

3.	MODELS FOR COUNT OUTCOMES	4
3.1	INTRODUCTION	4
3.1.1	<i>Poisson distribution</i>	4
3.1.2	<i>Negative binomial distribution</i>	7
3.1.3	<i>Adaptive versus non-adaptive quadrature</i>	7
3.2	TWO-LEVEL MODELS FOR COUNT OUTCOMES FROM NESARC DATA	9
3.2.1	<i>The data</i>	9
3.2.1.1	Exploring the data	10
3.2.2	<i>A 2-level Poisson model with 2 predictors</i>	11
3.2.2.1	The model	11
3.2.2.2	Setting up the analysis.....	13
3.2.2.3	Discussion of results.....	15
	Program information and syntax	15
	Data summary	16
	Descriptive statistics	17
	Fixed effects results	17
	Random effects results	19
	Level-1 variation for Poisson distribution.....	19
3.2.2.4	Interpreting the results.....	20
	Estimated outcomes for groups: unit-specific results	20
	Level 2 ICC.....	22
3.2.3	<i>A 2-level negative binomial model with 2 predictors</i>	22
3.2.3.1	The model	22
3.2.3.2	Setting up the analysis.....	23
3.2.3.3	Discussion of results.....	25
	Fixed and random effect results	25
3.2.4	<i>Weighted 2-level models</i>	26
3.2.4.1	The data.....	26
3.2.4.2	Setting up the analysis.....	26
3.2.4.3	Discussion of results.....	27
3.3	TWO-LEVEL MODELS FOR COUNT OUTCOMES FROM ASPART DATA	29
3.3.1	<i>The data</i>	29
3.3.2	<i>A 2 level Poisson model with random intercept</i>	30
3.3.2.1	The model	30
3.3.2.2	Setting up the analysis.....	31
3.3.2.3	Discussion of results.....	33
	Data summary	33
	Descriptive statistics and starting values.....	34
	Fixed and random effect results	35
	Correlation of estimates	36
3.3.2.4	3.3.2.4 Interpreting the results.....	36
	Estimated outcomes for groups: unit-specific results	36
	Estimated outcomes for groups: population-average results	38
3.3.3	<i>A 2-level Poisson log model with an offset variable</i>	40
3.3.3.1	The model	40
3.3.3.2	Setting up the analysis.....	41

3.3.3.3	Discussion of results.....	42
	Fixed and random effect results	42
3.3.3.4	3.3.3.4 Interpreting the results.....	43
	Estimated outcomes for groups: unit-specific results	43
	Level 2 Bayes results	45
	Graphical displays.....	46
REFERENCES		49

3. MODELS FOR COUNT OUTCOMES	4
FIGURE XXX.1: POISSON PROBABILITIES FOR VARIOUS VALUES OF λ	5
FIGURE XXX.2: LOG LINK FUNCTION	6
FIGURE XXX.3: BAR CHART FOR COUNT VARIABLE N_DEP	11
TABLE XXX.1: ESTIMATED NUMBER OF EPISODES UNDER THE POISSON LOG MODEL	21
FIGURE XXX.4: EXPECTED NUMBER OF EPISODES FOR TWO GROUPS	21
TABLE XXX.2: COMPARISON OF RESULTS FOR WEIGHTED AND UNWEIGHTED POISSON MODELS	27
TABLE XXX.2: ESTIMATED UNIT-SPECIFIC AND POPULATION AVERAGE RESULTS FOR GROUPS.....	39
TABLE XXX.3: COMPARISON OF RESULTS FOR POISSON MODELS	44
FIGURE XXX.5: PREDICTED AVERAGE NUMBER OF HEADACHES FOR PLACEBO AND ASPARTAME.....	46
FIGURE XXX.6: FITTED AND OBSERVED TRAJECTORIES	48
REFERENCES	49

3. Models for count outcomes

3.1 Introduction

A count variable counts the number of discrete occurrences of a characteristic of interest that takes place during a time interval. Examples are the occurrence of cancer cases in a hospital during a given period of time, the number of cars that pass through a toll station per day, and the phone calls at a call center.

The most common distribution for a count variable is the Poisson distribution. Besides the Poisson distribution, negative binomial distributions may also be used to describe the properties of count variables. In this chapter, models for count data, utilizing both the Poisson and negative binomial distributions, are discussed. For further information regarding these distributions, please refer to chapter XXX.

3.1.1 Poisson distribution

The Poisson distribution is a discrete probability distribution. It is appropriate for expressing the probability of a number of events occurring in a fixed time period with a known average rate, under the assumption that the occurrences are independent of one another.

The probability of k occurrences can be expressed as

$$f(k; \lambda) = \frac{e^{-\lambda} \lambda^k}{k!} \text{ for } k = 0, 1, 2, \dots$$

where k is a non-negative integer and λ is a positive real number, which equals the expected number of occurrences during the given interval.

The cumulative probability function is

$$\Pr(k; \lambda) = \sum_{i=0}^k \frac{e^{-\lambda} \lambda^i}{i!} \quad \text{for } k = 0, 1, 2, \dots,$$

with the single parameter λ . A Poisson distribution has an important property: the mean number of occurrences λ is equal to the variance: $E(f) = \text{var}(f) = \lambda$. Figure XXX.1 shows Poisson probabilities $f(k)$ and cumulative probabilities $g(k)$ for $\lambda = 0.5, 2$ and 5 .

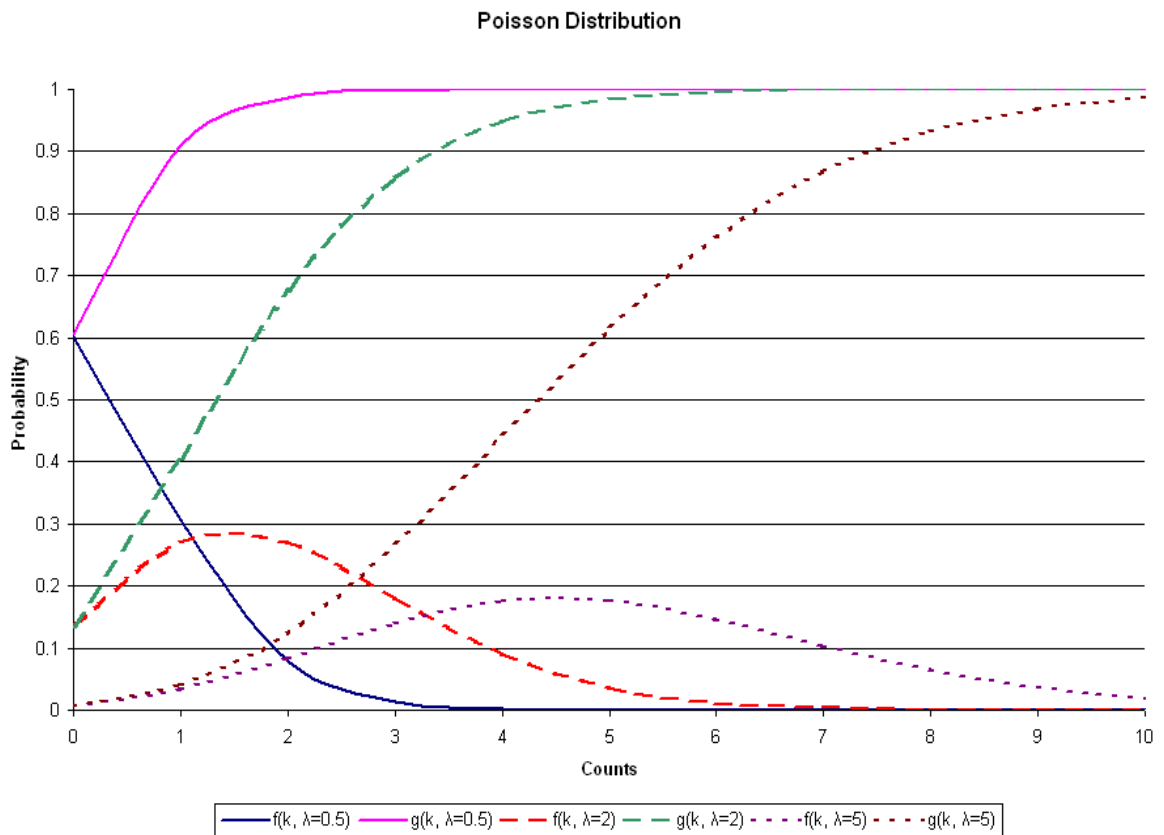


Figure XXX.1: Poisson probabilities for various values of λ

As shown above, the smaller λ is, the more skewed to the right the probability distribution is. When λ is large, the Poisson distribution is close to the normal distribution.

The log link function is generally used for the Poisson distribution. Assume the response measurements for a count variable y_1, \dots, y_n are independent and

$$y_i \sim Poi(\lambda_i), \quad \text{where} \quad \lambda_i = e^{\beta_1 x_{i1} + \dots + \beta_p x_{ip}}$$

The natural logarithm of the above equation is used to define the link function:

$$\log(\lambda_i) = \beta_1 x_{i1} + \dots + \beta_p x_{ip}$$

As shown in Figure XXX.2, using the log link function maps the mean of the count variable λ with an open interval $(0, +\infty)$ to the set of real numbers $(-\infty, +\infty)$.

