Exponential curve for simulated data

Contents

1. Introduction

In this example we consider the fitting of a exponential curve to simulated longitudinal data. Values were simulated for 5 time points and 100 cases. Data are given in **exponen.lsf** and the data for the first five cases are shown below.

EXPONEN.lsf

	id2	id1	y	time
1	1.00	1.00	30.87	0.00
2	1.00	2.00	39.11	3.00
3	1.00	3.00	47.55	6.00
4	1.00	4.00	62.06	9.00
5	1.00	5.00	76.50	12.00
6	2.00	1.00	33.35	0.00
7	2.00	2.00	37.38	3.00
8	2.00	3.00	40.78	6.00
9	2.00	4.00	48.13	9.00
10	2.00	5.00	53.68	12.00
11	3.00	1.00	28.65	0.00
12	3.00	2.00	34.10	3.00
13	3.00	3.00	45.01	6.00
14	3.00	4.00	53.42	9.00
15	3.00	5.00	62.11	12.00
16	4.00	1.00	27.20	0.00
17	4.00	2.00	32.98	3.00
18	4.00	3.00	43.19	6.00
19	4.00	4.00	67.64	9.00
20	4.00	5.00	85.63	12.00
21	5.00	1.00	23.64	0.00
22	5.00	2.00	35.59	3.00
23	5.00	3.00	46.24	6.00
24	5.00	4.00	60.67	9.00
	5.00	5.00	85.21	12.00

The model is defined as

$$y = b_1 * \exp(-b_2 * Time) + e$$

with b_1 = 30.0, b_2 = -0.075 used for simulation, where

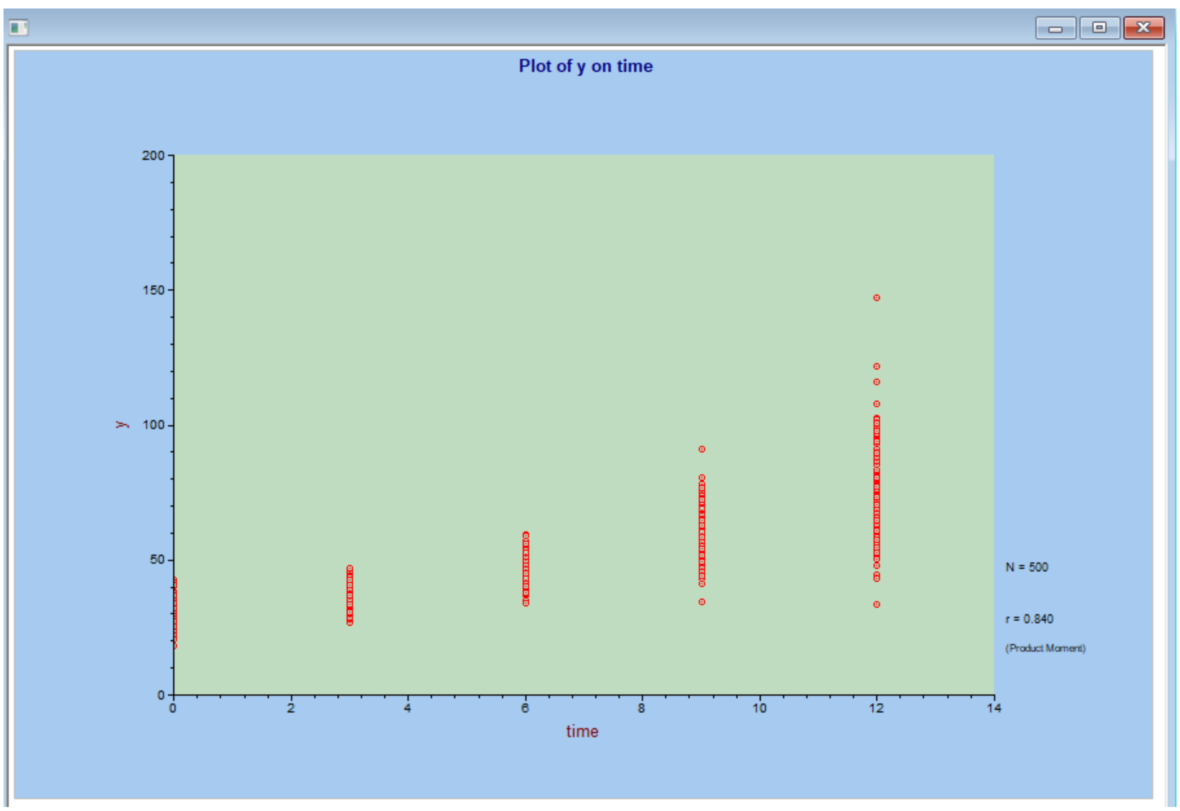
$$b_1 = \beta_1 + u_1$$
$$b_2 = \beta_2 + u_2$$

