Scientific Software International

Nonlinear curve for predicting height

Contents

1. Introduction

In this example we consider the fitting of a double exponential model to the simulated height measurements of 150 males between the ages of 2 and 22. Data are given in **mal_height.lsf** and the data for the first male are shown below.

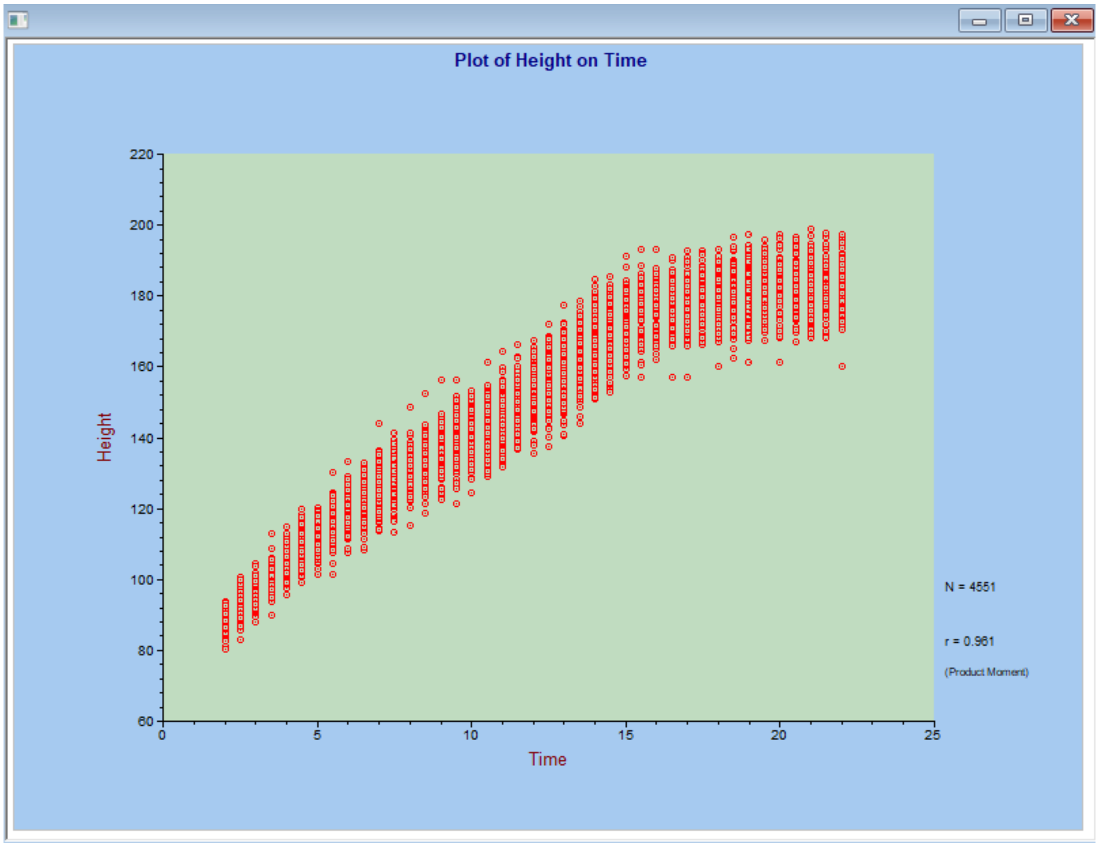
mal_height.lsf

	Case	Occasio	Height	Time
1	1.00	1.00	88.30	2.00
2	1.00	4.00	100.74	3.50
3	1.00	5.00	104.87	4.00
4	1.00	6.00	108.13	4.50
5	1.00	7.00	113.64	5.00
6	1.00	8.00	117.38	5.50
7	1.00	9.00	120.07	6.00
8	1.00	11.00	125.39	7.00
9	1.00	13.00	132.51	8.00
10	1.00	14.00	134.39	8.50
11	1.00	15.00	134.94	9.00
12	1.00	16.00	138.62	9.50
13	1.00	17.00	141.74	10.00
14	1.00	18.00	146.03	10.50
15	1.00	19.00	146.38	11.00
16	1.00	21.00	156.22	12.00
17	1.00	22.00	156.28	12.50
18	1.00	23.00	160.71	13.00
19	1.00	24.00	165.53	13.50
20	1.00	25.00	167.85	14.00
21	1.00	26.00	173.31	14.50
22	1.00	27.00	175.22	15.00
23	1.00	28.00	176.44	15.50
24	1.00	29.00	180.52	16.00
25	1.00	31.00	181.51	17.00
26	1.00	32.00	181.16	17.50
27	1.00	33.00	184.27	18.00
28	1.00	34.00	182.42	18.50
29	1.00	35.00	184.17	19.00
30	1.00	37.00	183.80	20.00
31	1.00	38.00	184.41	20.50
32	1.00	40.00	186.99	21.50
33	1.00	41.00	183.54	22.00

The variables of interest are:

- Case: the identifier of each individual
- Occasio: the measurement occasion
- Height: the height of the individual at that point in time
- Time: the time of measurement (in years)

A scatterplot of the height measurements over time is shown below. The relationship between height and time curve is nonlinear. The relationship between height and time curve is nonlinear. Moreover, the curve shows two inflection points, at approximately 8 and again at approximately 15 years of age. While a nonlinear curve like the logistic curve has a single inflection point, multiple inflection points are best handled by fitting multiple component curves.



2. Double logistic curve