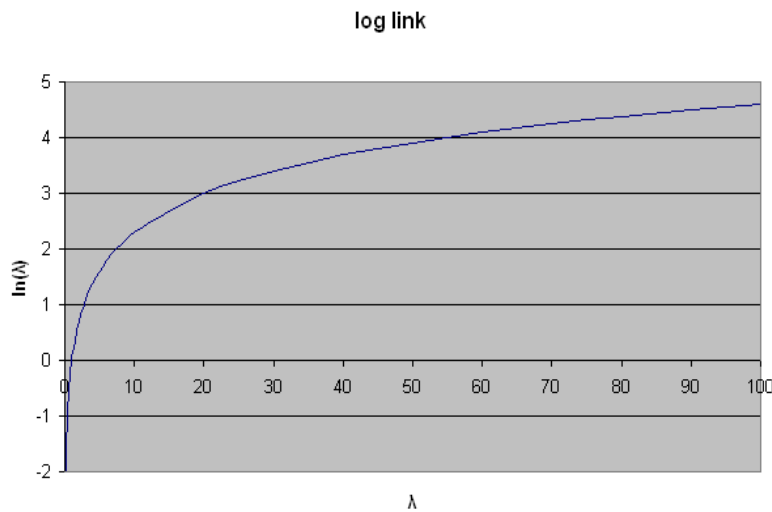


Figure XXX.2: Log link function

3.1.2 Negative binomial distribution

The negative binomial distribution is a probability distribution used to describe a certain number of failures and successes in a series of independent and identically distributed Bernoulli trials. Specifically, for $k + r$ Bernoulli trials with success probability p , the negative binomial gives the probability of k failures and r successes, with success on the last trial. In other words, the negative binomial distribution is the probability distribution of the number of failures before the r^{th} success in a Bernoulli process, with probability p of success on each trial.

The negative binomial distribution can be expressed as

$$f(y_i) = \frac{\Gamma(y_i + 1/\alpha)}{\Gamma(y_i + 1)\Gamma(1/\alpha)} \times \frac{(\alpha\mu_i)^{y_i}}{(1 + \alpha\mu_i)^{y_i + 1/\alpha}}$$

with $\Sigma(y_i) = \mu_i + \alpha\mu_i^2$, where $\Gamma(\)$ is the gamma function or generalized factorial from advanced calculus, and where α denotes an additional parameter and it can no longer be assumed that the variance is a known function of the mean. In the example to follow, α is assumed to have a fixed value.

3.1.3 Adaptive versus non-adaptive quadrature

Ordinary quadrature is a numeric method for evaluating multi-dimensional integrals. For mixed-effect models with count and categorical outcomes, the log-likelihood function is expressed as the sum of the logarithm of integrals, where the summation is over higher-level units, and the dimensionality of the integrals equals the number of random effects.

A problem with ordinary quadrature is that it assumes a common location and scale for each level-2 unit. This assumption often requires the use of a large number of quadrature points to calculate the log-likelihood and derivatives to an acceptable

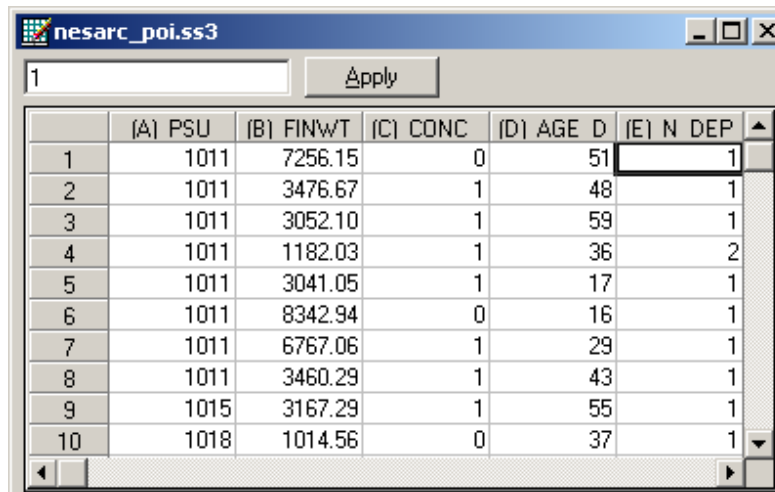
level of accuracy. To overcome this problem with ordinary quadrature, SuperMix also offers a numeric integration procedure called adaptive quadrature. The adaptive quadrature procedure uses the empirical Bayes means and covariances, updated at each iteration to essentially shift and scale the quadrature locations of each higher-level unit in order to place them under the peak of the corresponding integral. To distinguish between the two quadrature methods, SuperMix uses the terminology non-adaptive quadrature (ordinary quadrature) and adaptive quadrature. To illustrate this, models in Section 3.2 will be fitted using adaptive quadrature, while models in Section 3.3 will use the default method of non-adaptive quadrature.

3.2 Two-level models for count outcomes from NESARC data

3.2.1 The data

The data set is from the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC), which was designed to be a longitudinal survey with its first wave fielded in 2001-2002. This data file has been used in some of the examples in Chapter XXX. Detailed information about the survey is available at <http://niaaa.census.gov/index.html>.

In Section 4 of the description of the NESARC study, information on the data regarding occurrences of major depression, family history of major depression and dysthymia are collected. This information was used, in combination with the demographic information provided in Section 1 of the study description, to produce the `nesarc_poi.ss3` data set used in this section. The image below shows the first ten records of this data set. There are 2339 dysthymia respondents in the survey; after listwise deletion, the sample size is 1981.



	(A) PSU	(B) FINWT	(C) CONC	(D) AGE D	(E) N DEP
1	1011	7256.15	0	51	1
2	1011	3476.67	1	48	1
3	1011	3052.10	1	59	1
4	1011	1182.03	1	36	2
5	1011	3041.05	1	17	1
6	1011	8342.94	0	16	1
7	1011	6767.06	1	29	1
8	1011	3460.29	1	43	1
9	1015	3167.29	1	55	1
10	1018	1014.56	0	37	1

The variables of interest are:

- PSU denotes the Census 2000/2001 Supplementary Survey (C2SS) primary sampling unit.
- FINWT represents the NESARC weights sample results used to form national level estimates. The final weight is the product of the NESARC base weight and other individual weighting factors.
- CONC_DEP contains the information captured in field S4CQ3A6 of the NESARC data. It represents the response to the statement "Often had trouble concentrating/keeping mind on things," with 1 indicating "Yes," and 0 indicating "No."
- AGE_DEP is based on field S4CQ7AR of the NESARC data. It represents the age at onset of first episode.
- N_DEP is recoded from field S4CQ6A of the NESARC data, and gives the number of depression/dysthymia episodes. This is the count variable we would like to use as outcome variable in the examples to follow.

3.2.1.1 Exploring the data

Inspecting the distribution of the intended outcome variable, N_DEP, before starting with the model is important. In the case of a count variable, this can easily be done by producing a bar chart of the observed frequencies of occurrence captured by the count variable. Select the **File, Data-based Graph, Univariate** option from the main SuperMix window and request a bar chart before clicking the **Plot** button.

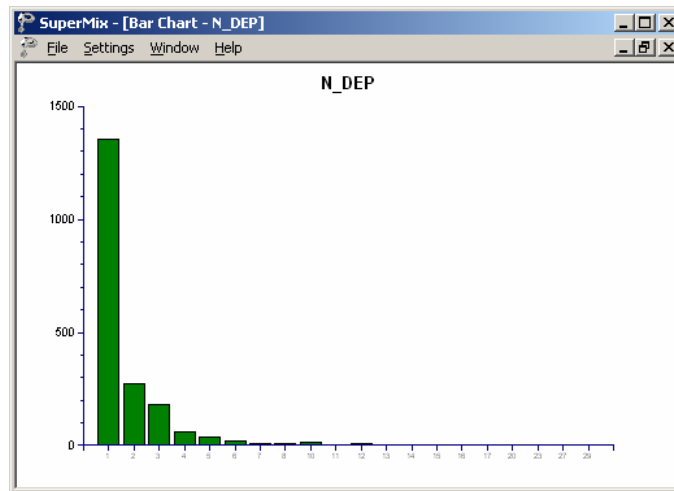


Figure XXX.3: Bar chart for count variable N_DEP

The frequency bar chart for the count variable N_DEP shown in Figure XXX.3 is obtained. We note that the number of depression episode ranges from 1 to 29, with most respondents having a small number of reported episodes of depression.

3.2.2 A 2-level Poisson model with 2 predictors

3.2.2.1 The model

The first model fitted to the data explores the relationship between N_DEP and the variables indicating concentration (or lack thereof) and age, as represented by the variables CONC_DEP and AGE_DEP.

The level-1 model is

$$\log(\lambda_{ij}) = \beta_0 + \beta_1 \times \text{CONC_DEP}_{ij} + \beta_2 \times \text{AGE_DEP}_{ij}$$

where the expected number of depression episodes is $\lambda_{ij} = E(N_DEP_{ij})$.