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

$$\begin{bmatrix} 30.000 & \\ 0.150 & 0.001 \end{bmatrix}$$

along with a value of 2.50 for σ^2.

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



2. Exponential curve

We now fit an exponential 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 **exponential.prl**. The variable ID2 is used as level-2 identifier.

exponential.prl

```
!-------------------------------------------------------------
!    Model simulated: 100 cases
!
!    y=  30.00*exp[-(-0.075)*t] , t=0,3,6,9,12
!        Covariance matrix of parameters
!           30.000
!            0.150     0.001
!        Level 1 error variance : 2.50
!-------------------------------------------------------------
OPTIONS METHOD = ML CONVERGE = 0.00001 MAXITER = 30 QUADPTS = 70;
TITLE = Simulated data Single Exponential;
SY=EXPONEN.LSF;
ID1 = id1;
ID2 = id2;
RESPONSE = y;
FIXED = time;
MODEL = Exponential;
```

The ML solution is as follows.

Coefficients	Beta	Std.Err.	Z-value	P > |z|
b1	29.97700	0.34520	86.83859	0.00000
b2	-0.07496	0.00214	-35.05943	0.00000

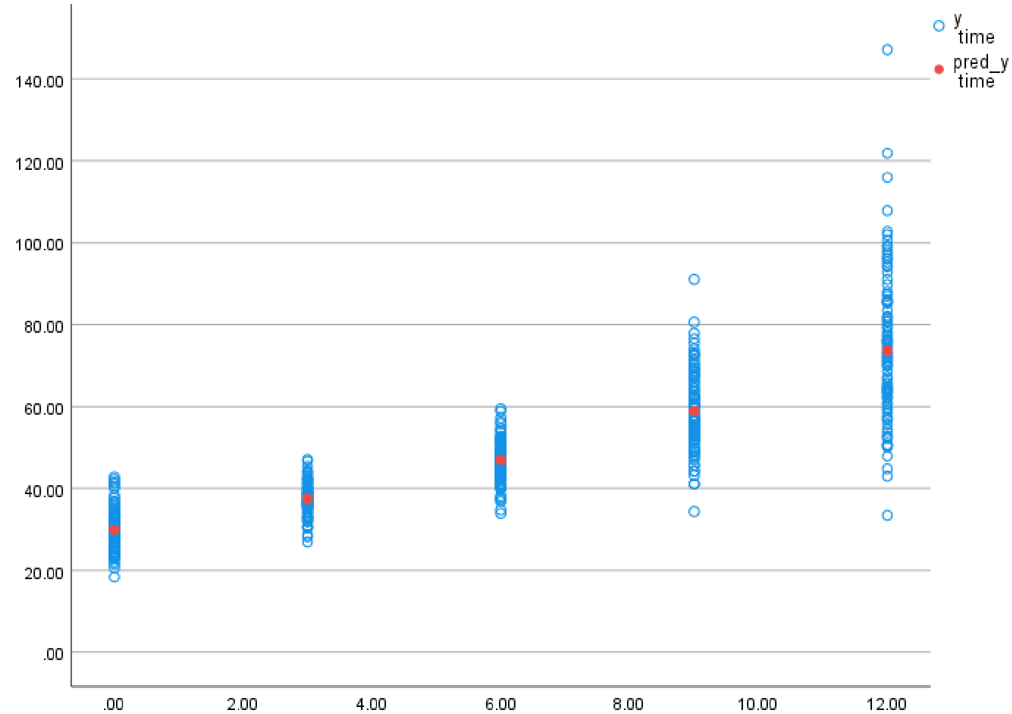
Variance estimate	Level 1	Std.Err.	Z-value	P > |z|
Sigma**2	2.55741	0.11439	22.35682	0.00000

