Logistic curve for simulated data

Contents

1. Introduction

In this example we consider the fitting of an logistic curve to simulated longitudinal data. Values were simulated for 10 time points and 100 cases. Data are given in **logistic.lsf** and the data for the first three cases are shown below.

LOGISTIC.lsf

	id2	id1	y	time
1	1.00	1.00	25.16	1.00
2	1.00	2.00	32.08	2.00
3	1.00	3.00	42.74	3.00
4	1.00	4.00	49.29	4.00
5	1.00	5.00	57.90	5.00
6	1.00	6.00	63.31	6.00
7	1.00	7.00	63.93	7.00
8	1.00	8.00	66.37	8.00
9	1.00	9.00	66.94	9.00
10	1.00	10.00	70.43	10.00
11	2.00	1.00	22.88	1.00
12	2.00	2.00	31.73	2.00
13	2.00	3.00	45.18	3.00
14	2.00	4.00	53.52	4.00
15	2.00	5.00	58.14	5.00
16	2.00	6.00	65.23	6.00
17	2.00	7.00	69.70	7.00
18	2.00	8.00	74.44	8.00
19	2.00	9.00	72.36	9.00
20	2.00	10.00	70.80	10.00
21	3.00	1.00	31.12	1.00
22	3.00	2.00	36.95	2.00
23	3.00	3.00	48.48	3.00
24	3.00	4.00	55.56	4.00
25	3.00	5.00	61.76	5.00
26	3.00	6.00	71.75	6.00
27	3.00	7.00	74.55	7.00
28	3.00	8.00	79.28	8.00
29	3.00	9.00	80.72	9.00
30	3.00	10.00	86.55	10.00

The model is defined as

$$y = b_1 / [1 + s * \exp(b_2 - b_3 * Time)] + e$$

with b_1 = 70.0, b_2 = 1.25 and b_3 = -0.5 used for simulation, where

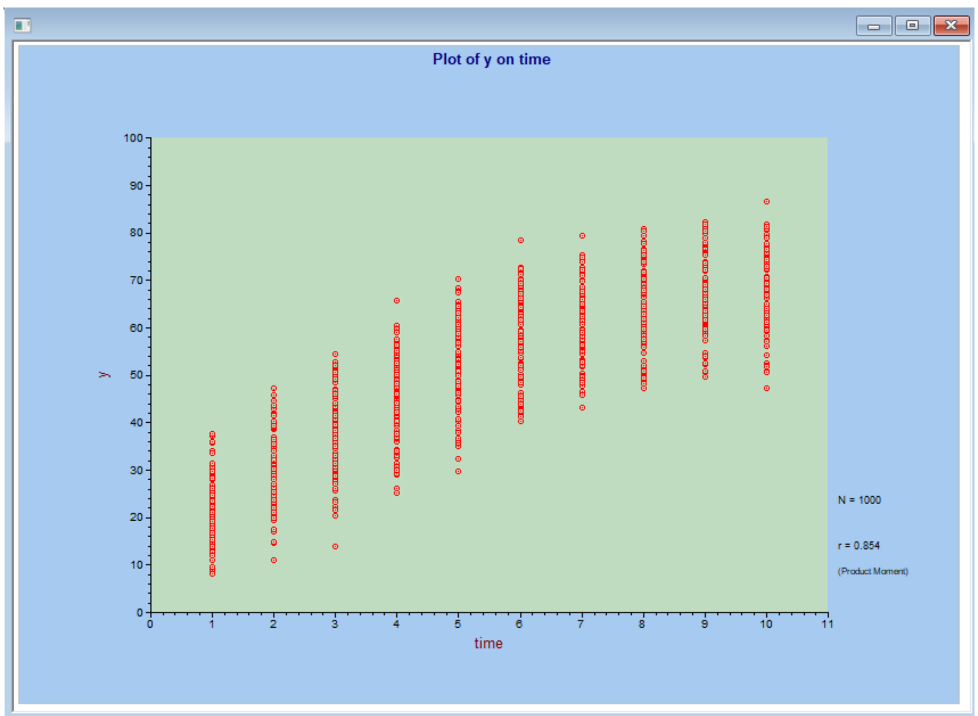
$$b_1 = \beta_1 + u_1$$
$$b_2 = \beta_2 + u_2$$
$$b_3 = \beta_3 + u_3$$

The covariance matrix of the parameters b_1, b_2 used for simulation was

$$\begin{bmatrix} 65.000 & & \\ 0.500 & 0.150 & \\ -0.090 & 0.009 & 0.010 \end{bmatrix}$$

along with a value of 2.50 for σ^2. The constant *s* has a value of 1 as function values increase monotonically over time.