The level-2 model is

$$\beta_0 = b_{00} + v_{i0}, \beta_1 = b_{10} \text{ and } \beta_2 = b_{20}.$$

Another way of writing the combined model is

$$\log(\lambda_{ij}) = b_{00} + b_{10} \times \text{CONC_DEP}_{ij} + b_{20} \times \text{AGE_DEP}_{ij} + v_{i0}.$$

In this model, $e^{b_{00}}$ denotes the average expected count of depression episodes, and b_{10} represents the estimated coefficient for the respondent's level of concentration.

Taking exponents on both sides, we also have

$$\begin{aligned} \lambda_{ij} &= e^{b_{00} + b_{10} \times \text{CONC_DEP}_{ij} + b_{20} \times \text{AGE_DEP}_{ij} + v_{i0}} \\ &= e^{b_{00}} e^{b_{10} \times \text{CONC_DEP}_{ij}} e^{b_{20} \times \text{AGE_DEP}_{ij}} e^{v_{i0}} \end{aligned}$$

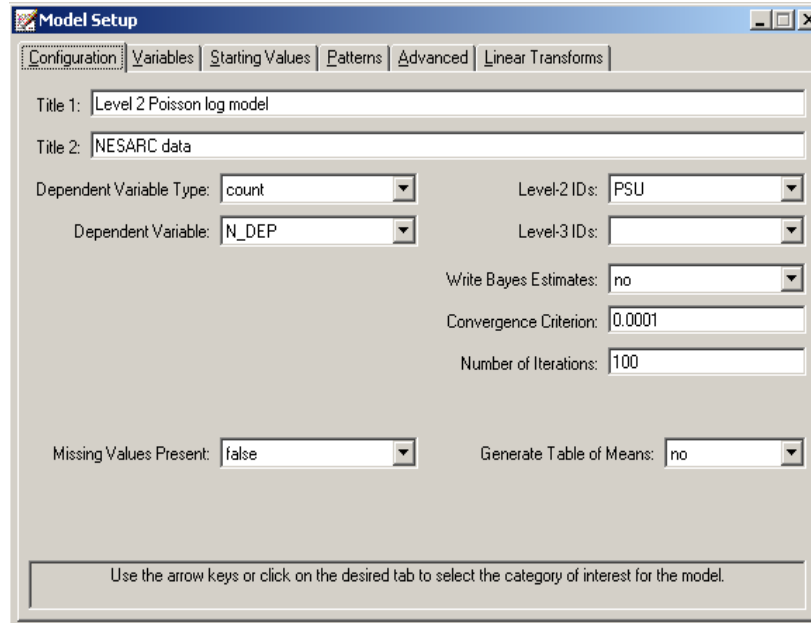
For a person who had problems concentrating ($\text{CONC_DEP} = 1$), the expected average number of episodes $e^{b_{00}}$ is multiplied by $e^{b_{10}}$, compared to an expected count of $e^{b_{00}}$ for a person for whom $\text{CONC_DEP} = 0$. Similarly, an increase of one year in age increases the estimated number of episodes by a factor of $e^{b_{20}}$. For example, a respondent with concentration problems who is two years older than another respondent who had no concentration problems is expected to have $e^{b_{00}} e^{b_{10}} e^{2b_{20}}$ episodes compared to only $e^{b_{00}}$ episodes for the younger person without concentration problems.

The random part of the model is represented by $e^{v_{i0}}$, which denotes the variation in average count of depression episodes over PSU and between respondents (or, in other words, over respondents nested within PSU). For a Poisson distribution, the assumption of normality at level 1 is not realistic, as the level-1 random effect can

only assume a number of distinct values. Thus, this random effect cannot have homogeneous variance.

3.2.2.2 Setting up the analysis

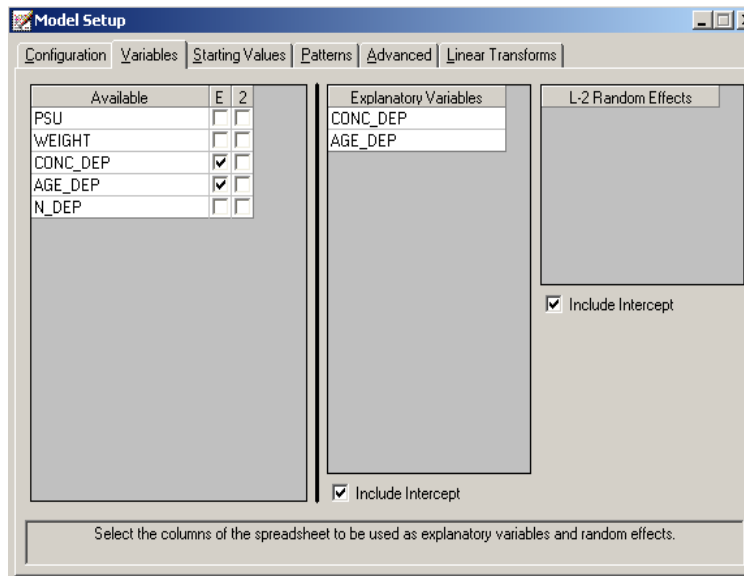
Open the SuperMix spreadsheet `nesarc_poi.ss3` used during the exploratory analysis. From the main menu bar, select the **File, New Model Setup** option. The **Model Setup** window that appears has six tabs. In this example, only three tabs are used: the **Configuration**, **Variables**, and **Advanced** tabs.



The **Configuration** screen is the first tab on the **Model Setup** window. It enables the user to define the outcome variable and the level-2 and level-3 IDs. Some other settings such as missing values, convergence criterion, number of iterations, etc. can be specified here. For all the available settings, please refer to chapter XXXX. To obtain the model we discussed, start by selecting the outcome variable `N_DEP` from the **Dependent Variable** drop-down list box. Indicate that it is a count variable by selecting the **count** option from the **Dependent Variable Type** drop-down list box. Next, describe the hierarchical structure of the data by selecting the level-2 ID, `PSU`,

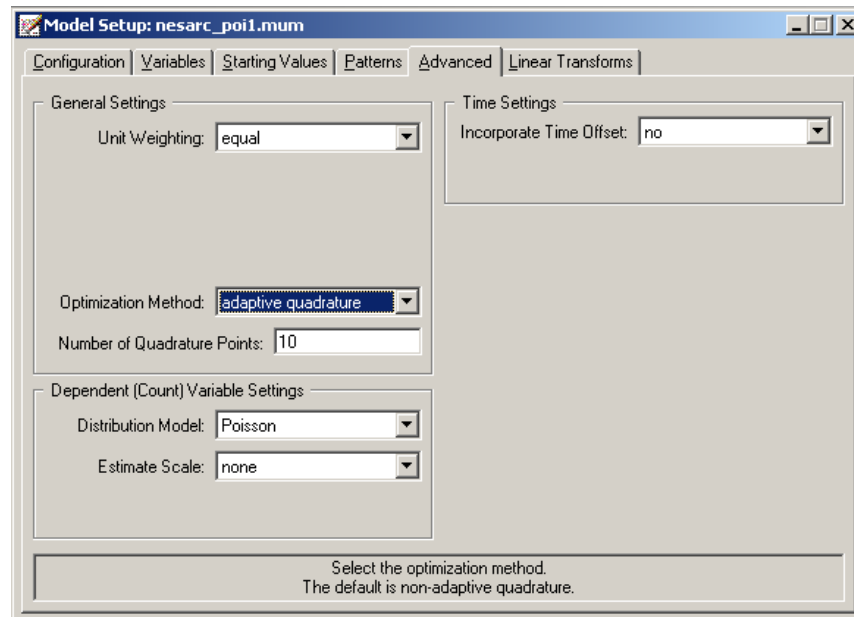
from the **Level-2 IDs** drop-down list box. Enter a title in the **Title** text boxes, and proceed to the **Variables** screen by clicking on this tab.

The **Variables** screen is used to specify the fixed and random effects to be included in the model. To include the variables CONC_DEP and AGE_DEP as predictor variables, check the **E** check boxes next to the variables' names. Note that, as the variables are selected, the selected variables are listed in the **Explanatory Variables** grid. After selection, the screen below is obtained. Note that the **Include Intercept** check boxes in the **Explanatory Variables** grid and **L-2 Random Effects** are checked by default, indicating that an intercept term will automatically be included in the fixed and random parts of the model.



As mentioned previously, the models based on the NESARC data are fitted using adaptive rather than full quadrature. The final step of model specification is to request the use of adaptive quadrature, and this is done on the **Advanced** tab.

Click on the **Advanced** tab, and request **adaptive quadrature** using the **Optimization Method** drop-down list box. Do not change the number of quadrature points from the default displayed (10).



Before running the analysis, the model specifications have to be saved. Select the **File, Save As** option, and provide a name (**nesarc_poi1.mum**) for the model specification file. Run the analysis by selecting the **Run** option from the **Analysis** menu.

3.2.2.3 Discussion of results

Portions of the output file **nesarc_poi.out** are shown below.

