A growth curve for the weights of chicks on different diets

Contents

1. Introduction

The data file **chicks1.lsf** contains the weights (in grams) of chicks placed on four different protein diets. Weight was measures on 12 occasions, namely on days 0, 2, ,4, 20, and 21. The question is whether any of the four diets is superior to the others in terms of promoting weight gain. Note that all chicks were not measured on all occasions. In that sense, this data may be viewed as an example of longitudinal data with missing values. Data for this example is from Crowder and Hand (1990, Example 5.3).

The variables are:

- Weight: the chick's weight at time of measurement
- Diet1 to Diet4: Indicator variables indicating which diet the chick was placed on. The first two chicks were on Diet1.
- Day: indicating the day since birth on which the measurement was made
- Day_Sq: the squared value of the day of measurement

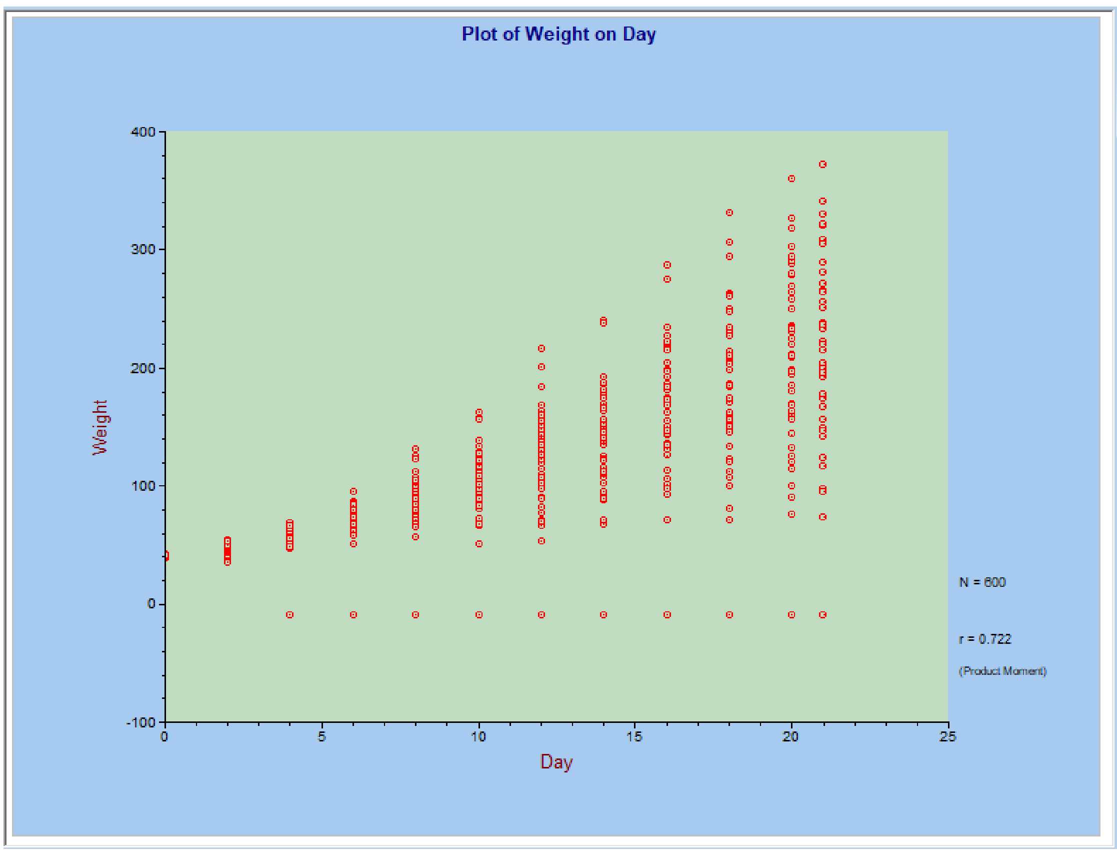
Group is coded as follows:

- Group = 1 if Diet1 = 1
- Group = 2 if Diet2 =1
- Group = 3 if Diet3 =1
- Group = 4 if Diet4 = 1.

Data for the first two chicks are shown below.

Chick_ID	Weight	Intcept	Diet1	Diet2	Diet3	Diet4	Day	Day_Sq	Group
1.00	42.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	1.00
1.00	51.00	1.00	1.00	0.00	0.00	0.00	2.00	4.00	1.00
1.00	59.00	1.00	1.00	0.00	0.00	0.00	4.00	16.00	1.00
1.00	64.00	1.00	1.00	0.00	0.00	0.00	6.00	36.00	1.00
1.00	76.00	1.00	1.00	0.00	0.00	0.00	8.00	64.00	1.00
1.00	93.00	1.00	1.00	0.00	0.00	0.00	10.00	100.00	1.00
1.00	106.00	1.00	1.00	0.00	0.00	0.00	12.00	144.00	1.00
1.00	125.00	1.00	1.00	0.00	0.00	0.00	14.00	196.00	1.00
1.00	149.00	1.00	1.00	0.00	0.00	0.00	16.00	256.00	1.00
1.00	171.00	1.00	1.00	0.00	0.00	0.00	18.00	324.00	1.00
1.00	199.00	1.00	1.00	0.00	0.00	0.00	20.00	400.00	1.00
1.00	205.00	1.00	1.00	0.00	0.00	0.00	21.00	441.00	1.00
2.00	40.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	1.00
2.00	49.00	1.00	1.00	0.00	0.00	0.00	2.00	4.00	1.00
2.00	58.00	1.00	1.00	0.00	0.00	0.00	4.00	16.00	1.00
2.00	72.00	1.00	1.00	0.00	0.00	0.00	6.00	36.00	1.00
2.00	84.00	1.00	1.00	0.00	0.00	0.00	8.00	64.00	1.00
2.00	103.00	1.00	1.00	0.00	0.00	0.00	10.00	100.00	1.00
2.00	122.00	1.00	1.00	0.00	0.00	0.00	12.00	144.00	1.00
2.00	138.00	1.00	1.00	0.00	0.00	0.00	14.00	196.00	1.00
2.00	162.00	1.00	1.00	0.00	0.00	0.00	16.00	256.00	1.00
2.00	187.00	1.00	1.00	0.00	0.00	0.00	18.00	324.00	1.00
2.00	209.00	1.00	1.00	0.00	0.00	0.00	20.00	400.00	1.00
2.00	215.00	1.00	1.00	0.00	0.00	0.00	21.00	441.00	1.00

A scatterplot of the weights against time of measurement is obtained by using the **Graphs**, **Bivariate** option from the main menu bar. Weight is specified as Y-variable, Day as x-variable.



The scatterplot indicates that the growth curve may be quadratic in nature. It also seems that the within chick error variance may be increasing over time.

2. Quadratic model

We first fit a model using the variables Day and Day_sq are predictors of weight and Diet1 to Diet4 are used to indicate the diet given:

$$y_{it} = \beta_1(Diet1)_{it} + \beta_2(Diet2)_{it} + \beta_3(Diet3)_{it} + \beta_4(Diet4)_{it}$$
$$+b_{1i}(Day)_{it} + b_{2i}(Day\_sq)_{it} + (Day)_{it}e_{it}$$

where

$b_{1i} = \beta_5 + u_{1i}$ represents the fixed and random effects of the variable Day, $b_{2i} = \beta_6 + u_{2i}$ represents the fixed and random effects of Day_Sq, and the term $(Day)_{it}e_{it}$ is used to indicate the level-1 error term.

