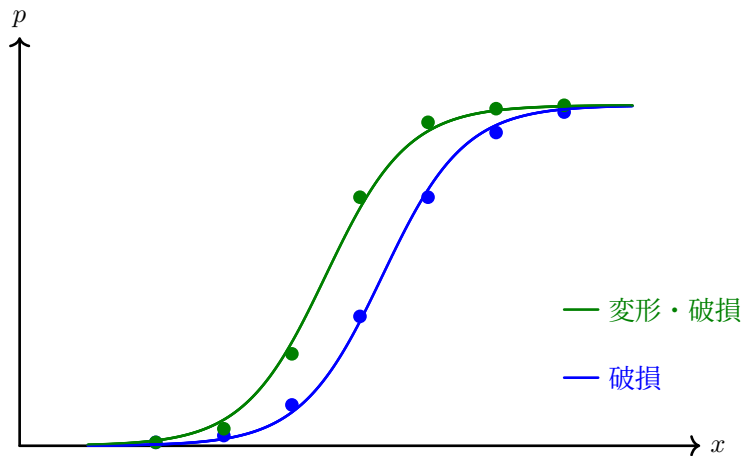
$$p_{(y_{i2}=1)} = \frac{1}{1 + \exp(-(\alpha + \alpha_1 + \beta x_i))}$$

とする。

加熱時間 1 単位ごとに、変形・破損する
オッズは 5.28 倍になる。

α	-22.172 967
α_1	1.388 330
β	1.663 165
$\ln L$	-389.470 984
$\exp(\beta)$	5.275 982

図8 順序ロジスティックモデルの予測値



4.6 多項ロジスティックモデル (参考)

カテゴリカル・データが従属変数である場合を紹介する。

i をデータの組を識別する添え字とする。分類の総数を K とする。データの組 i が k 番目の分類に属する確率を p_{ik} とする。ここで p_{iK} を規準として $\frac{p_{ik}}{p_{iK}}$ を求める。この値を真数とする自然対数を従属変数とし、 x_i を独立変数とする線型モデルとして多項ロジスティックモデルを定義する。

$$\ln \left(\frac{p_{i1}}{p_{iK}} \right) = \alpha_1 + \beta_1 x_i \quad (49)$$

$$\ln \left(\frac{p_{i2}}{p_{iK}} \right) = \alpha_2 + \beta_2 x_i \quad (50)$$

$$\vdots$$

$$\ln \left(\frac{p_{ik}}{p_{iK}} \right) = \alpha_k + \beta_k x_i \quad (51)$$

$$\vdots$$

$$\ln \left(\frac{p_{iK-1}}{p_{iK}} \right) = \alpha_{K-1} + \beta_{K-1} x_i \quad (52)$$

4.6.1 影響力の指標

x_i のときの分類 k におけるモデル式を指数関数であらわす。

$$\left. \frac{p_{ik}}{p_{iK}} \right|_{x_i} = \exp(\alpha_k + \beta_k x_i) \quad (53)$$

ここで x_i が 1 単位増えた式を考える。

$$\left. \frac{p_{ik}}{p_{iK}} \right|_{x_i+1} = \exp(\alpha_k + \beta_k (x_i + 1)) \quad (54)$$

$$= \exp(\alpha_k + \beta_k x_i + \beta_k) \quad (55)$$

そして (55) を (53) で割ると

$$\frac{\exp(\alpha_k + \beta_k x_i + \beta_k)}{\exp(\alpha_k + \beta_k x_i)} = \exp(\beta_k) \quad (56)$$

$\exp(\beta_k)$ は x_i が 1 単位増えたときの分類 K に属する確率 (p_{iK}) と分類 k に属する確率 (p_{ik}) の比の倍率であり、分類 K と分類 k との間の影響力の指標として利用される。

この値を『分類 K に対する分類 k のオッズ比』とよぶ。

4.6.2 $K = 3$ の多項ロジスティックモデル

$K = 3$ として多項ロジスティックモデルを説明する。分類先をあらわす変数として y を採用する。 y は 1, 2, 3 のどれかを取る値である。 $y = k$ のときの確率を p_{ik} とあらわす。 $\Omega = \{y|1, 2, 3\}$ とすると、 p_{ik} は確率の公理を満たすので

$$1 = p_{i1} + p_{i2} + p_{i3} \quad (57)$$

である。

この条件のもとで多項ロジスティックモデルを記述すると

$$\begin{cases} \ln \left(\frac{p_{i1}}{p_{i3}} \right) = \alpha_1 + \beta_1 x_i & (58) \\ \ln \left(\frac{p_{i2}}{p_{i3}} \right) = \alpha_2 + \beta_2 x_i & (59) \\ 1 = p_{i1} + p_{i2} + p_{i3} & (60) \end{cases}$$

