

ロジスティック回帰分析（目的変数は任意の確率値）

1. 目的

目的変数の数量（0～1の確率値）を推測するものです。
モデル式は

$$P = \frac{1}{1 + \exp(- (a_0 + a_1x_1 + a_2x_2 + \cdots + a_mx_m))}$$

であり、説明変数は x_1, x_2, \dots です。

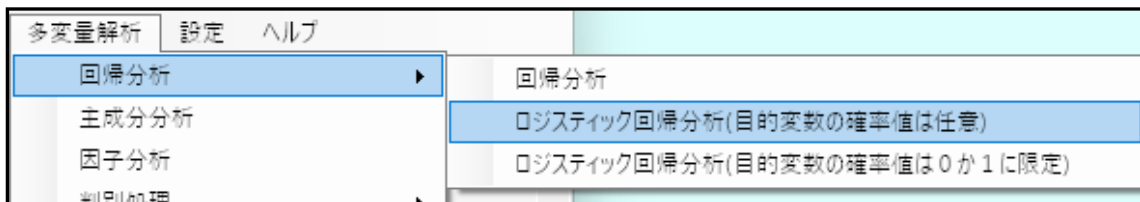
計算の過程で、目的変数の確率値が 0 または 1 の場合、計算は困難なのですが、あえてここでは数値的な工夫をして、解を求めますが、ほとんどすべての確率値が 0 または 1 の場合には、この機能でなく、