The PRELIS syntax file **chicks0.prl** shows the syntax for this model:

```
OPTIONS MAXITER = 100;
TITLE = Weights of 50 chicks on 4 diets;
SY=Chicks1.lsf;
MISSING_DAT = -9.0;
ID1 = Day;
ID2 = Chick_ID;
RESPONSE = Weight;
FIXED = Diet1 Diet2 Diet3 Diet4 Day Day_Sq;
RANDOM1 = Day;
RANDOM2 = intcept Day Day_sq;
```

The following output is obtained. The diet that seems to produce the most weight gain is Diet4, with an estimated coefficient of 40.29. All effects are statistically significant, so the assumption of a quadratic growth curve seems reasonable. The same applies to the assumption of a linearly increasing level-1 error variance. Significant variation in the intercept, slope for Day and slope for Day_Sq is observed over the chicks.

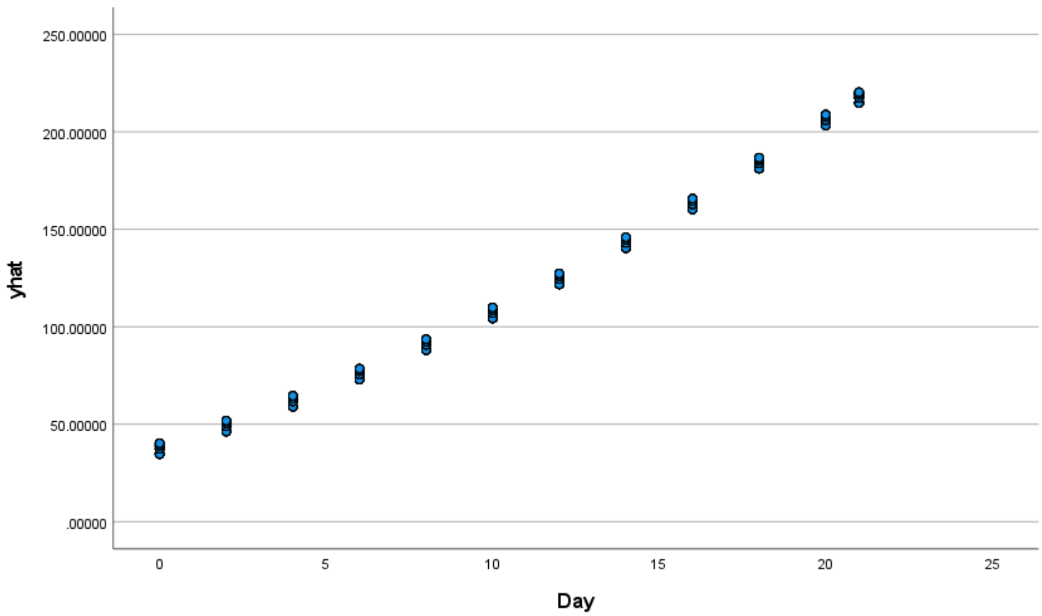
```
 ITERATION NUMBER    13

                              +-----------------------+
                              | FIXED PART OF MODEL   |
                              +-----------------------+

------------------------------------------------------------------------------
  COEFFICIENTS               BETA-HAT      STD.ERR.      Z-VALUE       PR > |Z|
------------------------------------------------------------------------------
  Diet1                     34.76194       1.11540      31.16547       0.00000
  Diet2                     37.47333       1.29962      28.83412       0.00000
  Diet3                     39.09500       1.29962      30.08193       0.00000
  Diet4                     40.29288       1.29992      30.99654       0.00000
  Day                        5.49498       0.42191      13.02406       0.00000
  Day_Sq                     0.14664       0.03172       4.62273       0.00000


                              +-----------------------+
                              |   -2 LOG-LIKELIHOOD   |
                              +-----------------------+

  DEVIANCE= -2*LOG(LIKELIHOOD) =     3963.22086782274
  NUMBER OF FREE PARAMETERS =                  13
```

```
                    +-----------------------+
                    | RANDOM PART OF MODEL  |
                    +-----------------------+
```

LEVEL 2	TAU-HAT	STD.ERR.	Z-VALUE	PR > |Z|
intcept /intcept	41.94086	9.79004	4.28403	0.00002
Day /intcept	-15.70037	3.90172	-4.02396	0.00006
Day /Day	7.77129	1.76392	4.40570	0.00001
Day_Sq /intcept	0.32469	0.22965	1.41384	0.15741
Day_Sq /Day	-0.31089	0.10712	-2.90225	0.00370
Day_Sq /Day_Sq	0.04603	0.00995	4.62416	0.00000

LEVEL 1	TAU-HAT	STD.ERR.	Z-VALUE	PR > |Z|
Day /Day	0.41172	0.02986	13.78701	0.00000

The curve fitted under this model is shown below.



3. Logistic model

Using Group as a covariate with recoded values

We next consider the logistic model

$$y_{it} = b_1 / \left[1 + s * \exp\left(b_2 - b_3 Day_{it}\right)\right] + e_{it}$$

where

$$b_1 = \beta_1 + \gamma_1 GROUP + u_{1i}$$
$$b_2 = \beta_2 + \gamma_2 GROUP + u_{2i}$$
$$b_3 = \beta_3 + \gamma_3 GROUP + u_{3i}$$

Syntax for this model is given in **chicks_log1.prl**, as shown below:

```
OPTIONS METHOD = ML CONVERGE = 0.00001 MAXITER = 100 QUADPTS =75;
TITLE = Weights of 50 chicks on 4 diets;
SY=newchicks.lsf;
MISSING_DAT = -9.0;
ID1 = Day;
ID2 = Chick;
RESPONSE = log_wgt;
FIXED = Day;
MODEL = Logistic;
COVARIATES b1 = Group
           b2 = Group
           b3 = Group;
```

The Logistic model is specified on the MODEL command and the COVARIATES command is used to indicate the categories of Group.

Note that GROUP was assigned the values 1, 2, 3 and 4 to denote the 4 diets. However, we cannot even consider GROUP as an ordinal variable, since we do not know if the diets were ordered from most to least weight loss.

Instead of using the values 1, 2, 3, and 4 for the variable Group, we recode this variable using the results obtained from the quadratic model.

COEFFICIENTS	BETA-HAT	STD.ERR.	Z-VALUE	PR > |Z|
Diet1	34.76194	1.11540	31.16547	0.00000
Diet2	37.47333	1.29962	28.83412	0.00000
Diet3	39.09500	1.29962	30.08193	0.00000
Diet4	40.29288	1.29992	30.99654	0.00000

For Diet1, the estimated coefficient was 34.76194. We now set Group = 0 for chicks assigned to Diet11. For chicks on Diet2, we set Group to 2.71, that is approximately equal to 37.47 – 34.76. Chicks on Diet3 are assigned the value 4.34 and those in Diet4 has a Group value of 5.53.

The maximum likelihood solution for this analysis using the revised data file **newchicks.lsf** is shown below.

Results are as follows:

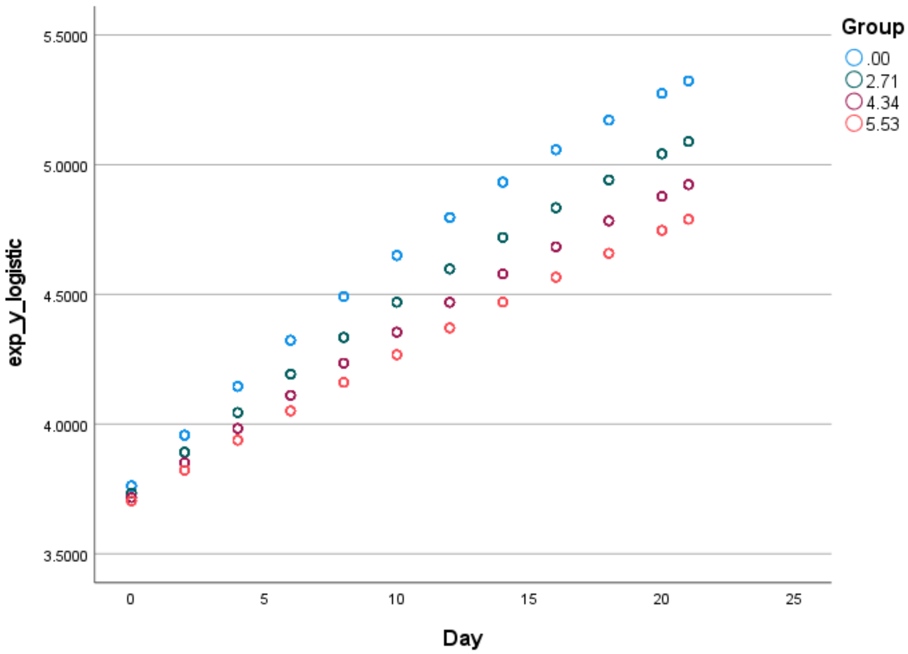
Coefficients	Beta	Std.Err.	Z-value	P > |z|
b1	6.10242	0.10470	58.28496	0.00000
b2	-0.47471	0.04662	-10.18304	0.00000
b3	0.06891	0.00518	13.29589	0.00000