である。

これを p_{ik} に関して解くことによって、分類 k に属する確率 p_{ik} を求めることができる。

これを p_{ik} に関して解くと

$$p_{i1} = \frac{1}{1 + (1 + \exp(\alpha_2 + \beta_2 x_i)) \exp(-(\alpha_1 + \beta_1 x_i))} \quad (61)$$

$$p_{i2} = \frac{1}{1 + (1 + \exp(\alpha_1 + \beta_1 x_i)) \exp(-(\alpha_2 + \beta_2 x_i))} \quad (62)$$

$$p_{i3} = 1 - p_{i1} - p_{i2} \quad (63)$$

である。

4.6.3 尤度と尤度関数と対数尤度関数

- y_{i1} を $y = 1$ のとき 1, それ以外のとき 0 をとるダミー変数とする。
- y_{i2} を $y = 2$ のとき 1, それ以外のとき 0 をとるダミー変数とする。
- そして尤度 l_i , 尤度関数 L , 対数尤度関数 $\ln L$ を以下のように定義する。

$$l_i = p_{i1}^{y_{i1}} p_{i2}^{y_{i2}} (1 - p_{i1} - p_{i2})^{1-y_{i1}-y_{i2}} \quad (64)$$

$$L = \prod_{i=1}^n p_{i1}^{y_{i1}} p_{i2}^{y_{i2}} (1 - p_{i1} - p_{i2})^{1-y_{i1}-y_{i2}} \quad (65)$$

$$\ln L = \sum_{i=1}^n (y_{i1} \ln p_{i1} + y_{i2} \ln p_{i2} + (1 - y_{i1} - y_{i2}) \ln (1 - p_{i1} - p_{i2})) \quad (66)$$

4.6.4 数値例

従属変数 (y) がカテゴリカル・データの場合

1. 和食
2. 中華
3. 洋食

各年齢から 100 名のサンプルを選び、選択肢を和食、中華、洋食 とし、一番好きな料理を調査した。

表 8 年齢と好む料理

年齢	和食	中華	洋食
20	1	8	91
25	4	17	79
30	9	34	57
35	21	46	33
40	43	42	15
45	67	27	6
50	85	13	2
55	94	5	1

4.6.5 推定値

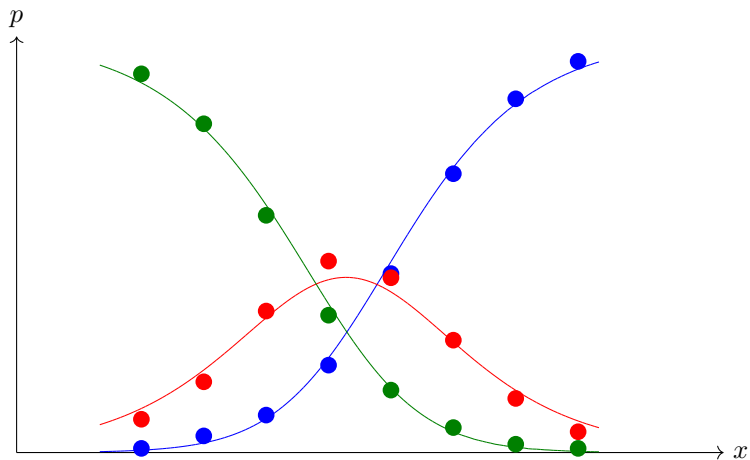
$$\ln \left(\frac{p_{i1}}{p_{i3}} \right) = \alpha_1 + \beta_1 x_i$$

$$\ln \left(\frac{p_{i2}}{p_{i3}} \right) = \alpha_2 + \beta_2 x_i$$

$$1 = p_{i1} + p_{i2} + p_{i3}$$

α_1	-11.191 683
α_2	-5.164 678
β_1	0.306 891
β_2	0.151 891
$\ln L$	-547.721 419
$\exp(\beta_1)$	1.359 192
$\exp(\beta_2)$	1.164 033
$\exp(\beta_1 - \beta_2)$	1.167 658

図9 選択肢が3つのときの好み



4.6.6 分類 1 と分類 2 の間のオッズ比

(58) を (59) を指数関数であらわすと

$$\frac{p_{i1}}{p_{i3}} = \exp(\alpha_1 + \beta_1 x_i) \quad (67)$$

$$\frac{p_{i2}}{p_{i3}} = \exp(\alpha_2 + \beta_2 x_i) \quad (68)$$

(67) を (68) で割ると、

$$\frac{\left(\frac{p_{i1}}{p_{i3}}\right)}{\left(\frac{p_{i2}}{p_{i3}}\right)} = \frac{\exp(\alpha_1 + \beta_1 x_i)}{\exp(\alpha_2 + \beta_2 x_i)} \quad (69)$$

整理し、 x_i のときの値としてあらわすと

$$\frac{p_{i1}}{p_{i2}} \Big|_{x_i} = \exp(\alpha_1 - \alpha_2 + (\beta_1 - \beta_2)x_i) \quad (70)$$