Program information and syntax

As shown below, the syntax for the model setup is printed in the output file. The first line of the syntax shows the option **Model = Count**, which indicates the outcome variable is a count variable. The **Options** syntax line corresponds to the settings on

the **Configuration** screen. The Link = log and Distribution = Poi options specify the use of a Poisson distribution with a log link function for the fitted model.

```

nesarc_poi1.out

The following lines were read from file C:\Program Files\SuperMix\count\nesarc_poi1.inp

Model=Count;
Options Converge=0.0001 Maxiter=100 Bayes=No NQuadPTS=-10;
Link=log;
Distribution=Poi;
Scale=none;
Varnames= PSU WEIGHT CONC_DEP AGE_DEP N_DEP intcept;
Title1=Level 2 Poisson log model;
Title2=NESARC data;
DataFile=C:\Program Files\SuperMix\count\nesarc_poi1.dat;
Level2ID= PSU;
Dependent= N_DEP;
Predictors= intcept CONC_DEP AGE_DEP;
L2Random= intcept;
FixPatType=Free;
Cov2PatType=Correlated;

Save As... Close

```

Data summary

```

nesarc_poi1.out

Model and Data Descriptions

Sampling Distribution          = Poisson
Link Function                  = Log
Number of Level-2 Units       = 395
Number of Level-1 Units       = 1981
Number of Level-1 Units per Level-2 Unit =
8  1  1  5  5  5  1  5  1  5  2  5
3  31 16  3  1  7  5  3  2  1  9  1
7  6  3  22 8  1  8  2  1  1  1  2
5  51 8  25 8 10  4  1  4  4 10  2
19 7  5  2  2 10  7  3  1  6  6  1
3  8  4  3 10  2  4  2  1  6  2  1
16 18 5  3  7  3  1  6  4  8  5  3
2  20 8  1  3  4  2  3  1  7  1  2
4  8  3  4  9  1  1  33 2  7  2  2
3  3  3  3 11  5  6  1  7  2  5  2

Save As... Close

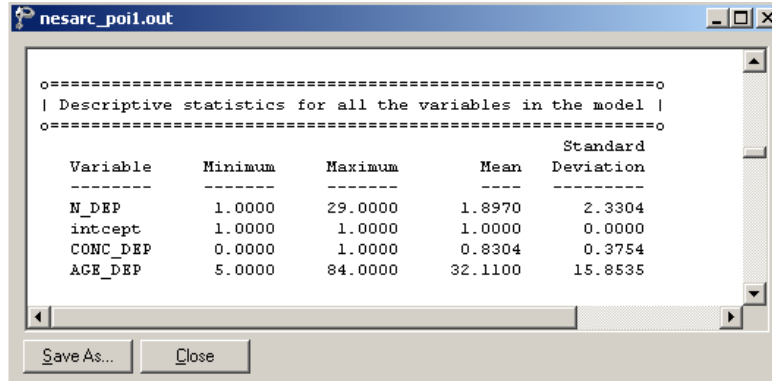
```

A description of the hierarchical structure follows the syntax: data from a total of 395 PSU and 1981 respondents were included at levels 2 and 1 of the model. In

addition, an enumeration of the number of respondents nested within each of the 395 PSUs is provided.

Descriptive statistics

The data summary is followed by descriptive statistics for all the variables included in the model. The mean of 1.8970 and standard deviation of 2.3304 are reported for the outcome N_DEP indicating that, on average, 1.8970 episodes of depression were recorded.



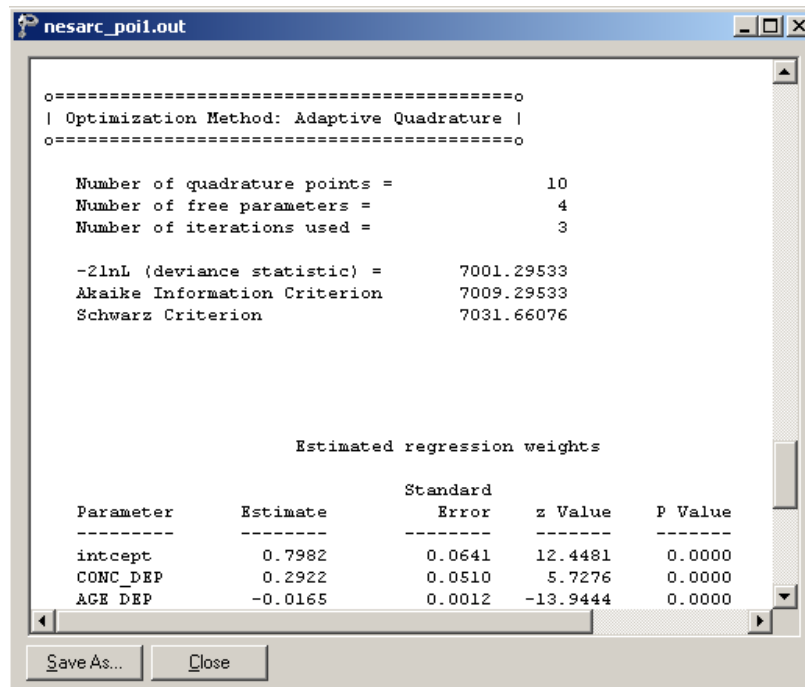
```
o=====o
| Descriptive statistics for all the variables in the model |
o=====o

Variable      Minimum      Maximum      Mean      Standard
-----      -
N_DEP         1.0000      29.0000      1.8970      2.3304
intcept       1.0000      1.0000      1.0000      0.0000
CONC_DEP      0.0000      1.0000      0.8304      0.3754
AGE_DEP       5.0000      84.0000      32.1100     15.8535
```

Descriptive statistics are followed by the results for a fixed-effects-only model, i.e. a model without random coefficients.

Fixed effects results

At the top of the final results, the number of iterations required for convergence of the iterative procedure is given. Next, the number of quadrature points per dimension is reported which, in this case, is the default number of points. The log likelihood and the deviance, which is defined as $-2\ln L$, are listed next. For a pair of nested models, the difference in $-2\ln L$ values has a χ^2 distribution, with degrees of freedom equal to the difference in number of parameters estimated in the models compared.

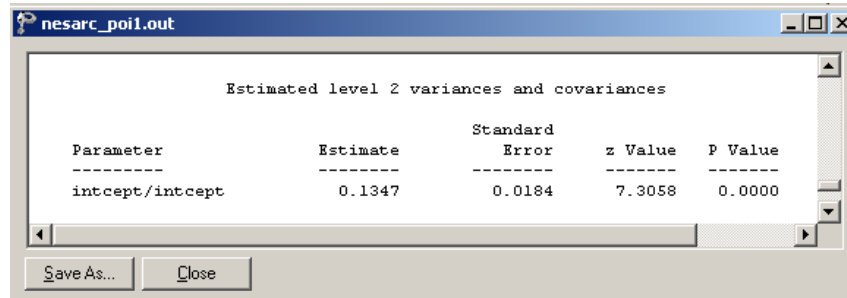


The estimated intercept is 0.7982, which means that the average number of depression episodes is $e^{0.7982}=2.2215$, implying that on average the number of episodes is about two. The estimated coefficient for CONC_DEP is 0.2922, which indicates that respondents who had trouble concentrating on things tended to have $2.2215e^{0.2922}=(2.2215)(1.3394)=2.9754$ episodes at the same age as respondents without concentration problems. The estimate of the effect of age at the onset of the first episode (AGE_DEP) shows that the onset age does not affect the number of episodes much, since $e^{-0.0165}=0.98$. A slight reduction in the expected number of episodes is expected with increasing age. If one compares two typical respondents with reported concentration problems, but with one respondent ten years older than the other, one would expect the older respondent to have $(2.2215)(1.3394)e^{10(-0.0165)}=2.5229$ episodes, compared to 2.9268 expected episodes for the younger respondent. In other words, the longer it takes for the first episode to occur, the fewer episodes a respondent is expected to have. Of course, it has to be kept in mind that the younger a respondent is at the first episode, the longer that

person must live with the condition and thus the more time there is for subsequent episodes to occur.

Random effects results

The output for the level-2 random effect variance term follows next. The estimated variation in the average estimated N_DEP at level 2 is 0.1347, which is highly significant. Respondents are different in terms of their average expected number of episodes, holding all other variables constant.



Parameter	Estimate	Standard Error	z Value	P Value
intcept/intcept	0.1347	0.0184	7.3058	0.0000

Level-1 variation for Poisson distribution

The variance-to-mean ratio is a measure of the dispersion of a probability distribution:

$$R = \text{variance-to-mean ratio} = \frac{\sigma^2}{\mu}$$

For the Poisson distribution, where the variance equals the mean, this implies $R = 1$. Thus, we use a value of one as our level-1 variation. In the cases when over-dispersion ($R > 1$) or under-dispersion ($R < 1$) is assumed, different level-1 variation values will apply. The details of these scenarios are not discussed in this guide. Please refer to XXXXX for more details.

