

多重ロジスティックモデル (Walker-Duncan 法)

青木繁伸

1 目的

ロジスティック回帰を Walker-Duncan 法により行う。
R では、`glm` 関数を用いて分析を行うこともできる。

2 使用法

```
from logistic_regression import logistic_regression
logistic_regression(data, plot=True, xlabel="lambda", ylabel="probability",
                    title="Logistic regression", verbose=True)
```

2.1 引数

<code>data</code>	データフレーム。最終列が従属変数 (0/1 データ, 0:生存, 1:死亡) であること。
<code>plot</code>	あてはめ結果の図を描く (描かないときは <code>False</code> を指定)
<code>xlabel</code>	x 軸の名前
<code>ylabel</code>	y 軸の名前
<code>title</code>	図のタイトル
<code>verbose</code>	必要最小限のプリント出力をする

2.2 戻り値

回帰係数

3 使用例

```
import pandas as pd
import sys
sys.path.append("statlib")
from logistic_regression import logistic_regression

data = pd.read_csv("data/lr.data", sep="\t")
a = logistic_regression(data)
```

***** 多重ロジスティック回帰

サンプルサイズ	98
生存 (打ち切り)	85

死亡 (故障) 13

 平均值 標準偏差

x1 132.673469 14.504728

x2 223.234694 49.225088

パラメータ推定値

	偏回帰係数	標準偏差	t 値	p 値	標準化偏回帰係数
x1	0.008297	0.021208	0.391221	0.6965	0.120347
x2	0.011386	0.005740	1.983755	0.0502	0.560501
定数項	-5.645581	3.048239	1.852079	0.0671	NaN

log likelihood = -36.092

deviance = 72.184

AIC = 78.184

	以上	以下	期待値	リスク	観察値	故障率	サンプルサイ
ズ							
0	-3.090951	-2.663472	0.540297	0.054030	0	0.000000	10
1	-2.661530	-2.490380	0.729186	0.072919	0	0.000000	10
2	-2.472638	-2.436890	0.712089	0.079121	0	0.000000	9
3	-2.398670	-2.202452	0.916765	0.091676	3	0.300000	10
4	-2.160525	-2.054780	1.097699	0.109770	1	0.100000	10
5	-2.029713	-1.945594	1.206158	0.120616	1	0.100000	10
6	-1.944181	-1.742756	1.350194	0.135019	1	0.100000	10
7	-1.724572	-1.549538	1.427380	0.158598	3	0.333333	9
8	-1.522529	-1.252166	1.986057	0.198606	0	0.000000	10
9	-1.238838	0.587855	3.034176	0.303418	4	0.400000	10

左の 2 列は、各区間のλの値 (最小値と最大値)

Logistic regression