ここで x_i が 1 単位増加した時を考えると

$$\frac{p_{i1}}{p_{i2}} \Big|_{x_i+1} = \exp(\alpha_1 - \alpha_2 + (\beta_1 - \beta_2)(x_i + 1)) \quad (71)$$

である。(71) を (70) で割ると『分類 2 に対する分類 1 のオッズ』を得ることができる。

4.6.7 p_1 と p_2 の関係

(71) を (70) で割ると

$$\frac{\left(\frac{p_{i1}}{p_{i3}}\right)\Big|_{x_i+1}}{\left(\frac{p_{i2}}{p_{i3}}\right)\Big|_{x_i}} = \frac{\frac{p_{i1}}{p_{i2}}\Big|_{x_i+1}}{\frac{p_{i1}}{p_{i2}}\Big|_{x_i}} = \frac{\exp(\alpha_1 - \alpha_2 + (\beta_1 - \beta_2)(x_i + 1))}{\exp(\alpha_1 - \alpha_2 + (\beta_1 - \beta_2)x_i)} = \exp(\beta_1 - \beta_2) \quad (72)$$

である。この値は $\Omega = \{y|1, 2, 3\}$ における、 $\frac{p_{i1}}{p_{i3}}$ と $\frac{p_{i2}}{p_{i3}}$ の比において、 x_i が 1 単位変化したときの影響力の指標である。

同時に、分類 2 に対する分類 1 のオッズ比である。

(70) において

$$\alpha = \alpha_1 - \alpha_2 \quad (73)$$

$$\beta = \beta_1 - \beta_2 \quad (74)$$

とおくと

$$\frac{p_{i1}}{p_{i2}} = \exp(\alpha + \beta x_i) \quad (75)$$

である。

$\Omega = \{y|1, 2\}$ とすると、確率の公理より

$$p_{i2} + p_{i1} = 1 \quad (76)$$

$$p_{i2} = 1 - p_{i1} \quad (77)$$

なので (75) は

$$\frac{p_{i1}}{1 - p_{i1}} = \exp(\alpha + \beta x_i) \quad (78)$$

である。この式を p_{i1} に関して解くと 2 項ロジスティックモデルを得る。

$$p_{i1} = \frac{1}{1 + \exp(-(\alpha + \beta x_i))} \quad (79)$$

そしてこの時のオッズ比は

$$\exp(\beta) = \exp(\beta_1 - \beta_2) \quad (80)$$

である。この値は (72) に等しく、 $\Omega = \{y|1, 2, 3\}$ における『分類 2 に対する分類 1 のオッズ比』は $\Omega = \{y|1, 2\}$ とした 2 項ロジスティックモデルにおける『オッズ比』と等しい。

このことは分類 3 の追加によって、分類 3 が選ばれる場合を除き、分類 1 と分類 2 が選ばれる確率の関係は影響を受けないことを示している。この性質を「無関係な選択対象からの独立性」という。

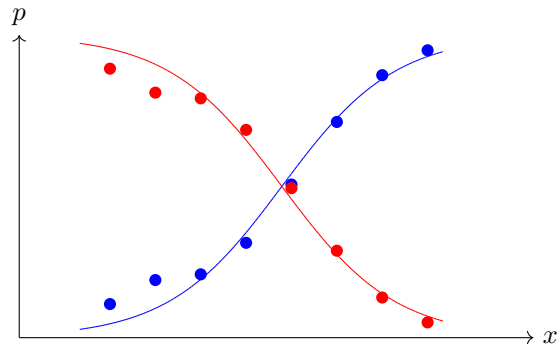
表 9 選択肢が 2 個の時の p_{ik}

x	$n_{y=1}$	$n_{y=2}$	計	$p_{y=1}$	$p_{y=2}$
20	1	8	9	0.111 111	0.888 889
25	4	17	21	0.190 476	0.809 524
30	9	34	43	0.209 302	0.790 698
35	21	46	67	0.313 433	0.686 567
40	43	42	85	0.505 882	0.494 118
45	67	27	94	0.712 766	0.287 234
50	85	13	98	0.867 347	0.132 653
55	94	5	99	0.949 495	0.050 505

表 10 選択肢が 2 個の時の推定値

α	-6.236 518
β	0.160 156
$\ln L$	-252.569 394
$\exp(\beta)$	1.173 693 747

図 10 選択肢が 2 つのときの好み



4.7 まとめ

- ロジスティックモデルはダミー変数だけを独立変数とすることができる。
- ロジスティックモデルを用いると交互作用の有無を分析可能。
- 序数データを従属変数とする場合。
 - 階層ロジスティックモデル。
 - 順序ロジスティックモデル。
- カテゴリカル・データを従属変数とする場合、多項ロジスティックモデル。
- 多項ロジスティックモデルは、無関係な選択対象からの独立性を持つ。