3.2.2.4 Interpreting the results

Estimated outcomes for groups: unit-specific results

First, we substitute the regression weights and obtain the following function for $\log(\widehat{N_DEP}_{ij})$:

$$\begin{aligned}\log(\widehat{N_DEP}_{ij}) &= \hat{b}_{00} + \hat{b}_{10} \times \text{CONC_DEP}_{ij} + \hat{b}_{20} \times \text{AGE_DEP}_{ij} \\ &= 0.7982 + 0.2922 \times \text{CONC_DEP}_{ij} - 0.0165 \times \text{AGE_DEP}_{ij}.\end{aligned}$$

For example, at age 40, the estimated $\log(\widehat{N_DEP}_{ij})$ for a typical respondent who does not often have trouble concentrating ($\text{CONC_DEP} = 0$), we find that

$$\begin{aligned}\log(\widehat{N_DEP}_{ij}) &= \hat{b}_{00} + \hat{\beta}_1 \times \text{CONC_DEP}_{ij} + \hat{\beta}_2 \times \text{AGE_DEP}_{ij} \\ &= 0.7982 + 0.2922 \times \text{CONC_DEP}_{ij} - 0.0165 \times \text{AGE_DEP}_{ij} \\ &= 0.7982 + 0.2922 \times 0 - 0.0165 \times 40 \\ &= 0.1382.\end{aligned}$$

Keeping in mind that we defined the relationship between λ and the predictors as

$$\log(\lambda_{ij}) = \beta_1 x_{i1} + \dots + \beta_p x_{ip},$$

it follows that

$$\hat{\lambda}_{ij} = e^{0.1382} = 1.1482.$$

We can estimate the count of the occurrence of depression episodes for typical individuals of different ages in the same way. Results are summarized in Table XXX.1. The results show a decrease in the expected number of episodes with increasing age, regardless of whether they had concentration problems or not.

Table XXX.1: Estimated number of episodes under the Poisson log model

AGE_DEP	10	20	30	40	50	60	70
CONC_DEP = 1	2.5229	2.1391	1.8138	1.5379	1.3040	1.1056	0.9374
CONC_DEP = 0	1.8836	1.5971	1.3542	1.1482	0.9736	0.8255	0.6999

The results in Table XXX.1 can also be presented graphically, as shown in Figure XXX.4. We clearly see that the correspondents who often had trouble concentrating (CONC_DEP = 1) have a higher estimated number of depression episodes. It also shows that the number of episodes is expected to decrease as people get older.

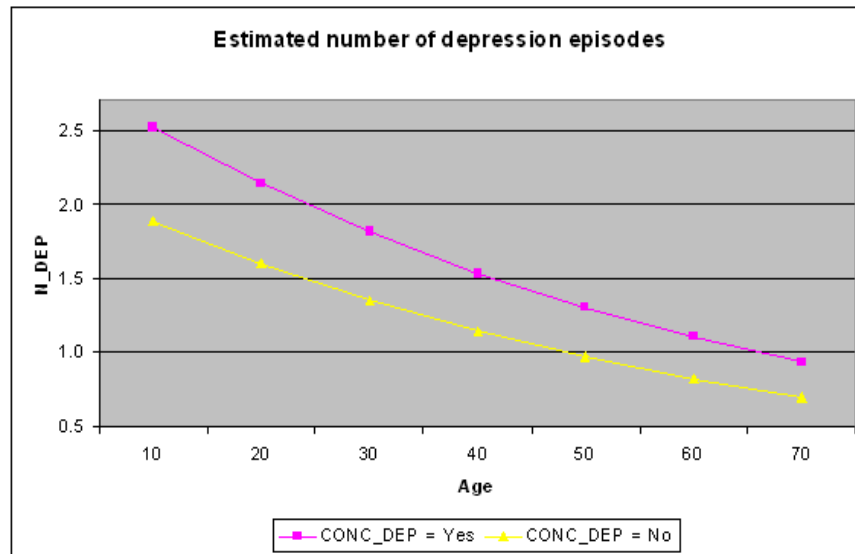


Figure XXX.4: Expected number of episodes for two groups

Level 2 ICC

The percentage of variance explained over level-2 units, or intraclass correlation coefficient (ICC), is calculated as

$$ICC = \frac{\text{level-2 variation}}{\text{level-1 variation} + \text{level-2 variation}}$$

In this example, under the assumption that the level-1 variation is fixed at a value of one, we have

$$ICC = \frac{0.1347}{1 + 0.1347} \times 100\% = 11.8\%$$

We can conclude that most of the unexplained variation in the outcome (approximately 78%) is between measurements at the lowest level of the hierarchy.

3.2.3 A 2-level negative binomial model with 2 predictors

3.2.3.1 The model

In Section 3.2.2, a Poisson model was fitted to the data. It was also noted that a Poisson distribution has an important property: the mean number of occurrences is equal to the variance. The negative binomial distribution is an alternative distribution that may also be used to describe the properties of count variables. If the assumption of a Poisson distribution is reasonable, one would expect a model using a negative binomial distribution with a very small dispersion parameter to produce results that correspond closely to those obtained for the Poisson model. In this section, we fit a negative binomial model, utilizing the same predictors and a small dispersion parameter, to the NESARC data. Again, adaptive quadrature is used as the method of optimization.

Recall that the negative binomial distribution can be expressed as

$$f(y_i) = \frac{\Gamma(y_i + 1/\alpha)}{\Gamma(y_i + 1)\Gamma(1/\alpha)} \times \frac{(\alpha\mu_i)^{y_i}}{(1 + \alpha\mu_i)^{y_i + 1/\alpha}}$$

with $\Sigma(y_i) = \mu_i + \alpha\mu_i^2$ where α denotes an additional parameter and it can no longer be assumed that the variance is a known function of the mean. We assume α to be a fixed parameter.

The model fitted to the data explores the relationship between N_DEP and the variables indicating concentration (or lack thereof) and age, as represented by the variables CONC_DEP and AGE_DEP.

The level-1 model is

$$\log[E(N_DEP_{ij})] = \beta_0 + \beta_1 \times \text{CONC_DEP}_{ij} + \beta_2 \times \text{AGE_DEP}_{ij}$$

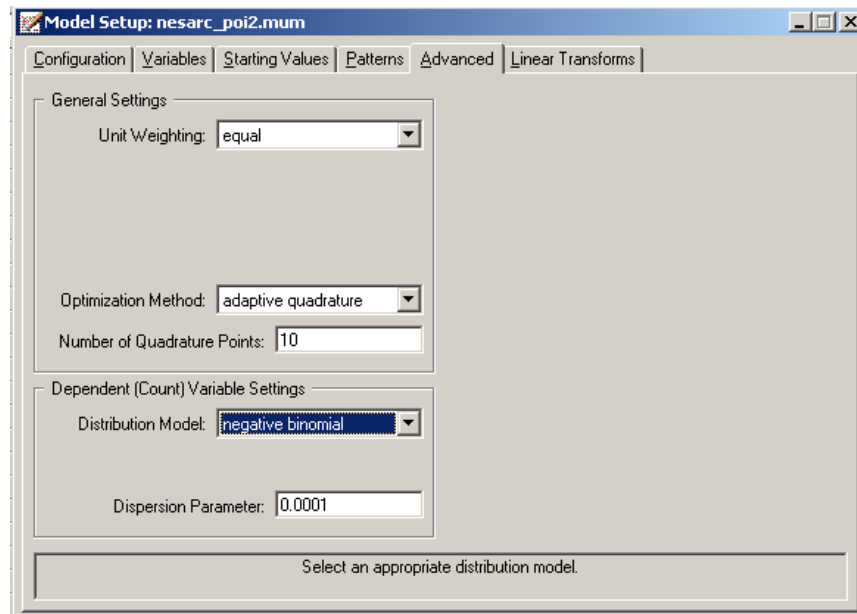
The level-2 model is

$$\beta_0 = b_{00} + v_{i0}, \beta_1 = b_{10} \text{ and } \beta_2 = b_{20}.$$

3.2.3.2 Setting up the analysis

Using the SuperMix spreadsheet **nesarc_poi.ss3** and model specification file **nesarc_poi1.mum** from the previous section, we now set up a negative binomial model for these data.

Start by saving the model specification file under the new name **nesarc_poi2.mum** using the **File, Save As** option. Next, click on the **Advanced** tab of the **Model Setup** window. This is the only tab on which modifications have to be made to change the previously specified Poisson model to a negative binomial model. Set the **Distribution Model** to **negative binomial**, and the **Dispersion Parameter** to 0.0001 to obtain an **Advanced** tab as shown below. Also, request **adaptive quadrature** using the **Optimization Method** drop-down list box. Do not change the number of quadrature points from the default displayed (10).



