

心理データ解析演習 2009.6.17

ロジスティック回帰分析

教育学研究科 M1

塩崎 雅恵



多変量解析法

一つの対象について3つ以上の測定値が得られている場合の分析

■ 予測のための手法

- ・重回帰分析
- ・数量化理論 類
- ・判別分析
- ・ロジスティック回帰分析

■ 分類のための手法

- ・主成分分析
- ・コレスポンデンス分析
- ・クラスター分析



ロジスティック回帰分析とは

■ 多変量解析法のひとつ

- ・ 目的変数が**質的変数**で、**名義尺度**の時使用

■ 次のような際に用いられる。

- ・ 比率のデータを目的変数として回帰分析を行う

- ・ 質的データを目的変数として回帰分析を行う

ロジスティック回帰分析とは

- もともとは冠状動脈性疾患の研究の中で、「病原菌」「発症」のような1:1関係ではなく、いくつかのリスクファクターの強弱から発症を説明するために使用。
- 重回帰分析

$$y = \beta_0 + \beta_1 X_1 + \dots + \beta_r X_r + \text{誤差}$$

この場合従属変数には体重や成績などの連続変数が仮定されている。ロジスティック回帰分析は主に「ある事象の生起の有無」を従属変数とする。



ロジスティック回帰分析とは

- 「判別分析」と同じパターンを持つが、ロジスティック回帰分析では説明変数が量的変数・質的変数・量/質混在のそれぞれで使用可能
(判別分析: 説明変数が量的変数)

ロジスティック回帰分析とは

- ある現象が発生する確率を p 、説明変数を x とすると

$$p = \frac{1}{1 + e^{\{-(b_0 + b_1x)\}}} = \frac{1}{1 + \exp\{- (b_0 + b_1x)\}}$$