A scatterplot of the *y* measurements over time is shown below. The relationship between *y* and time is nonlinear.



2. Logistic curve

We now fit a logistic curve to the data in order to examine how the estimated and simulated coefficients compare. The syntax file for this model is shown in the syntax file **logistic.prl**. The variable ID2 is used as level-2 identifier.

```
LOGISTIC.prl
!------------------------------------------------------------
!   Model simulated: Logistic, 100 cases
!
!   y=  70.00/(1+exp(1.25-0.5*t)) , t=1,2,...,10
!       Covariance matrix of parameters
!          65.000
!           0.500      0.150
!          -0.090      0.009        0.010
!       Level 1 error variance : 2.50
!------------------------------------------------------------
OPTIONS METHOD = ML CONVERGE = 0.000010 MAXITER = 30 QUADPTS = 40;
TITLE = Simulated data Single Logistic;
SY=LOGISTIC.LSF;
ID1 = id1;
ID2 = id2;
RESPONSE = y;
FIXED = time;
MODEL = Logistic;
```

The ML solution is as follows.

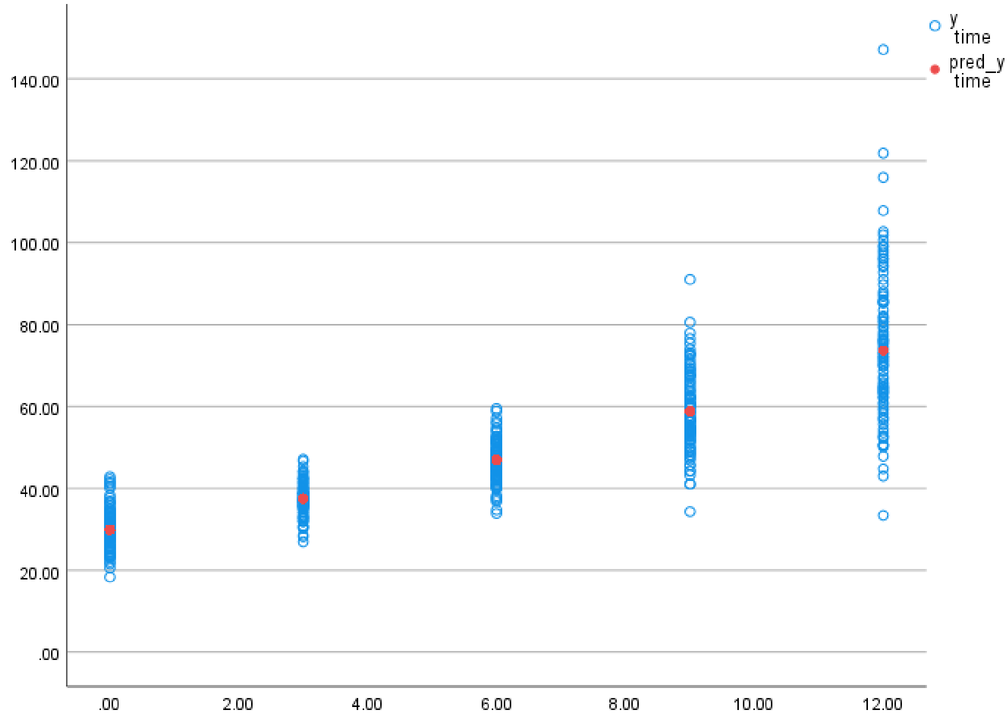
Coefficients	Beta	Std.Err.	Z-value	P > |z|
b1	69.962	0.57571	121.52	0.0000
b2	1.3118	0.31015E-01	42.296	0.0000
b3	0.50240	0.67406E-02	74.534	0.0000

Variance estimate	Level 1	Std.Err.	Z-value	P > |z|
Sigma**2	2.1505	0.68164E-01	31.548	0.19039-217