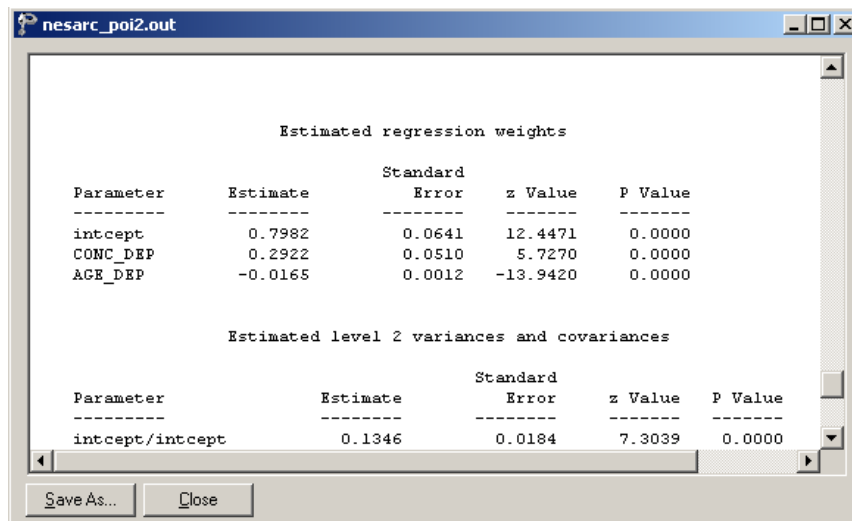
Save the revised model specification file, and click the **Analysis, Run** option to start the iterative process.

3.2.3.3 Discussion of results

Portions of the output file `nesarc_poi.out` are shown below.

Fixed and random effect results

The estimated regression coefficients for fixed effects in the model are shown below. Recall that the estimated coefficients of the intercept, `CONC_DEP`, and `AGE_DEP` under the Poisson model in Section 3.2.2 were 0.7982, 0.2922, and -0.0165 respectively. The estimated variation in the average estimated `N_DEP` at level-2 was 0.1347, and highly significant. The similarity of the results obtained under these two models indicate that the specification of a Poisson distribution model is reasonable for this data.



```
nesarc_poi2.out
```

Estimated regression weights

Parameter	Estimate	Standard Error	z Value	P Value
intcept	0.7982	0.0641	12.4471	0.0000
CONC_DEP	0.2922	0.0510	5.7270	0.0000
AGE_DEP	-0.0165	0.0012	-13.9420	0.0000

Estimated level 2 variances and covariances

Parameter	Estimate	Standard Error	z Value	P Value
intcept/intcept	0.1346	0.0184	7.3039	0.0000

Save As... Close

3.2.4 Weighted 2-level models

3.2.4.1 The data

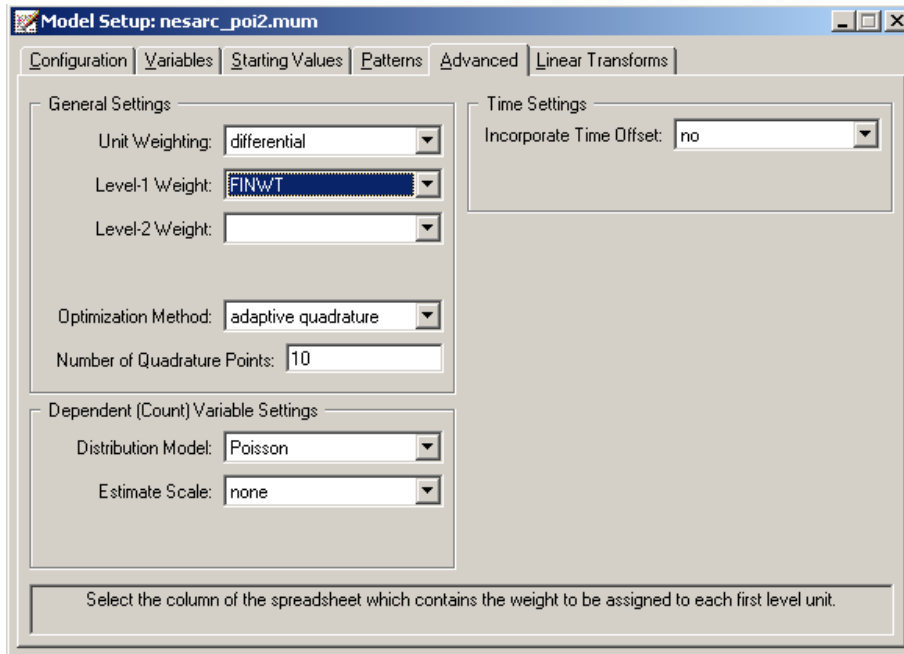
The sampling frame of many multistage surveys frequently entails selection of units with known, but unequal, selection probabilities. This situation is the result of a number of design factors, of which the cost of doing the survey is an important consideration. When this is the case, it is appropriate to weight observations in order to produce unbiased estimates of population parameters.

Recall from Section 3.2.1 that the data also included a weight variable. The variable FINWT represents the NESARC weights sample results used to form national-level estimates. The final weight is the product of the NESARC base weight and other individual weighting factors. In this section, we explore the effect of inclusion of the weights on the results obtained in Sections 3.2.2 and 3.2.3.

3.2.4.2 Setting up the analysis

The models remain the same, with only the selection of the weight variable on the **Advanced** tab of the **Model Specification** screen to be added. Below, we show how this is done in the case of the Poisson distribution model.

Open the model specification file for the Poisson distribution model (`nesarc_poi1.mum`) and click on the **Advanced** tab. Change the **Unit Weighting** field from its default value of equal to differential. Next, select the variable FINWT from the **Assigned Weight** drop-down list box that appears when the selection has been made in the **Unit Weighting** field. The completed **Advanced** tab is shown below.



Save the specification file as **nesarc_poi3.mum**, and run the analysis.

3.2.4.3 Discussion of results

Results for this analysis are reported in Table XXX.2 below. The results from the unweighted Poisson distribution model are included in order to facilitate evaluation of the impact of the weights on the results.

Table XXX.2: Comparison of results for weighted and unweighted Poisson models

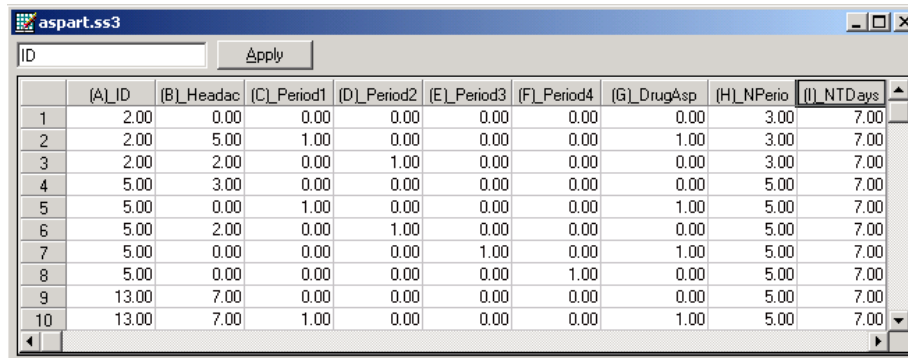
Parameter	Unweighted model		Weighted model	
	Estimate	Standard error	Estimate	Standard error
intcept	0.7982	0.0641	0.7225	0.0660
CONC_DEP	0.2922	0.0510	0.3055	0.0532
AGE_DEP	-0.0165	0.0012	-0.0156	0.0013
Level-2 variance	0.1347	0.0184	0.1378	0.1089

Results for the two models are very similar, and interpretation of the results of both models would lead to the same conclusions, both in terms of significance and in terms of the expected number of depression episodes. However, this is more the exception than the rule – users are cautioned to use weight variables whenever they are available in order to prevent skewed or biased results that may occur when weights are excluded in the analysis of a disproportionately drawn sample.

3.3 Two-level models for count outcomes from ASPART data

3.3.1 The data

The data for this example are taken from a paper by McKnight and Van Den Eeden (1993), who reported on the number of headaches in a two treatment, multiple period crossover trial. Specifically, the number of headaches per week was repeatedly measured for 27 patients. Following a seven day placebo run-in period, subjects received either aspartame or placebo in four seven-day treatment periods according to the double-blind crossover treatment design. Each treatment period was separated by a washout day. The sample size is 122. Data for the first 10 observations of all the variables used in this section are shown below in the form of a SuperMix spreadsheet window for **aspart.ss3**.



	(A)_ID	(B)_Headac	(C)_Period1	(D)_Period2	(E)_Period3	(F)_Period4	(G)_DrugAsp	(H)_NPerio	(I)_NTDAYS
1	2.00	0.00	0.00	0.00	0.00	0.00	0.00	3.00	7.00
2	2.00	5.00	1.00	0.00	0.00	0.00	1.00	3.00	7.00
3	2.00	2.00	0.00	1.00	0.00	0.00	0.00	3.00	7.00
4	5.00	3.00	0.00	0.00	0.00	0.00	0.00	5.00	7.00
5	5.00	0.00	1.00	0.00	0.00	0.00	1.00	5.00	7.00
6	5.00	2.00	0.00	1.00	0.00	0.00	0.00	5.00	7.00
7	5.00	0.00	0.00	0.00	1.00	0.00	1.00	5.00	7.00
8	5.00	0.00	0.00	0.00	0.00	1.00	0.00	5.00	7.00
9	13.00	7.00	0.00	0.00	0.00	0.00	0.00	5.00	7.00
10	13.00	7.00	1.00	0.00	0.00	0.00	1.00	5.00	7.00

The variables of interest are:

- ID is the patient ID (27 patients in total).
- Headache is the number of headaches during the week (from 0 to 7).
- Period1 is a period 1 treatment indicator (1 for the first treatment period and 0 otherwise).
- Period2 is a period 2 treatment indicator (1 for the second treatment period and 0 otherwise).