$$\ln \left(\frac{p}{1 - p} \right) = b_0 + b_1x \quad \text{という回帰式と同義}$$



ロジット変換とオッズ

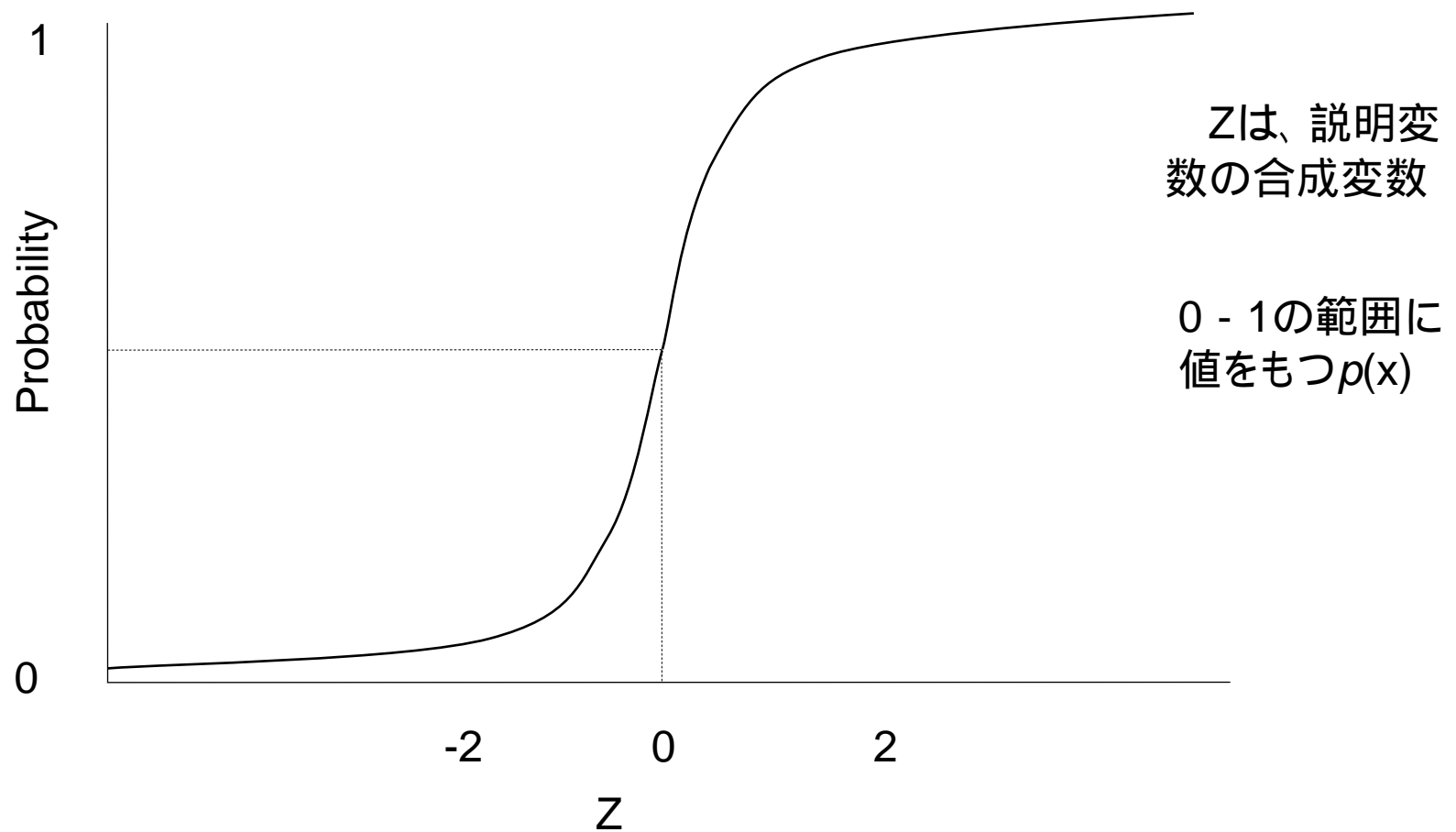
$$\ln \left(\frac{p}{1-p} \right)$$

…ロジット変換

$$\frac{p}{1-p}$$

…オッズ

ロジスティック曲線



オッズ odds と オッズ比 odds rate

- $P(Y=1|x)$ の $1-P(Y=1|x)$ に対するオッズ。
- オッズ比

$$\frac{\frac{P(Y=1|x')}{1 - P(Y=1|x')}}{\frac{P(Y=1|x)}{1 - P(Y=1|x)}} = \exp(\beta)$$



最尤法 maximum likelihood method

- 観測値が適切に理論値を推測している、という「尤度 likelihood」を最大にする方法。
- モデルに含まれるパラメータ数 r よりも、標本数 n が 4 ~ 5 倍あるのが望ましい。



SPSSをつかって

■ 二値データについて

例) 年齢 & 年収からみるマンション購入(内田
2004)

「分析」 「回帰」 「二項ロジスティック」
「共変量」に説明変数を、「従属変数」に目的
変数を投入。

「分類プロット」にチェック。

「残差のケースごとの出力」にチェックを入れ、
「全てのケース」を選択。

結果について

- ダミー変数に変換していない場合、従属変数コーディングをチェック。(ロジスティック回帰分析では、データエディタで、目的変数を文字のままにすることも可能。その場合、1となる方のグループになる確率を計算している)

従属変数コーディング

元の値	内部値
A	0
B	1

ここでは、

A(購入予定ない) = 0

B(購入予定ある) = 1

結果について

- 回帰式のあてはまりの良さについて
求めた回帰式がどれほどよくQ3の結果を予測する？
 - ・ 「分類表」をチェック

分類表*

	予測値		正分類パーセント
	Q3		
	A	B	
A	13	2	86.7
B	2	13	86.7

86.7%の正解率をもつ回帰式が得られた！！



結果について

- 誤判別をどれだけしているか？
 - ・ 「ケースごとのリスト」をチェック

- 変数ごとの有意性

それぞれの変数が、どれだけ結果を予測しているかを見る。

- ・ 「有意確率」をチェック。ロジスティック回帰の場合、0.2前後で判断。



結果

- オッズ比($\exp(\quad)$)

オッズ比を見ると、その説明変数の値が一つ増えると、オッズが何倍になるかがわかる。オッズ比が1以上のときはBである確率が、1以下のときはBではない(Aである)確率が高まる。

。



多群の判別

- 「ある」か「ない」ではなく、3つ以上の判別の際には二項ロジスティック回帰は使えない。

多項ロジスティック分析

- 五段階評定などのように、順序尺度が目的変数であるとき

累積ロジスティック分析



感度分析

- 実際に予測する値が、それまでの回帰式の算出に用いたデータから外れる場合のために、他の関数を用いてそれによる違いを見る。
- 具体的にはプロビット回帰分析など
プロビット回帰分析 $F(x) = \Phi(z)$



参考資料

- 狩野裕 2008 心理・教育測定特論講義資料
- 小塩真司 2007 SPSSとAmosによる心理・調査データ解析 因子分析・共分散構造分析まで 東京図書
- 丹後俊郎・山岡和枝・高木晴良 1996 統計ライブラリー ロジスティック回帰分析 SASを利用した統計解析の実際 朝倉出版
- 内田治 2004 すぐわかるSPSSによるアンケートの多変量解析 東京図書