We now fit a double logistic curve to these data. The 4551 observations, nested within 150 level-2 units, were simulated according to the following model:

$$y = b_1 / (1 + s * \exp(b_2 - b_3 * time) + c_1 / (1 + s * \exp(c_2 - c_3 * time) + e$$

The level-2 model is:

$$b_1 = \beta_1 + u_1$$
$$b_2 = \beta_2 + u_2$$
$$b_3 = \beta_3 + u_3$$
$$c_1 = \beta_3 + u_3$$
$$c_2 = \beta_4 + u_4$$
$$c_3 = \beta_5 + u_5$$

As we have a monotonic increase in function values over time, $s = 1$ for both components. The data values used in the simulation study were:

male_height.OUT

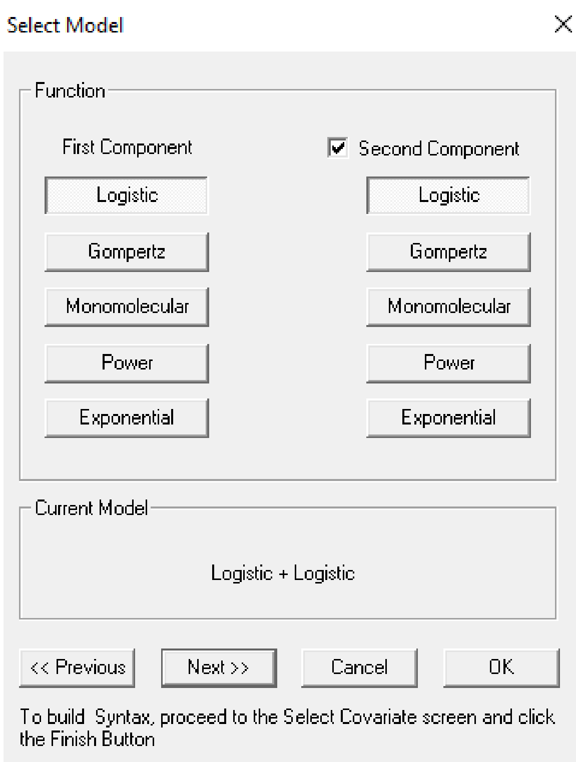
```
!  Values used for the simulation study:
!
!    --------------------------------
!      Coefficients          Beta
!    --------------------------------
!                  b1      160.13756
!                  b2        0.23788
!                  b3        0.21693
!                  c1       24.82734
!                  c2       14.32199
!                  c3        1.07272
!    --------------------------------
!      Variance estimate     Level 1
!    --------------------------------
!       Sigma**2             1.19317
!
!    --------------------------------
!     Covariances           Level 2
!    --------------------------------
!               u1,u1       73.20048
!               u2,u1        0.50440
!               u2,u2        0.00612
!               u3,u1       -0.08085
!               u3,u2       -0.00036
!               u3,u3        0.00043
!               u4,u1      -18.91052
!               u4,u2       -0.13033
!               u4,u3        0.02089
!               u4,u4       14.30563
!               u5,u1        8.96789
!               u5,u2        0.06675
!               u5,u3       -0.02324
!               u5,u4       -4.53906
!               u5,u5        6.83625
!               u6,u1        0.29343
!               u6,u2        0.00348
!               u6,u3       -0.00031
!               u6,u4       -0.16496
!               u6,u5        0.39301
!               u6,u6        0.02662
```