Covariate Names	GAMMAH	Std.Err.	Z-value	P > |z|
Group	0.00340	0.02943	0.11549	0.90805
Group	0.00720	0.01322	0.54481	0.58588
Group	0.00506	0.00152	3.33056	0.00087

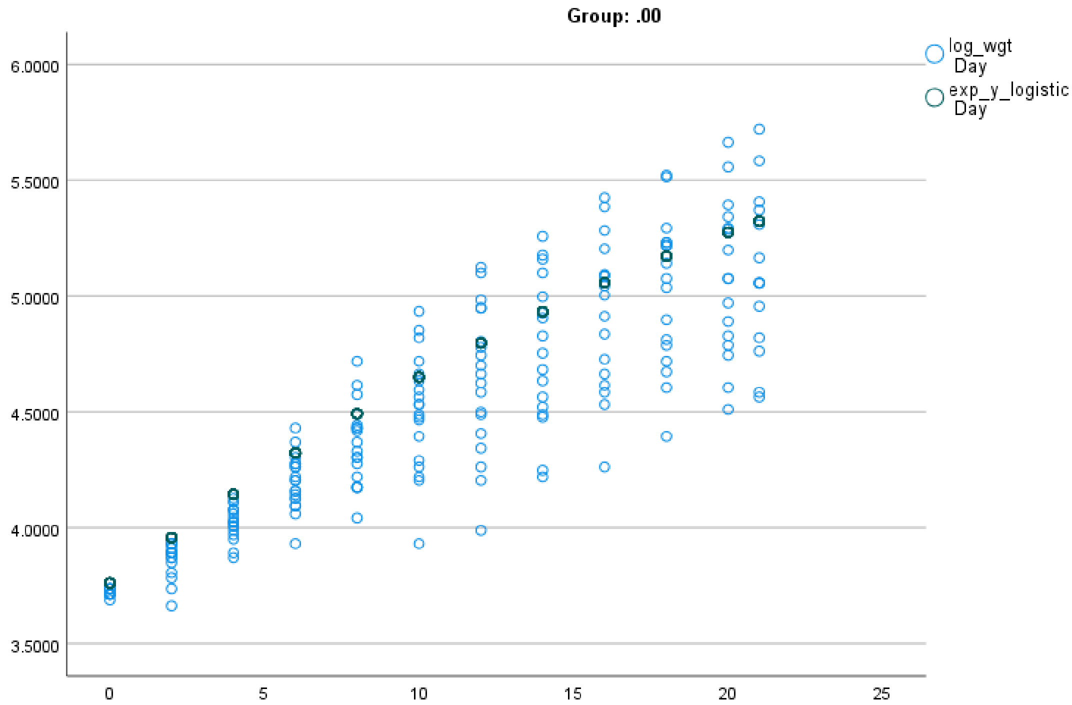
Variance estimate	Level 1	Std.Err.	Z-value	P > |z|

```
---------------------------------------------------------------------------
   Sigma**2            0.00307       0.00013       23.41348        0.00000

---------------------------------------------------------------------------
  Covariances          Level 2       Std.Err.       Z-value        P > |z|
---------------------------------------------------------------------------
          u1,u1        0.35283       0.05856        6.02468        0.00000
          u2,u1        0.16031       0.02612        6.13754        0.00000
          u2,u2        0.07519       0.01191        6.31180        0.00000
          u3,u1       -0.01333       0.00268       -4.98090        0.00000
          u3,u2       -0.00551       0.00116       -4.73964        0.00000
          u3,u3        0.00102       0.00016        6.34258        0.00000
```

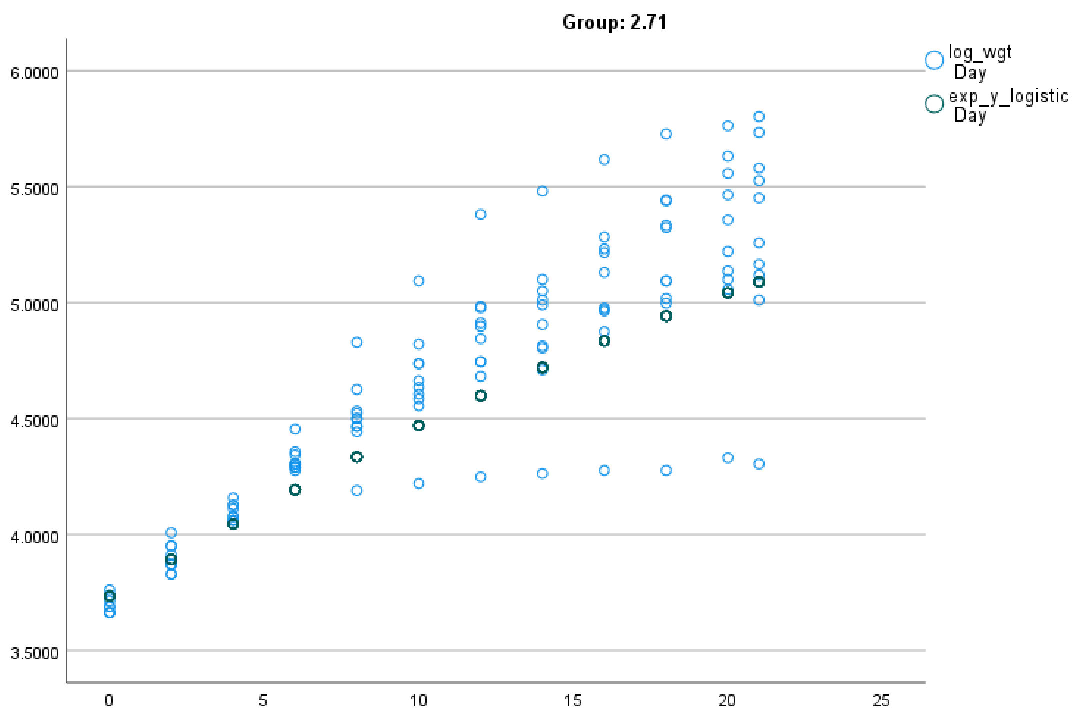
The expected log_wgt of the four groups of chicks (Diet1 through Diet4) as obtained under the logistic model are plotted below over time. It seems as if there is a difference between the 4 diets.