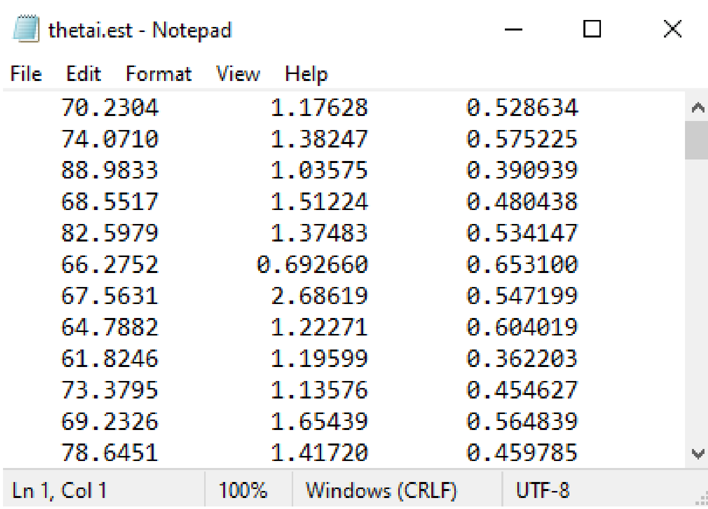
Covariances	Level 2	Std.Err.	Z-value	P > |z|
u1,u1	63.112	6.6128	9.5440	0.13744E-20
u2,u1	0.28523	0.25310	1.1269	0.25977
u2,u2	0.18348	0.19236E-01	9.5382	0.14539E-20
u3,u1	-0.12427	0.56261E-01	-2.2088	0.27188E-01
u3,u2	0.36273E-02	0.29888E-02	1.2136	0.22488
u3,u3	0.75134E-02	0.90694E-03	8.2843	0.11877E-15

The beta coefficients and the estimated σ^2 reported closely well to the values used in simulation.

A scatterplot of the predicted outcome and the simulated outcome over time is given below.



The program also writes the estimated parameters for each level-2 unit to an external text file named **thetai.est**. The contents of this file for the first few cases are as follows:



thetai.est - Notepad

File Edit Format View Help

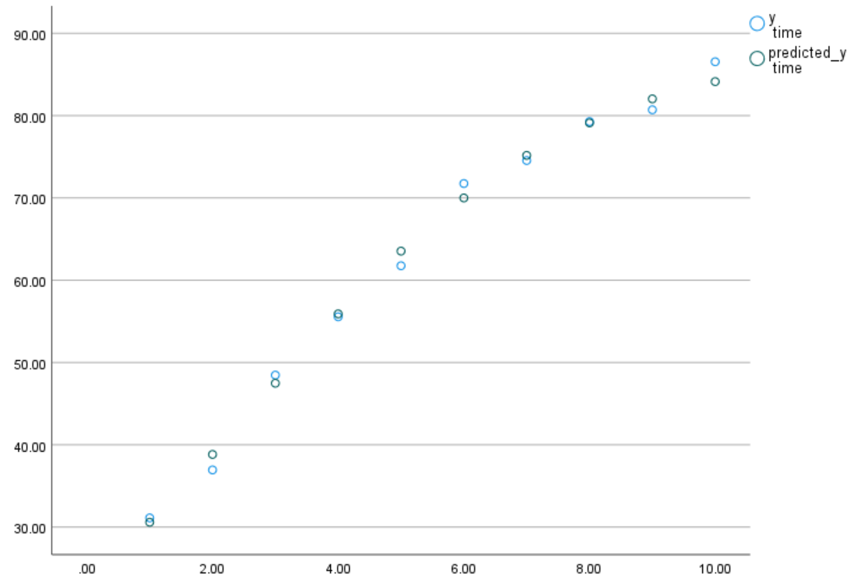
70.2304	1.17628	0.528634
74.0710	1.38247	0.575225
88.9833	1.03575	0.390939
68.5517	1.51224	0.480438
82.5979	1.37483	0.534147
66.2752	0.692660	0.653100
67.5631	2.68619	0.547199
64.7882	1.22271	0.604019
61.8246	1.19599	0.362203
73.3795	1.13576	0.454627
69.2326	1.65439	0.564839
78.6451	1.41720	0.459785

Ln 1, Col 1 100% Windows (CRLF) UTF-8

Using these results, the predicted outcome of, for example, the third case can be expressed as

$$\hat{y} = 88.8933 / [1 + \exp(1.03575 - 0.390939 * Time)]$$

which, when plotted, produces the following graph showing that the simulated and predicted y are very close at all time points.


90.00
80.00
70.00
60.00
50.00
40.00
30.00
.00
2.00
4.00
6.00
8.00
10.00
y
time
predicted_y
time