The syntax file for this model is shown in the syntax file **male_height.prl**. The variable Case is used as level-2 identifier (ID2).

male_height.prl

```
OPTIONS METHOD = ML CONVERGE = 0.00010 MAXITER = 50 QUADPTS = 8 ;
TITLE = Double Logistic Curve fitted to male height measurements ;
SY=mal_height.lsf;
ID1 = Occasio;
ID2 = Case;
RESPONSE = Height;
FIXED = Time;
MODEL = Logistic + Logistic;
```

The Model statement, specifying the fitting of a double logistic model, corresponds to the setting shown below on the **Select Model** dialog box accessed via the **Multilevel**, **Nonlinear Regression** option from the main menu bar.



The data summary at the top of the output file shows the number of measurements for each individual. Note that not all individuals had complete data over the time period: some individuals had only 24 measurements, while others had up to 37 measurements.

```
o-------------o
 | DATA SUMMARY |
 o-------------o
 NUMBER OF LEVEL 2 UNITS :      150
 NUMBER OF LEVEL 1 UNITS :     4551

 N2  :       1       2       3       4       5       6       7       8
 N1  :      33      29      27      25      32      30      31      29
```

N2	:	9	10	11	12	13	14	15	16
N1	:	34	32	36	28	27	31	31	24
N2	:	17	18	19	20	21	22	23	24
N1	:	32	32	27	33	26	26	28	27
N2	:	25	26	27	28	29	30	31	32
N1	:	32	28	26	29	27	30	33	30
N2	:	33	34	35	36	37	38	39	40
N1	:	30	31	32	33	31	27	26	32
N2	:	41	42	43	44	45	46	47	48
N1	:	34	28	37	28	35	33	28	29
N2	:	49	50	51	52	53	54	55	56
N1	:	31	31	30	31	33	32	31	27
N2	:	57	58	59	60	61	62	63	64
N1	:	27	31	31	29	32	33	29	31
N2	:	65	66	67	68	69	70	71	72
N1	:	31	28	31	28	33	30	31	34
N2	:	73	74	75	76	77	78	79	80
N1	:	24	29	28	29	35	29	34	25
N2	:	81	82	83	84	85	86	87	88
N1	:	32	30	29	33	37	32	28	33
N2	:	89	90	91	92	93	94	95	96
N1	:	33	35	30	34	29	32	30	32
N2	:	97	98	99	100	101	102	103	104
N1	:	32	30	25	27	33	28	26	29
N2	:	105	106	107	108	109	110	111	112
N1	:	24	25	29	29	33	35	27	29
N2	:	113	114	115	116	117	118	119	120
N1	:	34	32	33	35	24	29	22	29
N2	:	121	122	123	124	125	126	127	128
N1	:	31	29	33	31	28	33	33	36
N2	:	129	130	131	132	133	134	135	136
N1	:	33	29	25	37	33	33	33	32
N2	:	137	138	139	140	141	142	143	144
N1	:	32	34	34	28	33	30	28	25
N2	:	145	146	147	148	149	150		
N1	:	29	28	31	35	32	28		

The ML solution is as follows.