ロジスティック回帰分析（目的変数の確率値は0か1に限定）
を利用してください。

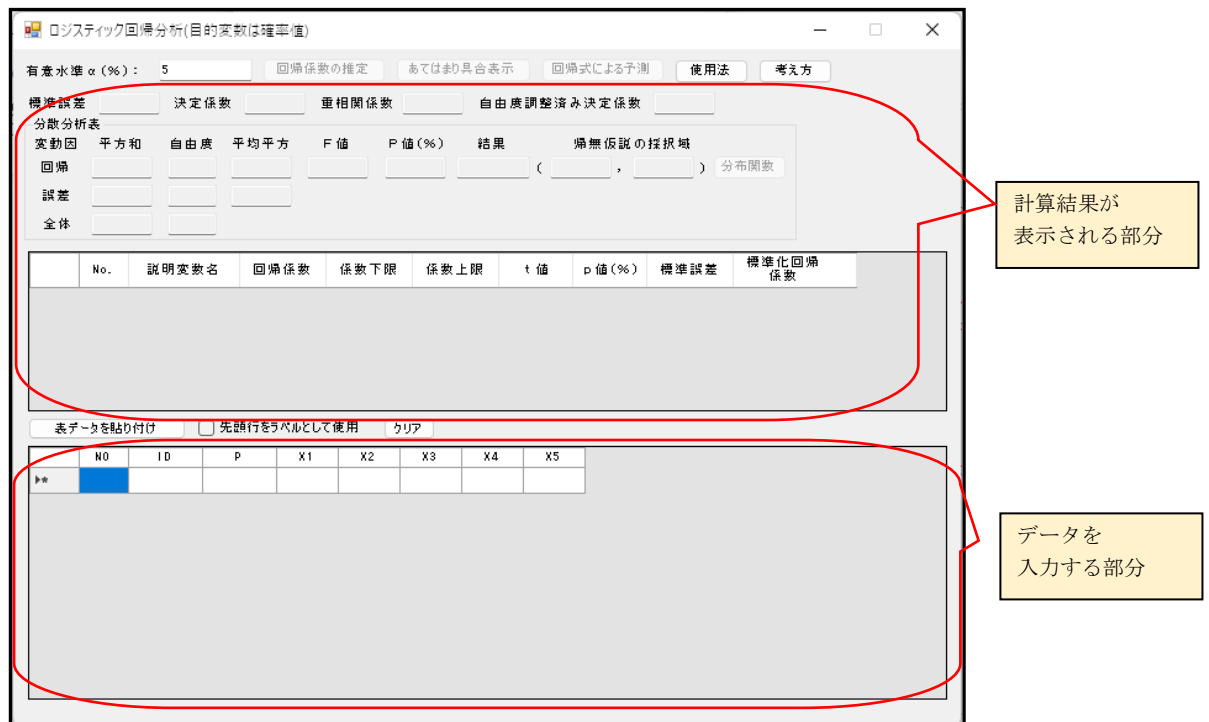
2. 使用法

(1) メニューの選択

メニューの「多変量解析→ロジスティック回帰分析」を選択します。



(2) パネルが表示されます。



有意水準 α (%) : 5 回帰係数の推定 あてはまり具合表示 回帰式による予測 使用法 考え方

標準誤差 決定係数 重相関係数 自由度調整済み決定係数

分散分析表

変動因	平方和	自由度	平均平方	F 値	P 値 (%)	結果	帰無仮説の採択域
回帰						(,)	分布関数
誤差							
全体							

表データを貼り付け ☐ 先頭行をラベルとして使用 クリア

No.	説明変数名	回帰係数	係数下限	係数上限	t 値	p 値 (%)	標準誤差	標準化回帰係数
1	X1							
2	X2							
3	X3							
4	X4							
5	X5							

(3) 計算対象データを入力

表計算ソフトに計算対象データを以下のように定義します。

ID	確率	×1
ID-1	0.05	10.00
ID-2	0.12	20.00
ID-3	0.3	30.00
ID-4	0.6	40.00
ID-5	0.8	50.00
ID-6	0.9	60.00
ID-7	0.95	70.00
ID-8	0.98	72.00

この場合、

- ・目的変数は 確率
- ・説明変数 1 は X_1

で、一番左に ID を記述します。

上記の表の 赤い部分を コピーして、表データを貼り付け をクリックすると、以下のように入力が完了します。

	NO	ID	確率	× 1
▶	1	ID-1	0.05	10.00
	2	ID-2	0.12	20.00
	3	ID-3	0.3	30.00
	4	ID-4	0.6	40.00
	5	ID-5	0.8	50.00
	6	ID-6	0.9	60.00
	7	ID-7	0.95	70.00
	8	ID-8	0.98	72.00

(4) 計算を実行

「回帰係数の推定」ボタンを押すことで、計算結果が表示されます。

標準誤差 0.2427189 決定係数 0.9913599 重相関係数 0.9956706 自由度調整済み決定係数 0.9899198

分散分析表

変動因	平方和	自由度	平均平方	F 値	P 値 (%)	結果	帰無仮説の採択域	
回帰	40.55738	1	40.55738	688.4346	2.458229E	有意	(0	3.865969)
誤差	0.3534748	6	0.0589124					分布関数
全体	40.91085	7						

No.	説明変数名	回帰係数	係数下限	係数上限	t 値	p 値 (%)	標準誤差	標準化回帰係数
0	定数項	-3.979024	-4.457572	-3.500475	-20.35234	9.145763E-05	0.195507	-0.6963506
1	x1	0.1047536	0.09498119	0.114526	26.23804	2.02219E-05	0.003992433	3.639482

(5) 計算結果の説明

1) 予測精度について

標準誤差	0.2427189	決定係数	0.9913599	重相関係数	0.9956706	自由度調整済み決定係数	0.9899198
------	-----------	------	-----------	-------	-----------	-------------	-----------

予測精度を示すものとして、

- ・ 重相関係数 0.9956
- ・ 決定係数 0.9913
- ・ 自由度調整済み決定係数 0.9999

と表示されています。

重相関係数はこの推定の当てはまり具合をしめすもので、0.0～1.0の値を取ります。1.0に近いほど、推定の精度が高いことを示します。この場合0.9956ですので、高い精度で推定していることを示します。決定係数は重相関係数の二乗で、自由度調整済み決定係数は説明変数の数を考慮して、決定係数に調整を加えたものです。

2) 分散分析表

分散分析表							
変動因	平方和	自由度	平均平方	F 値	P 値(%)	結果	帰無仮説の採択域
回帰	40.55738	1	40.55738	688.4346	2.458229E	有意	(0 , 3.865969) 分布関数
誤差	0.3534748	6	0.0589124				
全体	40.91085	7					

前提としている予測式が意味を持つのか？を示すものです。予測が意味を持たないとは、「回帰係数 a_1 が0である」ことを帰無仮説として検定を行うもので、上記の「結果」欄が“有意でない”場合に、帰無仮説が採択されます。