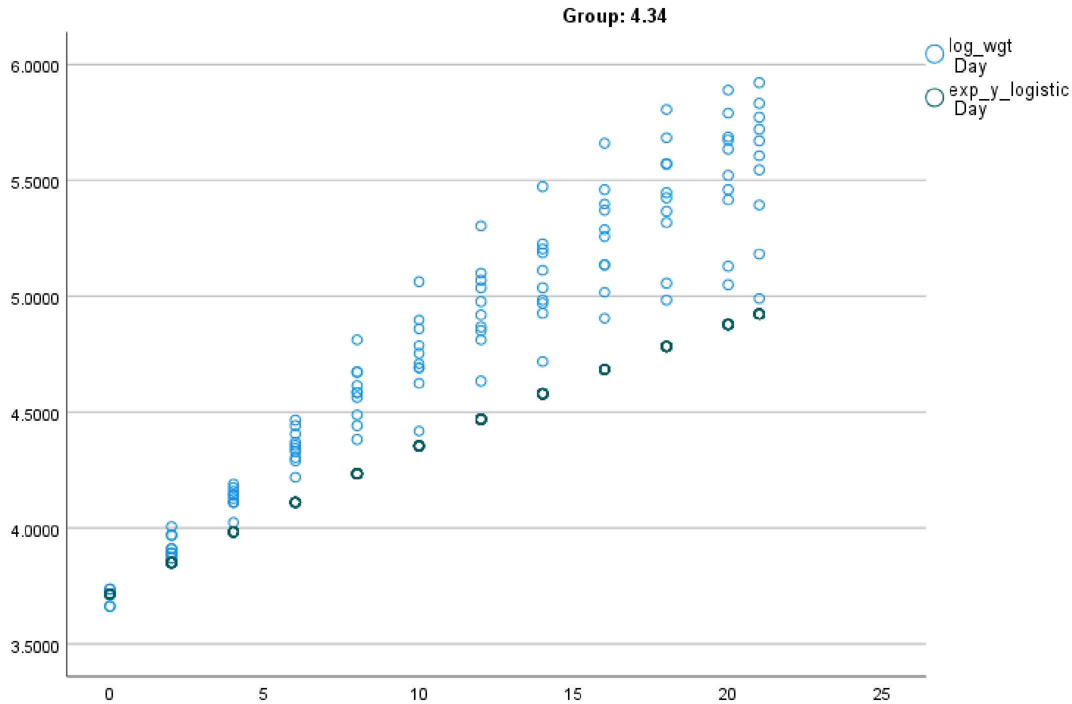
Turning our attention to the four diet groups, we now compare the estimated and observed log_wgt over time for each group.

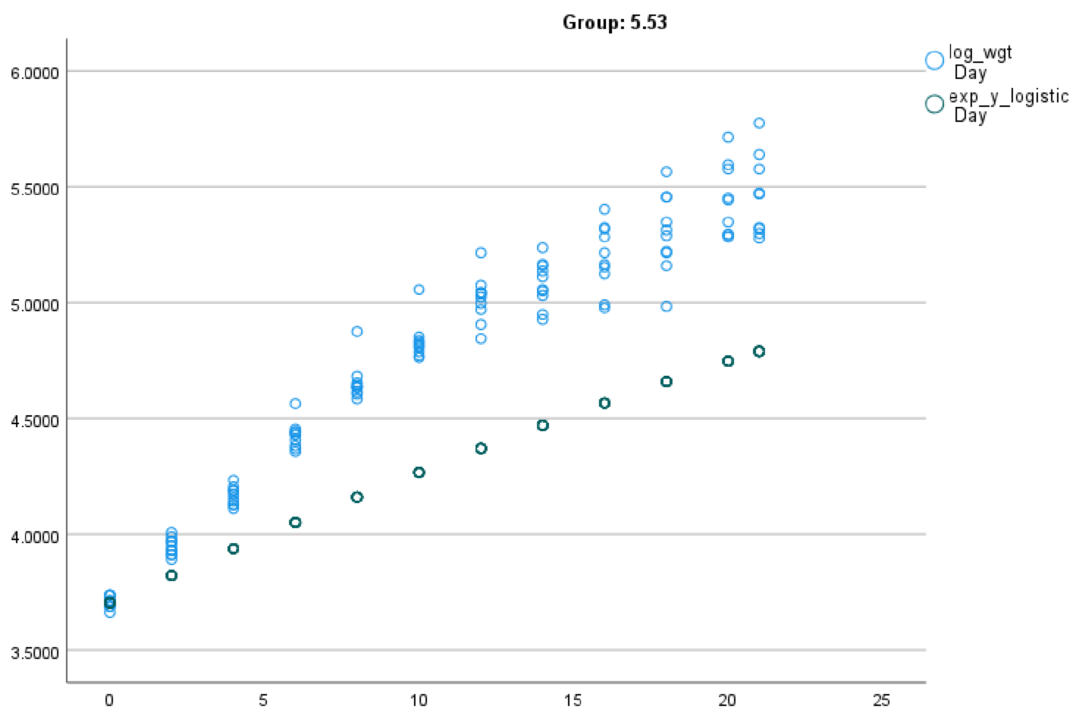


Group: .00
log_wgt
Day
exp_y_logistic
Day
6.0000
5.5000
5.0000
4.5000
4.0000
3.5000
0
5
10
15
20
25



Group: 2.71
log_wgt
Day
exp_y_logistic
Day
6.0000
5.5000
5.0000
4.5000
4.0000
3.5000
0
5
10
15
20
25





While the form of the fitted curve is similar to that observed in the raw data, it is clear the expected outcome under the current model generally underestimates the log_wgt, especially for the last two groups.

4. Gompertz model

Another model that may be used to describe the growth over time, is the Gompertz curve. The Gompertz curve we fit here is defined as

$$log\_wgt_{it} = b_1 * Exp\left(-b_2 * \exp(-s * b_3 * Day_{it}\right)$$

where

$$b_1 = \beta_1 + \gamma_{h1} * Group_i + u_{1i}$$
$$b_2 = \beta_2 + \gamma_{h2} * Group + u_{2i}$$
$$b_3 = \beta_3 + \gamma_{h3} * Group + u_{3i}$$

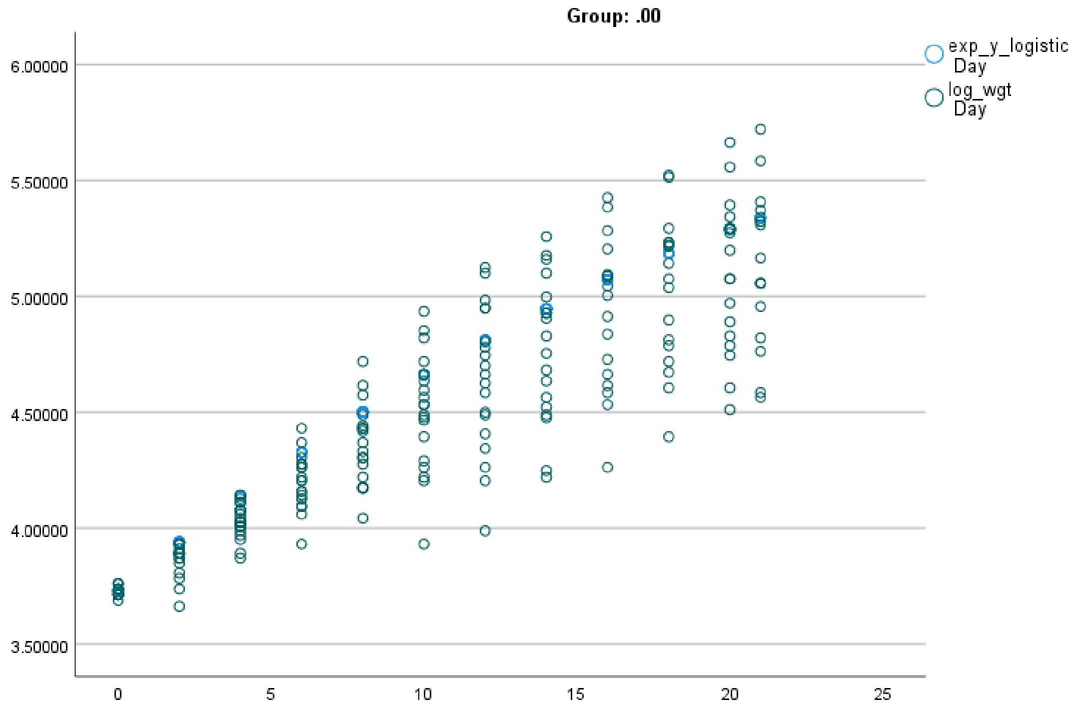
Syntax is given in **chicks_gompertz.prl** (shown below).

