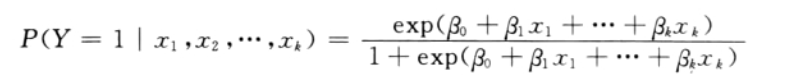
# Logistic回归

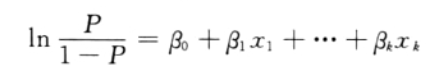
## Logistic模型

回归分析中，y可能有两种情况：①y为定量的变量，可用regress进行回归。②y为定性的变量，只有0，1取值，那么就需要进行Logistic回归。Logisitic回归主要研究某些现象发生的概率P ，比如股票涨还是跌，公司成功或失败的概率。

Logistic回归模型基本形式为：

其中，参数βi类似于多元线性回归中的参数。该式表示：当自变量为xi时，Y=1的概率。

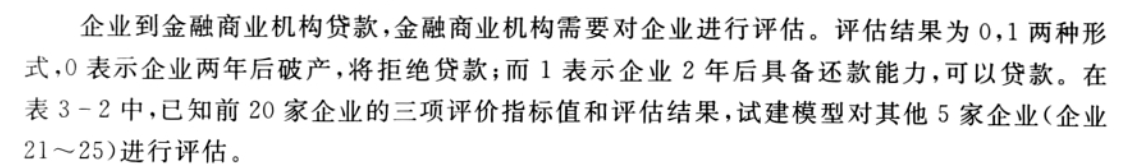
进行对数变换有：



此处，P为单调连续的概率函数，由P对Y的映射，即可得到Y。

下面举一个例子：

（其中还有matlab的文件操作，要会搞）



数据（由excel表格得到）

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 企业编号 | X1 | X2 | X3 | Y | 预测值 |
| 1 | -62.8 | -89.5 | 1.7 | 0 | 0 |
| 2 | 3.3 | -3.5 | 1.1 | 0 | 0 |
| 3 | -120.8 | -103.2 | 2.5 | 0 | 0 |
| 4 | -18.1 | -28.8 | 1.1 | 0 | 0 |
| 5 | -3.8 | -50.6 | 0.9 | 0 | 0 |
| 6 | -61.2 | -56.2 | 1.7 | 0 | 0 |
| 7 | -20.3 | -17.4 | 1 | 0 | 0 |
| 8 | -194.5 | -25.8 | 0.5 | 0 | 0 |
| 9 | 20.8 | -4.3 | 1 | 0 | 0 |
| 10 | -106.1 | -22.9 | 1.5 | 0 | 0 |
| 11 | 43 | 16.4 | 1.3 | 1 | 1 |
| 12 | 47 | 16 | 1.9 | 1 | 1 |
| 13 | -3.3 | 4 | 2.7 | 1 | 1 |
| 14 | 35 | 20.8 | 1.9 | 1 | 1 |
| 15 | 46.7 | 12.6 | 0.9 | 1 | 1 |
| 16 | 20.8 | 12.5 | 2.4 | 1 | 1 |
| 17 | 33 | 23.6 | 1.5 | 1 | 1 |
| 18 | 26.1 | 10.4 | 2.1 | 1 | 1 |
| 19 | 68.6 | 13.8 | 1.6 | 1 | 1 |
| 20 | 37.3 | 33.4 | 3.5 | 1 | 1 |
| 21 | -49.2 | -17.2 | 0.3 |  | 0 |
| 22 | -19.2 | -36.7 | 0.8 |  | 0 |
| 23 | 40.6 | 5.8 | 1.8 |  | 1 |
| 24 | 34.6 | 26.4 | 1.8 |  | 1 |
| 25 | 19.9 | 26.7 | 2.3 |  | 1 |

Matlab代码：

clear all;

clc;

X0=xlsread('E:\顾子涵专用文件夹\学习\matlab学习\Logistic回归数据.xlsx',1,'B2:D21');

XE=xlsread('E:\顾子涵专用文件夹\学习\matlab学习\Logistic回归数据.xlsx',1,'B2:D26');

Y0=xlsread('E:\顾子涵专用文件夹\学习\matlab学习\Logistic回归数据.xlsx',1,'E2:E21');

%数据转化

n=size(Y0,1);

%由于Y的值只能等于1或0

%但由于回归模型中Y不可以等于1或0

%因此将Y进行变换，变换为概率函数，在此，取中值进行计算。

for i=1:n

if Y0(i)==0

Y1(i,1)=0.25;

else

Y1(i,1)=0.75;

end

end

%进行普通的多元线性回归

%构建常数项系数

X1=ones(size(X0,1),1);

X=[X1,X0];

%在此处Y1是概率函数，Y是即将进行线性回归的因变量

Y=log(Y1./(1-Y1));

b=regress(Y,X);

%理论结果

%b=[-0.63656,0.004127,0.01692,0.53305]';

%模型的应用

for i=1:size(XE,1)

%这是Logistic回归的基本模型

Pai0=1/(1+exp(-(b(1)+b(2)\*XE(i,1)+b(3)\*XE(i,2)+b(4)\*XE(i,3))));

%基本分界点，小于0.5为0，大于0.5为1

if Pai0<0.5

P(i)=0;

else

P(i)=1;

end

end

%视察验证结果

disp(['回归系数：',num2str(b')])

disp(['评价结果：',num2str(P)])