Covariances	Level 2	Std.Err.	Z-value	P > |z|
u1,u1	22.78270	2.38375	9.55749	0.00000
u2,u1	0.12157	0.01368	8.88636	0.00000
u2,u2	0.00090	0.00009	9.84432	0.00000

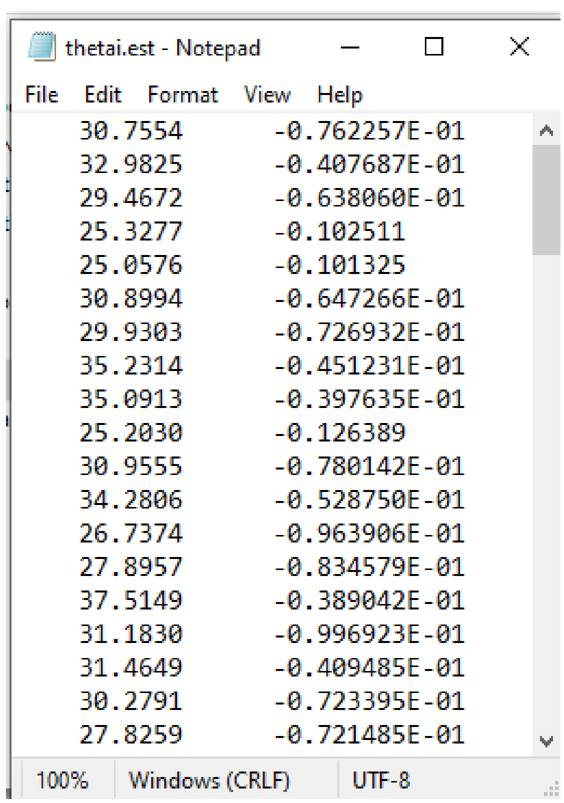
Note: ML estimates of individual coefficients written to file THETAI.EST

The beta coefficients and the estimated σ^2 reported closely correspond to the values used in simulation. The estimated $\mathrm{var}(u_1, u_1)$ of 22.78 is lower than the simulation value. For the purpose of calculating confidence intervals, the estimated standard error of b1 is 4.77, whereas the simulated value was 5.48. As such, calculated confidence intervals would still be reasonably similar. The other two variance components are close to the simulation parameters.

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:

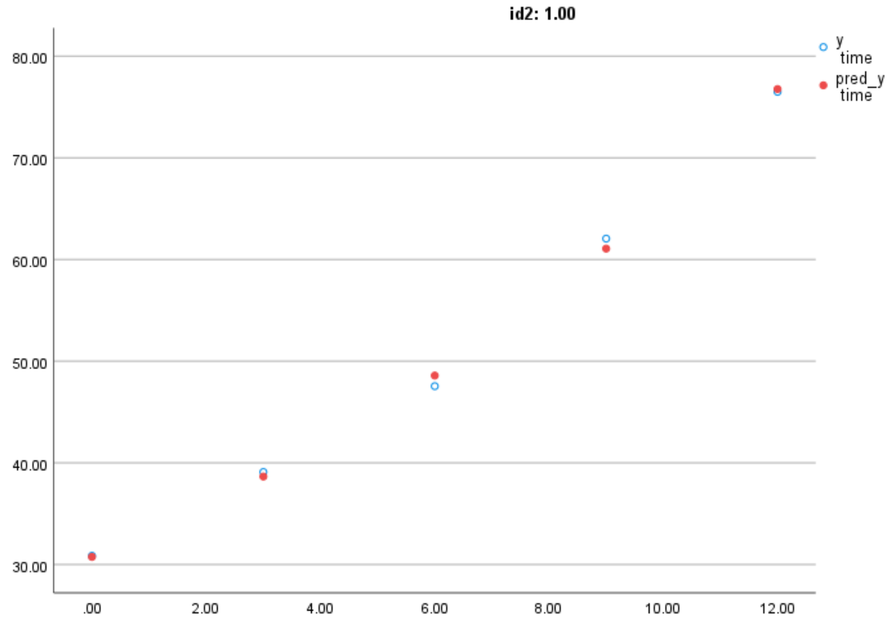


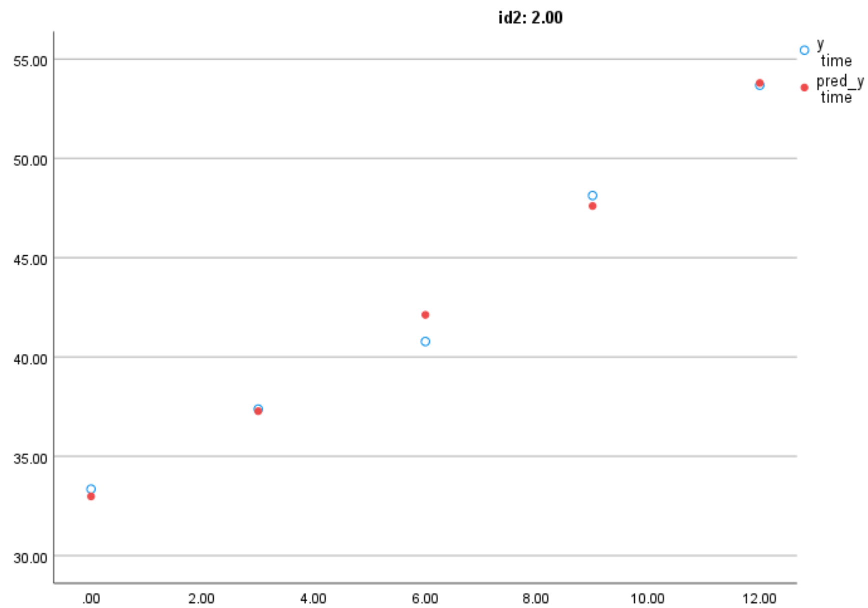
30.7554	-0.762257E-01
32.9825	-0.407687E-01
29.4672	-0.638060E-01
25.3277	-0.102511
25.0576	-0.101325
30.8994	-0.647266E-01
29.9303	-0.726932E-01
35.2314	-0.451231E-01
35.0913	-0.397635E-01
25.2030	-0.126389
30.9555	-0.780142E-01
34.2806	-0.528750E-01
26.7374	-0.963906E-01
27.8957	-0.834579E-01
37.5149	-0.389042E-01
31.1830	-0.996923E-01
31.4649	-0.409485E-01
30.2791	-0.723395E-01
27.8259	-0.721485E-01

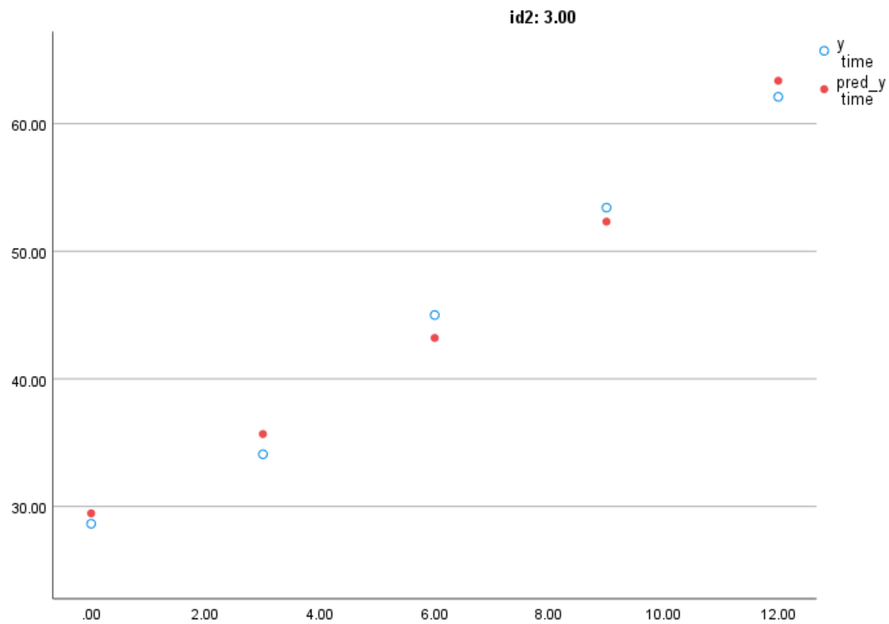
Using these results, the predicted outcome of, for example, the third case can be expressed as:

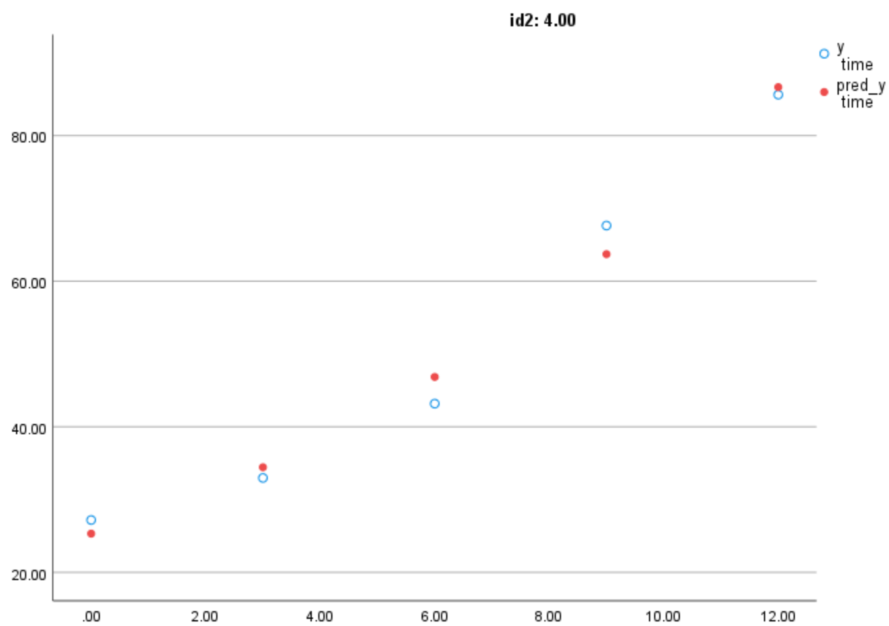
$$\hat{y} = 29.4672 * \exp(0.06381 * Time) + e$$

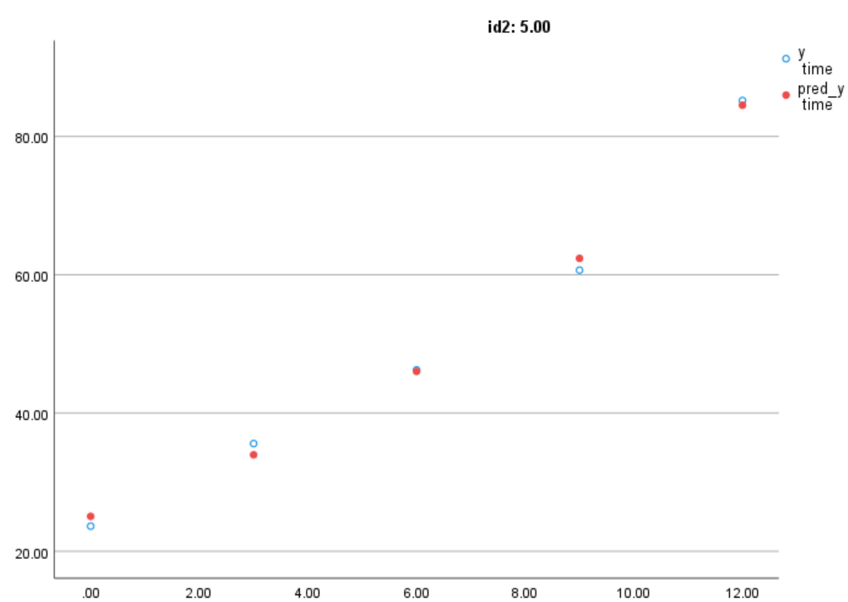
When the observed and predicted outcomes are plotted, we see that the fitted curve describes the data well, as illustrated by the plots for the first 5 cases shown below.






id2: 3.00
y
time
pred_y
time
60.00
50.00
40.00
30.00
.00
2.00
4.00
6.00
8.00
10.00
12.00


id2: 4.00
y
time
pred_y
time
80.00
60.00
40.00
20.00
.00
2.00
4.00
6.00
8.00
10.00
12.00


id2: 5.00
y
time
pred_y
time
80.00
60.00
40.00
20.00
.00
2.00
4.00
6.00
8.00
10.00
12.00