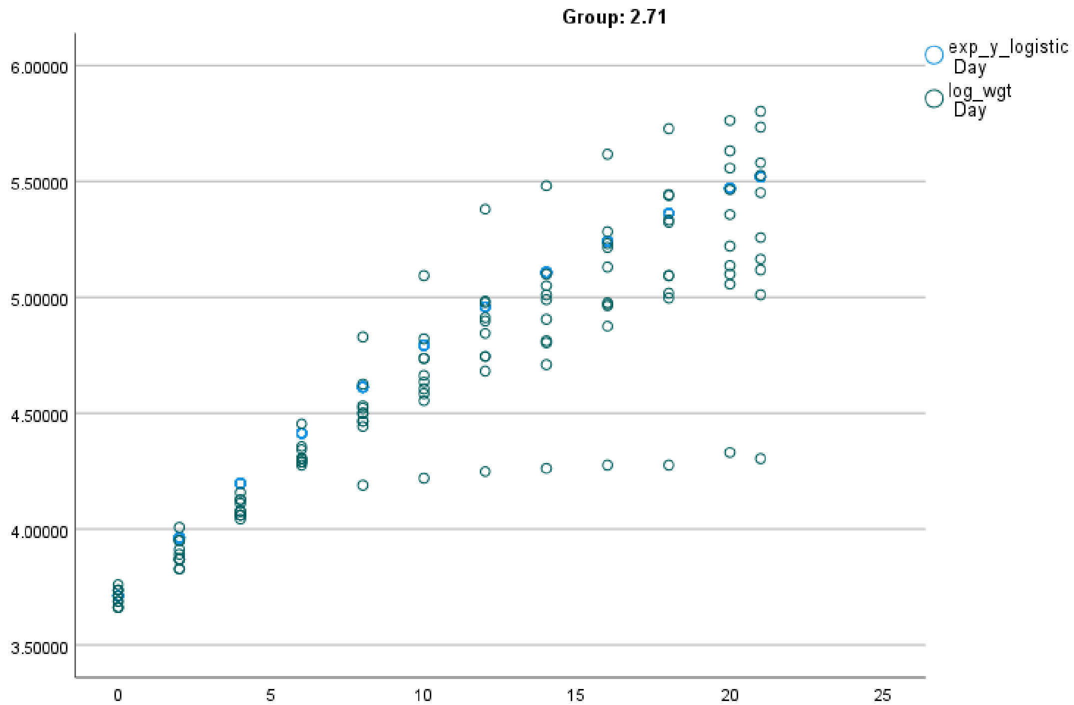
```
OPTIONS METHOD = ML CONVERGE = 0.00001 MAXITER = 1000 QUADPTS =100;
TITLE = Weights of 50 chicks on 4 diets;
SY=newchicks.lsf;
MISSING_DAT = -9.0;
ID1 = Day;
ID2 = Chick;
RESPONSE = log_wgt;
FIXED = Day;
MODEL = Gompertz;
COVARIATES b1 = Group
           b2 = Group
           b3 = Group;
```

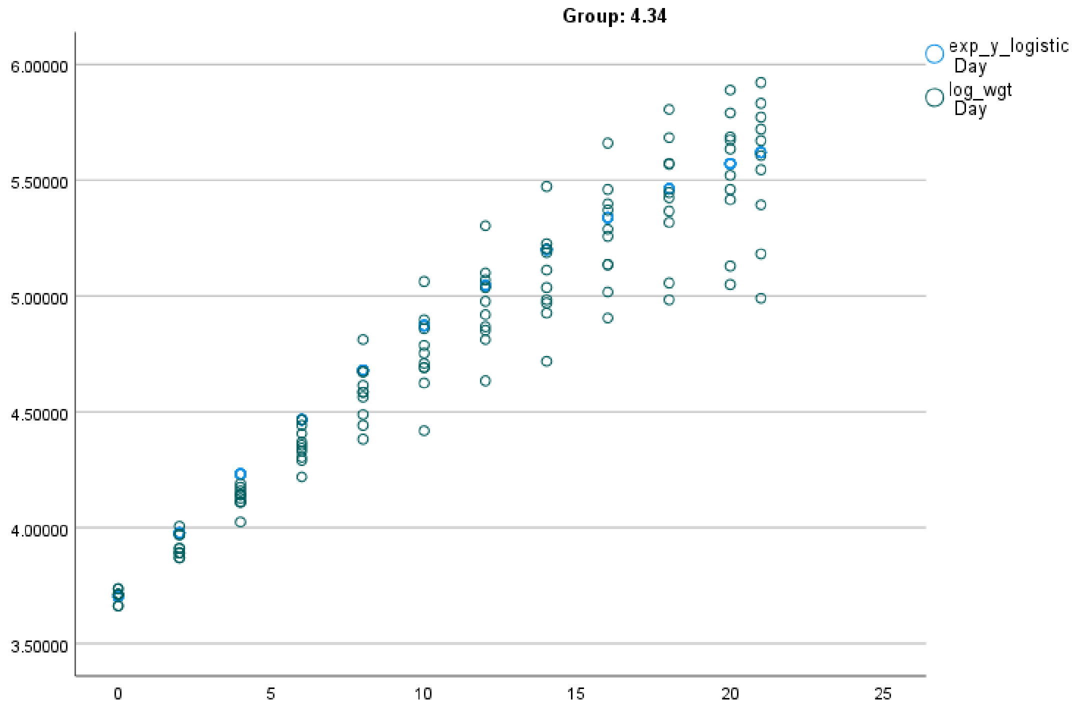
We obtain the following maximum likelihood solution. When we have monotonically increasing function values with an increase over time, $s = 1$.

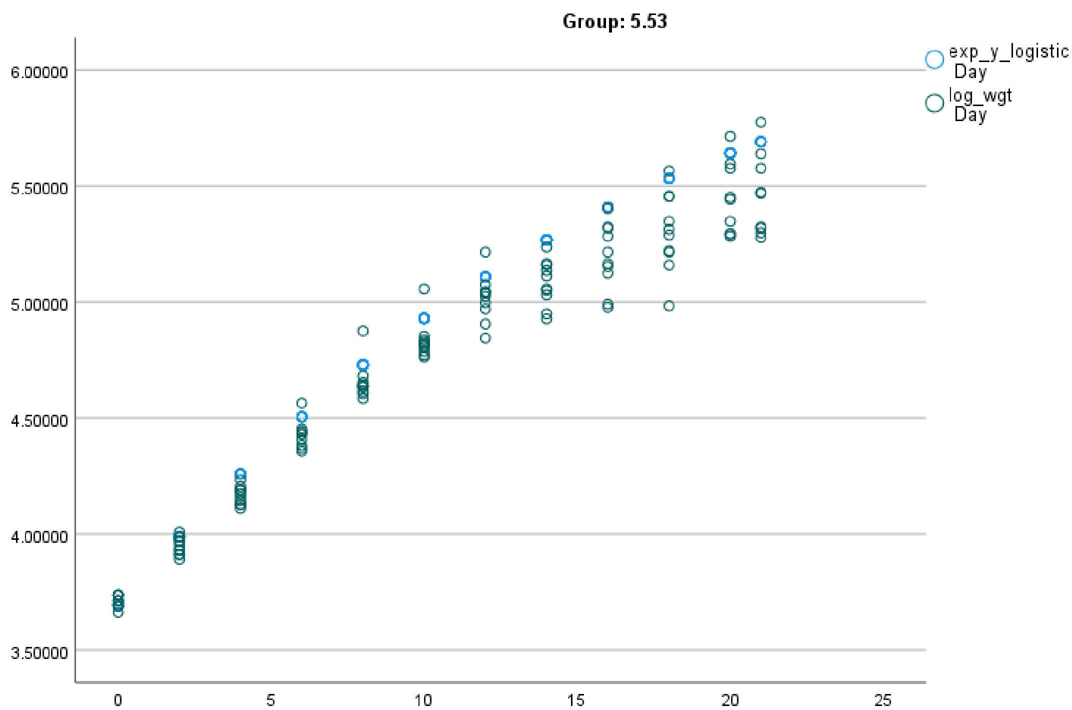
```
   --------------------------------------------------------------------------
   Coefficients          Beta          Std.Err.       Z-value        P > |z|
  --------------------------------------------------------------------------
              b1         6.24295       0.13695        45.58499       0.00000
              b2         0.51536       0.02266        22.74520       0.00000
              b3         0.05677       0.00504        11.27342       0.00000

  --------------------------------------------------------------------------
   Covariate Names       GAMMAH        Std.Err.       Z-value        P > |z|
  --------------------------------------------------------------------------
 Group                   0.01993       0.03832        0.52016        0.60295
 Group                   0.00479       0.00637        0.75229        0.45188
 Group                   0.00345       0.00146        2.35898        0.01833

  --------------------------------------------------------------------------
   Variance estimate     Level 1       Std.Err.       Z-value        P > |z|
  --------------------------------------------------------------------------
    Sigma**2             0.00321       0.00013        24.15818       0.00000
```