上記では「結果」欄が“有意”となっているので、

予測が意味を持たないとは言えない
ことを意味します。

3) 説明変数の評価

	No.	説明変数名	回帰係数	係数下限	係数上限	t 値	p 値(%)	標準誤差	標準化回帰係数
▶	0	定数項	-3.979024	-4.457572	-3.500475	-20.35234	9.145763E-05	0.195507	-0.6963506
	1	X1	0.1047536	0.09498119	0.114526	26.23804	2.02219E-05	0.003992433	3.639482

モデル式 $P = \frac{1}{1 + \exp(-(a_0 + a_1 x_1))}$ に おいて

回帰係数は、

定数項 a_0 は -3.979024

X1 a_1 は 0.1047536

であることを示しています。

また各回帰係数の t 値は、回帰係数としての有意性を検定するもので、

X1 について 26.23

この値は回帰係数が 0 であるかを示すもので、t 値がおおむね 2.0 以上ならば 0 でないと判断します。

少し詳しく言うと、 $t \text{ 値} = (\text{回帰係数} - 0.0) / \text{標準誤差}$ は、自由度 $(n - k - 1)$ (n : 個数、 k : 変数) の t 分布に従うので、上記で計算される t 値が、「回帰係数 = 0.0 とする帰無仮説」の採択域から外れていれば、帰無仮説は棄却されます。

(つまり回帰係数は 0 でないと言える)

さらに各回帰係数の p 値は、t 値から計算される有意確率です。

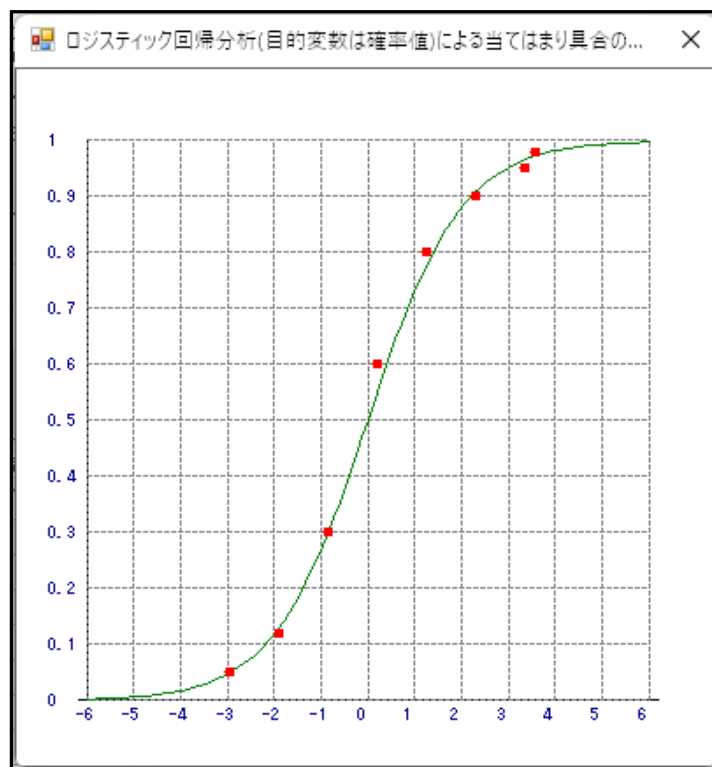
X1 について 2.02E-05 (%)

で 有意水準 (この場合 5 %) 以下ですので、説明変数として意味を持つことを示します。

(6) 数式の当てはまり具合を表示

あてはまり具合表示

ボタンを押すことで表示されます。



入力値の順に左からデータがプロットされています。
赤い点は入力値、緑の線は推計曲線を示すものです。
推計がうまくいっているのかを 直観で分かりやすいように表示しています。

(7) 回帰式による予測

回帰式による予測

を押すことで、推計を実行します。

NO	説明変数名	値
1	x1	-

計算 確率:

X1に 数値を代入して、確率を 計算できます。
X1に 4 5 を指定して、計算すると 0.6758 と確率値が計算されます。

NO	説明変数名	値
1	x1	45

計算 確率: 0.675877

(8) 変数を計算対象から外す

入力データによっては、説明変数が回帰分析において有意でない場合があります。

その場合、説明変数を省いて 再計算することができます。
普通の回帰分析と同様です。