- Period3 is a period 3 treatment indicator (1 for the third treatment period and 0 otherwise).
- Period4 is a period 4 treatment indicator (1 for the fourth treatment period and 0 otherwise).
- DrugAsp indicates the type of drug being used for the treatment, (0 = placebo and 1 = aspartame). 75 observations used placebo and 47 used aspartame.
- Nperiods is the number of periods the individual was observed (from 2 to 5).
- NTDays is the number of treatment days in the period (from 1 to 7).

3.3.2 A 2 level Poisson model with random intercept

3.3.2.1 The model

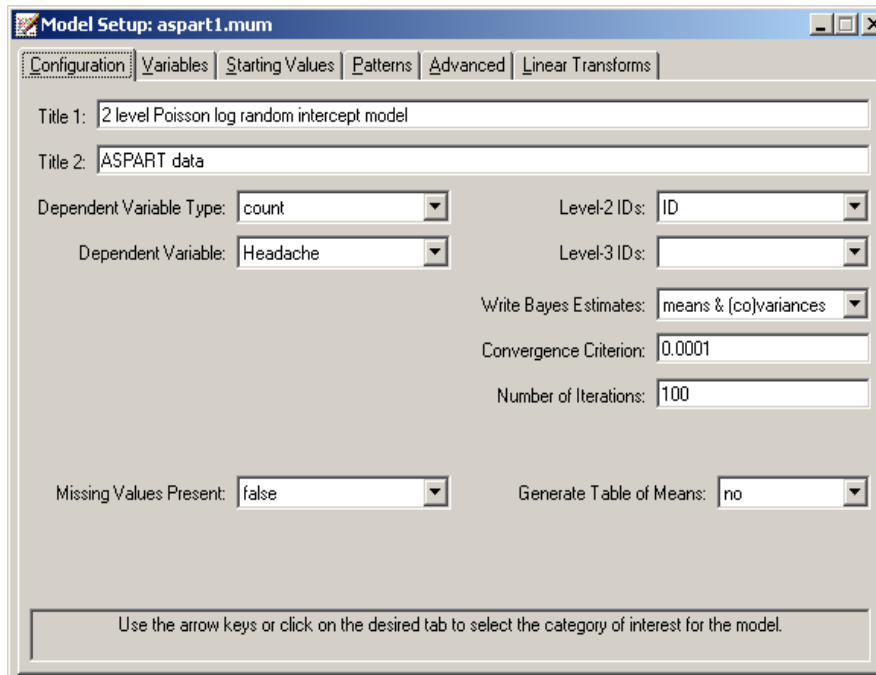
To model the relationship between the number of headaches during the week (Headache) and the treatment indicators (Period1 to Period4) and the type of drug administered (DrugAsp), the following Poisson regression model with a random intercept may be used:

$$\log(\hat{\lambda}_{ij}) = \beta_0 + \beta_1 \times \text{Period1}_{ij} + \beta_2 \times \text{Period2}_{ij} + \beta_3 \times \text{Period3}_{ij} \\ + \beta_4 \times \text{Period4}_{ij} + \beta_5 \times \text{DrugAsp}_{ij} + v_{i0}$$

where $\hat{\lambda}_{ij}$ denotes the estimated mean number of headaches of patient i for treatment period j ; Period1_{ij} , Period2_{ij} , Period3_{ij} and Period4_{ij} denote the values of the dummy variables Period1, Period2, Period3 and Period4 for patient i for treatment period j respectively; DrugAsp_{ij} denotes the value of the DrugAsp for patient i for treatment period j ; β_0 , β_1 , β_2 , β_3 , β_4 and β_5 denote unknown parameters; and v_{i0} denotes the random intercept for patient i for $i=1,2,\dots,27$ and $j=0,1,2,3$. This model is fitted to the data in **aspart.ss3** as described below.

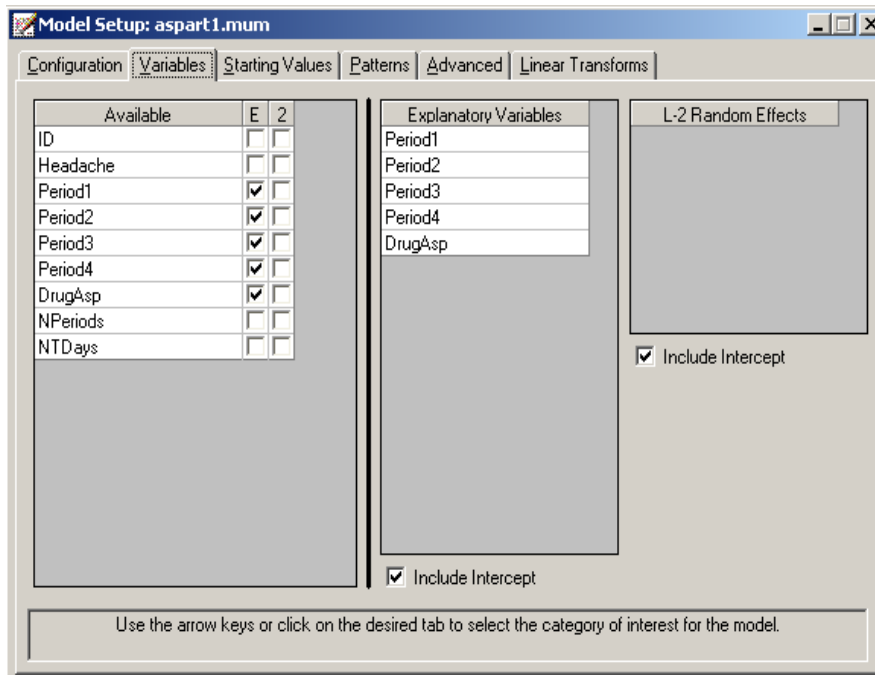
3.3.2.2 Setting up the analysis

Start by opening the SuperMix spreadsheet **aspart.ss3**. Select the **New Model Setup** option on the **File** menu to load the **Model Setup** window. On the **Configuration** tab, enter the titles 2 level Poisson log random intercept model and ASPART data for the analysis in the **Title 1** and **Title 2** text boxes respectively. The count outcome variable Headache is selected from the **Dependent Variable** drop-down list box. The **Dependent Variable Type** drop-down list box is used to indicate that the outcome variable is a count. The variable ID, which defines the levels of the hierarchy, is selected as the **Level-2 ID** from the **Level-2 IDs** drop-down list box.

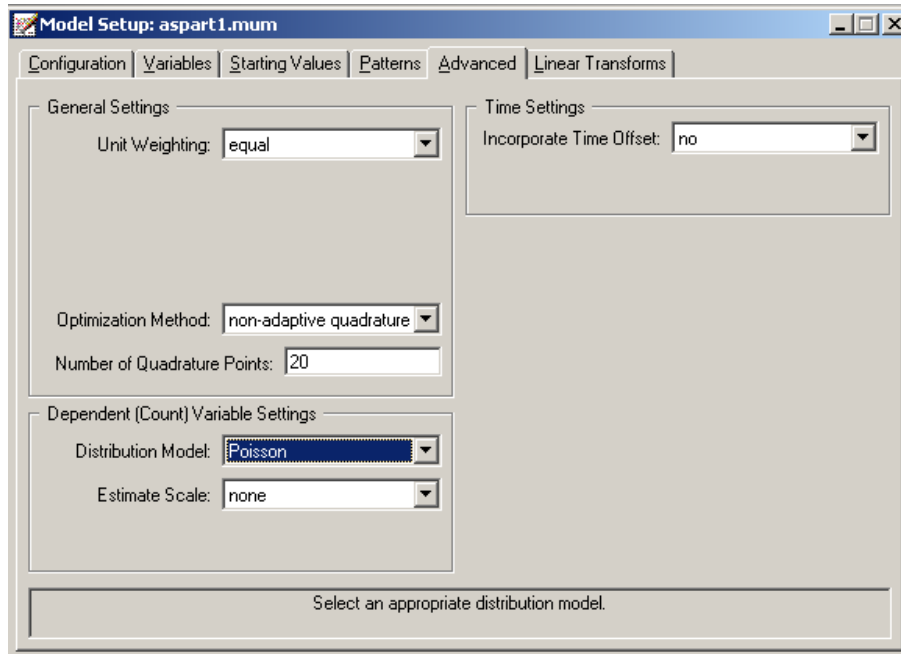


Next, click on the **Variables** tab to proceed with variable selection. The variables Period1, Period2, Period3, Period4, and DrugAsp are specified as the fixed effects of the model by checking the **E** check boxes for Period1, Period2, Period3, Period4, and DrugAsp in the **Available** grid. These actions produce the following **Variables** tab. By

default, an intercept model is included in the fixed part of the model, along with a random intercept at level 2.