The expected and observed log_wgt of the four groups of chicks (Diet1 through Diet4) as obtained under the Gompertz model are plotted below over time. When compared to the similar graphs obtained under the logistic model, the Gompertz model clearly does a better job at estimating the expected outcome.









Comparing the average log_wgt for the four groups over the entire period, we note that those obtained under the Gompertz are closer to the observed values too. In all four groups the expected outcome is higher than the observed.

Descriptive Statistics[a]

	N	Minimum	Maximum	Mean	Std. Deviation
log_wgt	218	3.66356	5.72031	4.4991228	.52215194
predicted(log_wgt)	218	3.72882	5.33884	4.6368993	.51484385

a. Group = .00

Descriptive Statistics[a]

	N	Minimum	Maximum	Mean	Std. Deviation
log_wgt	120	3.66356	5.80212	4.6438209	.58207257
predicted(log_wgt)	120	3.71258	5.51943	4.7794168	.58130301

a. Group = 2.71

Descriptive Statistics[a]

	N	Minimum	Maximum	Mean	Std. Deviation
log_wgt	120	3.66356	5.92158	4.7713420	.63776068
predicted(log_wgt)	120	3.70271	5.62021	4.8475374	.61650719

a. Group = 4.34

Descriptive Statistics[a]

	N	Minimum	Maximum	Mean	Std. Deviation
log_wgt	118	3.66356	5.77455	4.7585140	.57528632
predicted(log_wgt)	118	3.69546	5.69032	4.8830451	.63850354

a. Group = 5.53

At the onset of the study, the average expected and observed log_wgt for the four groups are very close:

Descriptive Statistics[a]

	N	Minimum	Maximum	Mean	Std. Deviation
log_wgt	19	3.68888	3.76120	3.7261332	.02016690
predicted(log_wgt)	19	3.72882	3.72882	3.7288237	.00000000

a. Group = .00

Descriptive Statistics[a]

	N	Minimum	Maximum	Mean	Std. Deviation
log_wgt	10	3.66356	3.76120	3.7056220	.03666899
predicted(log_wgt)	10	3.71258	3.71258	3.7125766	.00000000

a. Group = 2.71

Descriptive Statistics[a]

	N	Minimum	Maximum	Mean	Std. Deviation
log_wgt	10	3.66356	3.73767	3.7083880	.02559316
predicted(log_wgt)	10	3.70271	3.70271	3.7027069	.00000000

a. Group = 4.34

Descriptive Statistics[a]

	N	Minimum	Maximum	Mean	Std. Deviation
log_wgt	10	3.66356	3.73767	3.7132710	.02591126
predicted(log_wgt)	10	3.69546	3.69546	3.6954565	.00000000

a. Group = 5.53

At the end of the period (Day = 21), differences are larger.

Descriptive Statistics[a]

	N	Minimum	Maximum	Mean	Std. Deviation
log_wgt	16	4.56435	5.72031	5.1285656	.33577946
predicted(log_wgt)	16	5.33884	5.33884	5.3388443	.00000000

a. Group = .00

Descriptive Statistics[a]

	N	Minimum	Maximum	Mean	Std. Deviation
log_wgt	10	4.30407	5.80212	5.2946670	.43720510
predicted(log_wgt)	10	5.51943	5.51943	5.5194271	.00000000

a. Group = 2.71

Descriptive Statistics[a]

	N	Minimum	Maximum	Mean	Std. Deviation
log_wgt	10	4.99043	5.92158	5.5631910	.29523877
predicted(log_wgt)	10	5.62021	5.62021	5.6202106	.00000000

a. Group = 4.34

Descriptive Statistics[a]

	N	Minimum	Maximum	Mean	Std. Deviation
log_wgt	9	5.27811	5.77455	5.4607489	.17411415
predicted(log_wgt)	9	5.69032	5.69032	5.6903227	.00000000

a. Group = 5.53

Regardless of these differences, it seems reasonable to conclude that the Gompertz model describes the data better than either of the previous two models. From the averages given above, we can also conclude that the group on Diet1 (*i.e.*, Group = 0) showed the lowest weight gain.

5. Individual curves

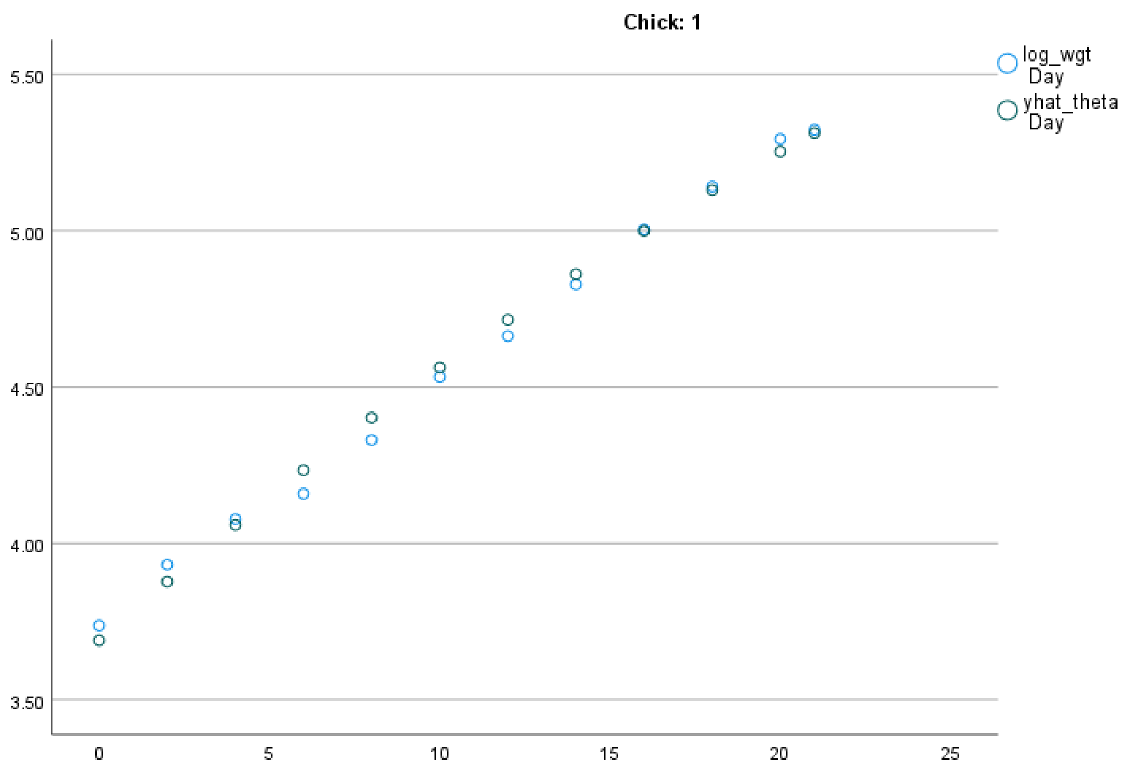
Finally, we look at the individual curves for each chick. As part of the output, information on the estimated parameters for each level-2 unit (in this case each chick) are written to a text file named **thetai.est**.