Coefficients	Beta	Std.Err.	Z-value	P > |z|
b1	160.61166	0.59720	268.94022	0.00000
b2	0.24397	0.00515	47.38377	0.00000
b3	0.21788	0.00158	138.29156	0.00000
c1	24.40840	0.31391	77.75523	0.00000
c2	14.88224	0.21439	69.41542	0.00000
c3	1.11554	0.01499	74.44330	0.00000

Variance estimate	Level 1	Std.Err.	Z-value	P > |z|
Sigma**2	1.20021	0.01793	66.95642	0.00000

Covariances	Level 2	Std.Err.	Z-value	P > |z|
u1,u1	85.39250	8.57691	9.95610	0.00000
u2,u1	0.59548	0.06630	8.98171	0.00000
u2,u2	0.00669	0.00064	10.38730	0.00000
u3,u1	-0.09588	0.01824	-5.25812	0.00000
u3,u2	-0.00039	0.00014	-2.76799	0.00564
u3,u3	0.00047	0.00006	8.11451	0.00000
u4,u1	-25.42991	3.83902	-6.62406	0.00000
u4,u2	-0.17733	0.03072	-5.77337	0.00000
u4,u3	0.02855	0.00915	3.12203	0.00180
u4,u4	16.27182	2.16655	7.51046	0.00000
u5,u1	12.49174	2.40547	5.19306	0.00000
u5,u2	0.08499	0.02008	4.23322	0.00002
u5,u3	-0.03219	0.00640	-5.02733	0.00000
u5,u4	-6.02829	1.22142	-4.93547	0.00000
u5,u5	6.10569	0.92313	6.61414	0.00000
u6,u1	0.64943	0.16097	4.03461	0.00005
u6,u2	0.00575	0.00139	4.12472	0.00004
u6,u3	-0.00121	0.00041	-2.92674	0.00343
u6,u4	-0.34338	0.08146	-4.21512	0.00002
u6,u5	0.36742	0.06175	5.95001	0.00000
u6,u6	0.02576	0.00432	5.96342	0.00000

```
Note: ML estimates of individual coefficients written to file THETAI.EST
```

The beta coefficients reported correspond well to the values used in simulation. Differences between estimated variance-covariance elements and the original values used in simulation are larger.

The average expected height can be estimated using the formula

$$\text{Predicted}(y) = 158.60887/(1+\exp(0.23233-0.22458*time)+26.01083/(1+\exp(13.69726-1.02737*time)$$

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 individuals are as follows:

*thetai.est - Notepad

File Edit Format View Help

162.505	0.269666	0.215670	24.4613	13.5744	1.03223
153.704	0.170042	0.237999	29.3896	13.1371	1.07212
168.944	0.323098	0.218352	23.5691	17.3560	1.25965
163.209	0.277383	0.234898	26.9652	13.0254	1.06260
161.140	0.301285	0.215223	24.7049	17.2298	1.34079
158.258	0.209421	0.203147	22.9422	15.9720	1.09563
156.295	0.148298	0.223612	28.1305	14.7448	1.11085
159.767	0.241107	0.238319	27.4208	13.4720	1.09314
161.176	0.222985	0.213063	24.7629	14.3016	1.01258
132.893	0.354063E-01	0.244149	29.1342	9.79453	0.800762
155.279	0.305864	0.241096	27.9643	14.0415	1.17340
158.030	0.211596	0.207123	21.6434	13.8562	0.966451
164.200	0.265891	0.217015	24.5890	15.6919	1.15548
163.219	0.259877	0.200835	21.2786	16.3698	1.19440
163.825	0.325297	0.215425	24.9229	17.0276	1.27582
170.360	0.320023	0.217444	26.2931	14.6768	1.10351

Ln 16, Col 76 100% Windows (CRLF) UTF-8

Using these results, the predicted height of, for example, the third male can be expressed as:

$$\text{Predicted}(y) = 168.944 / (1 + \exp(0.323098 - 0.218352 * time) + 23.5691 / (1 + \exp(17.3560 - 1.25965 * time)$$