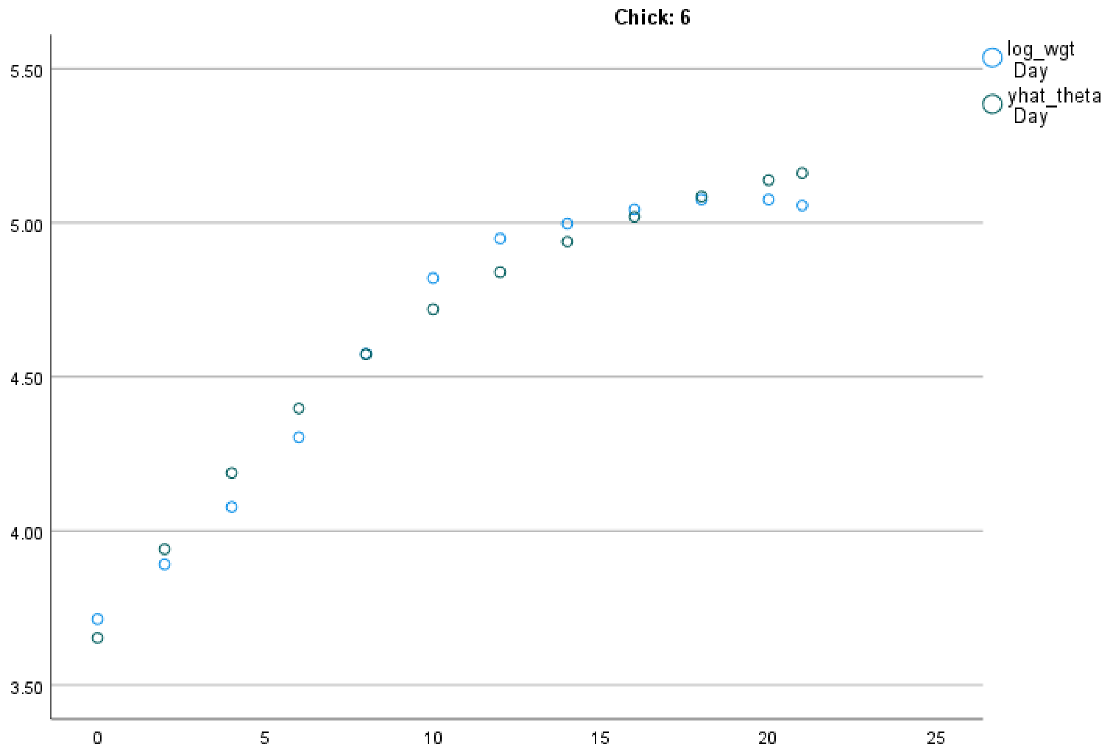
The contents of this file is shown below.

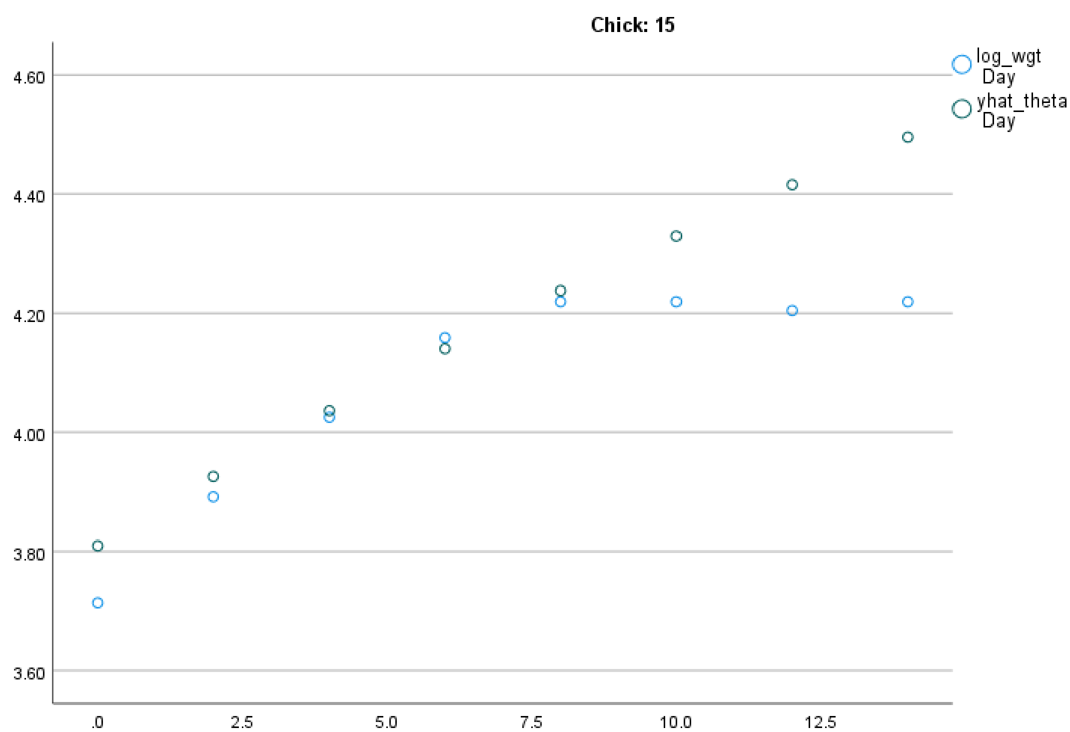
```
5.43966      0.382525     0.627245E-01
5.47698      0.402857     0.120372
7.11405      0.667473     0.410178E-01
5.77105      0.416918     0.349234E-01
6.52092      0.577131     0.675713E-01
5.50050      0.367326     0.428245E-01
6.68330      0.573525     0.114576E-01
5.36337      0.357489     0.724754E-01
6.93967      0.626932     0.296457E-01
6.40994      0.537441     0.309960E-01
6.51440      0.586808     0.836274E-01
5.69048      0.413315     0.666434E-01
5.71970      0.426786     0.731350E-01
5.76958      0.389967     0.197781E-01
```

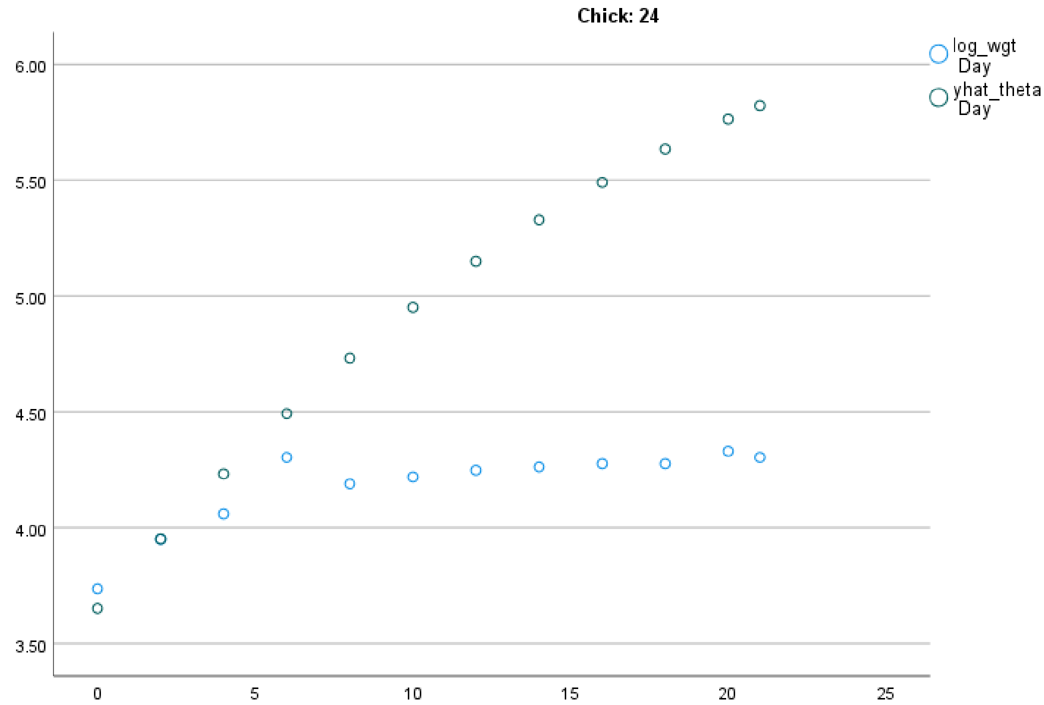
6.65032	0.594303	0.601422E-01
6.83035	0.619306	0.509631E-01
6.32304	0.536660	0.529147E-01
6.47212	0.570830	0.600726E-01
7.41734	0.723262	0.469870E-01
5.51051	0.396977	0.790953E-01
7.02934	0.643836	0.457248E-01
6.82662	0.621412	0.609520E-01
5.23286	0.363208	0.122968
7.08776	0.664186	0.587662E-01
6.88565	0.637731	0.719179E-01
6.26630	0.536733	0.687889E-01
6.79245	0.605968	0.390800E-01
7.22614	0.683506	0.484108E-01
7.02912	0.646785	0.480861E-01
7.01023	0.644499	0.559651E-01
5.64726	0.424703	0.974290E-01
6.61427	0.585100	0.630249E-01
5.50742	0.410701	0.138988
5.28295	0.360426	0.118193
5.73777	0.438858	0.815428E-01
6.23491	0.521265	0.678396E-01
5.76217	0.442358	0.873255E-01
7.11586	0.670596	0.554218E-01
5.96270	0.479120	0.832166E-01
6.44084	0.552226	0.668135E-01

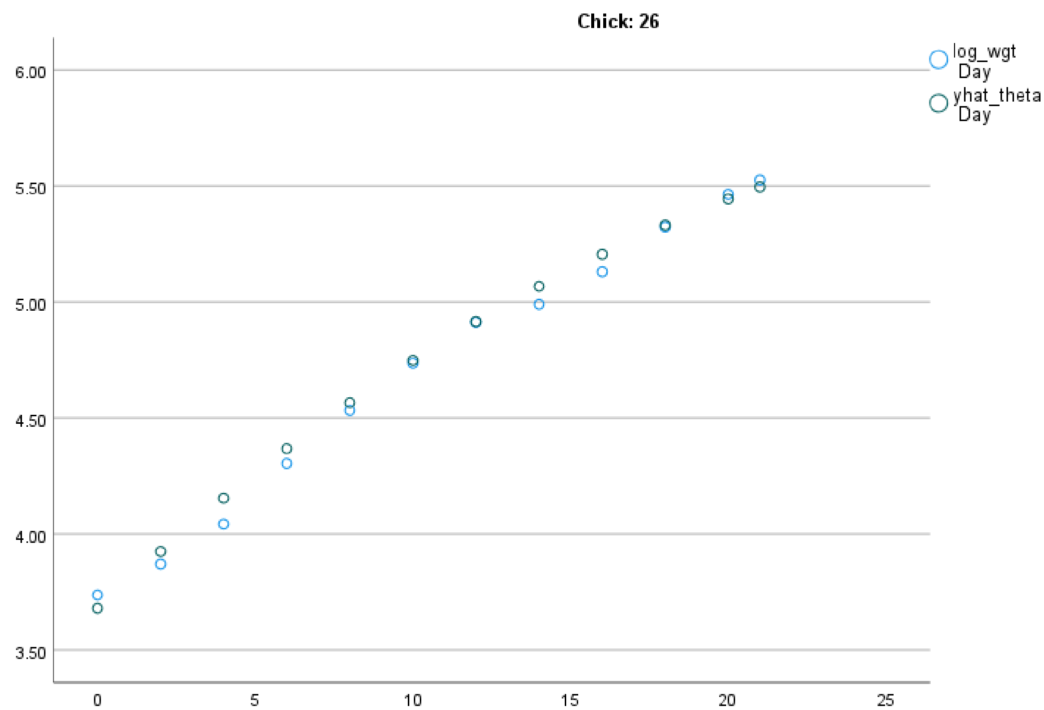
These estimates can be used to get individual curves when used in combination with the observed data. Here we show a few curves for chicks on different diets. In general, we see close agreement between the observed and expected log_wgt over the time of measurement. Exceptions are where chicks either lost weight or stopped picking up weight.

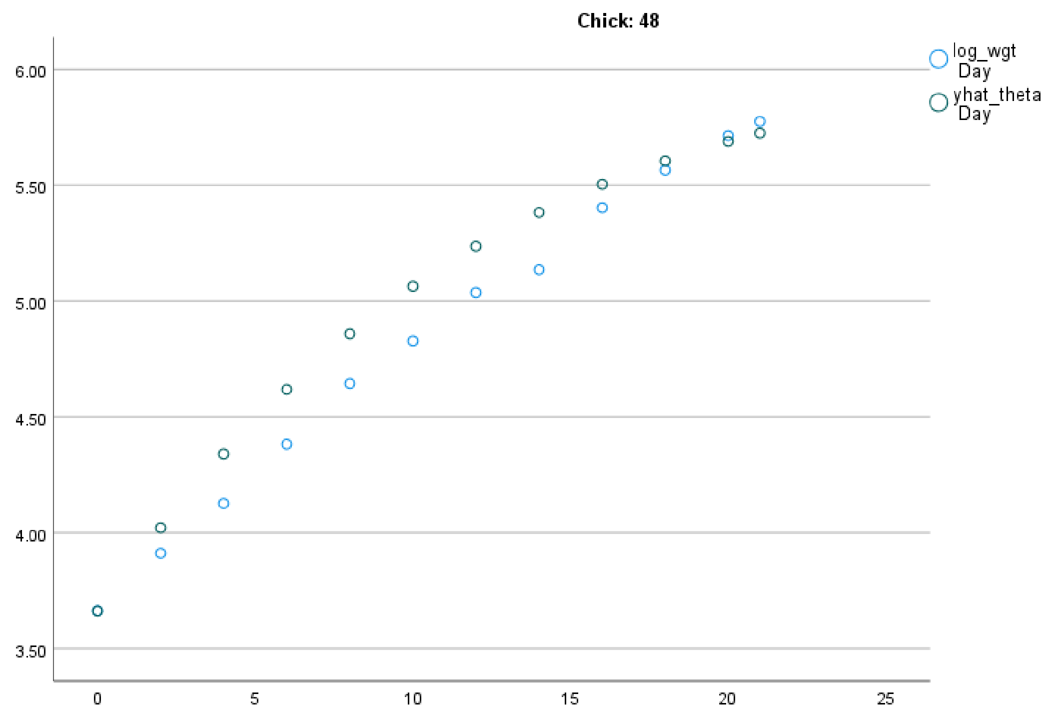



Chick: 6
log_wgt
Day
yhat_theta
Day
5.50
5.00
4.50
4.00
3.50
0
5
10
15
20
25


Chick: 15
log_wgt
Day
yhat_theta
Day
4.60
4.40
4.20
4.00
3.80
3.60
.0
2.5
5.0
7.5
10.0
12.5


Chick: 24
log_wgt
Day
yhat_theta
Day
6.00
5.50
5.00
4.50
4.00
3.50
0
5
10
15
20
25


Chick: 26
log_wgt
Day
yhat_theta
Day
6.00
5.50
5.00
4.50
4.00
3.50
0
5
10
15
20
25


Chick: 48
log_wgt
Day
yhat_theta
Day
6.00
5.50
5.00
4.50
4.00
3.50
0
5
10
15
20
25