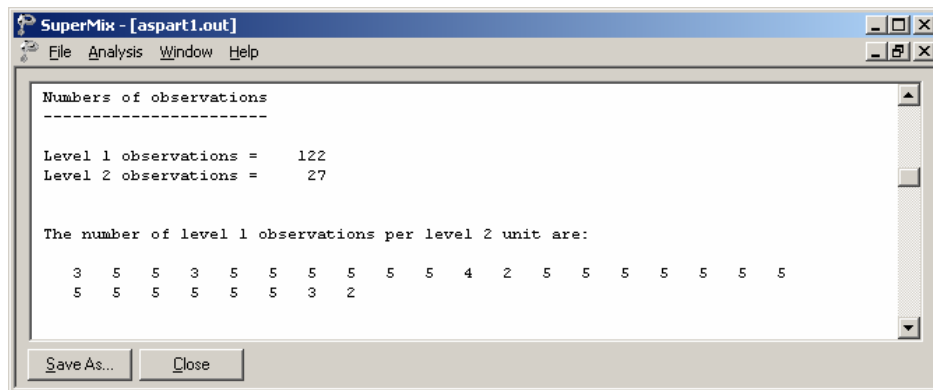
Finally, we click on the **Advanced** screen and keep all the default settings as shown below, except for the quadrature points which are set to 20. Before we can run the analysis, we have to save the model specifications to a file. This is accomplished by using the **Save** option on the **File** menu to open a **Save Mixed Up Model** dialog box. First enter the name **aspart1.mum** in the **File name** text box and then click on the **Save** button to save the file. The analysis is run by selecting the **Run** option from the **Analysis** menu. This produces the corresponding output file **aspart1.out**.



3.3.2.3 Discussion of results

Portions of this output file are shown below.

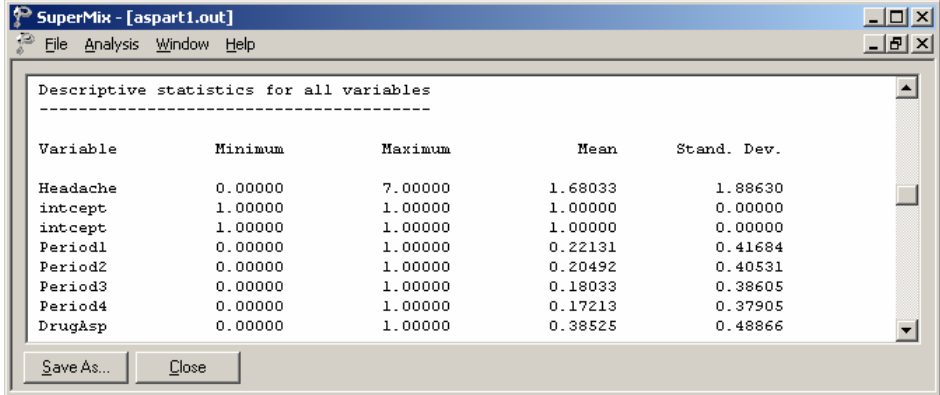
Data summary



The output file indicates that there are 27 subjects with 122 observations nested within them. The number of observations per subject varies between 2 and 5.

Descriptive statistics and starting values

The descriptive statistics for all the variables is shown next. The variance of Headache is $1.8863^2 = 3.5581$, which is substantially larger than the mean 1.6803. This might conflict with our assumption that the Poisson distribution is an appropriate choice for these data. As pointed out in Section 3.2.3, this can be verified by fitting a negative binomial model with a small dispersion parameter.



SuperMix - [aspart1.out]

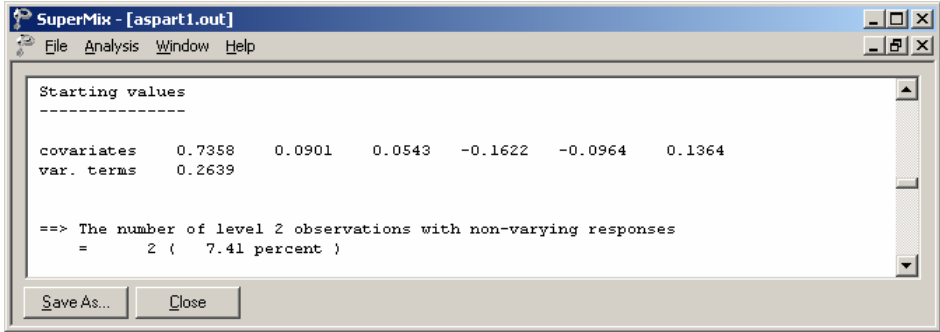
File Analysis Window Help

Descriptive statistics for all variables

Variable	Minimum	Maximum	Mean	Stand. Dev.
Headache	0.00000	7.00000	1.68033	1.88630
intcept	1.00000	1.00000	1.00000	0.00000
intcept	1.00000	1.00000	1.00000	0.00000
Period1	0.00000	1.00000	0.22131	0.41684
Period2	0.00000	1.00000	0.20492	0.40531
Period3	0.00000	1.00000	0.18033	0.38605
Period4	0.00000	1.00000	0.17213	0.37905
DrugAsp	0.00000	1.00000	0.38525	0.48866

Save As... Close

The starting values are given next.



SuperMix - [aspart1.out]

File Analysis Window Help

Starting values

covariates	0.7358	0.0901	0.0543	-0.1622	-0.0964	0.1364
var. terms	0.2639					

==> The number of level 2 observations with non-varying responses
= 2 (7.41 percent)

Save As... Close

Following the listing of the starting values, SuperMix indicates that of the 27 subjects, 2 had response vectors that were non-varying. Thus, 2 subjects gave identical responses at all time points that they were measured on.

Fixed and random effect results

The final results are shown next. The number of iterations needed for convergence and the deviance information are given first, followed by the estimates.

```

-----
* Final Results - Maximum Marginal Likelihood Estimates *
-----
Total Iterations = 34
Quad Pts per Dim = 20
Log Likelihood = -203.102
Deviance (-2logL) = 406.205
Ridge = 0.000

Variable      Estimate      Stand. Error      Z      p-value
-----
intcept       0.24035      0.14955      1.60713      0.10803 (2)
Period1       0.08031      0.23458      0.34236      0.73208 (2)
Period2       0.03412      0.22142      0.15410      0.87753 (2)
Period3      -0.22923      0.24952     -0.91867      0.35827 (2)
Period4      -0.16071      0.24717     -0.65019      0.51557 (2)
DrugAsp       0.21536      0.16278      1.32303      0.18582 (2)

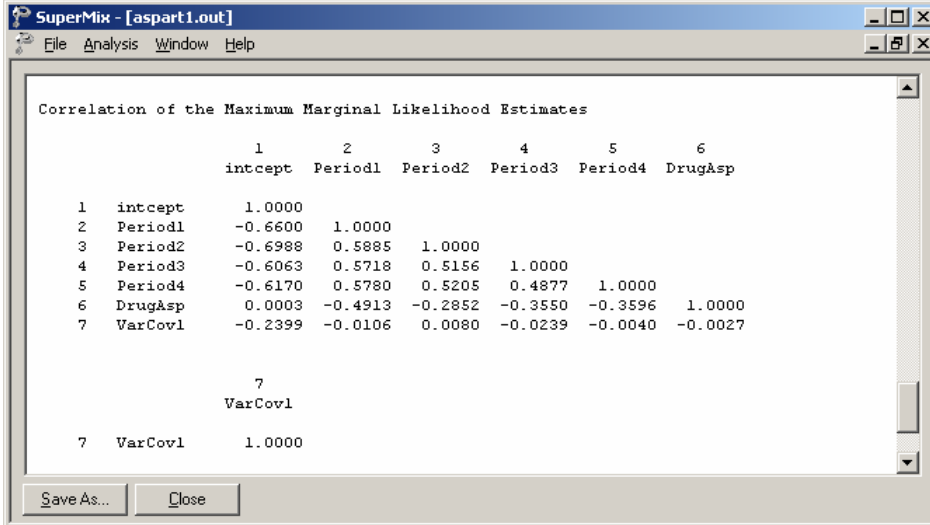
random effect variance term: expressed as a standard deviation
intcept       0.64277      0.05511     11.66395      0.00000 (1)

note: (1) = 1-tailed p-value
      (2) = 2-tailed p-value
  
```

The random-effect standard deviation is estimated as .643, and although a Wald test rejects the hypothesis that this parameter equals 0, use of the Wald test for testing whether variance parameters equal zero is questionable, since the Wald test is based on the assumption that parameters can assume any real value. Regarding the regression coefficients, all effects are non-significant. The results indicate that the model does not fit the data very well.

Correlation of estimates

The output provides the correlation matrix of the maximum likelihood estimates as shown below.



Correlation of the Maximum Marginal Likelihood Estimates

	1	2	3	4	5	6			
	intcept	Period1	Period2	Period3	Period4	DrugAsp			
1	intcept	1.0000							
2	Period1	-0.6600	1.0000						
3	Period2	-0.6988	0.5885	1.0000					
4	Period3	-0.6063	0.5718	0.5156	1.0000				
5	Period4	-0.6170	0.5780	0.5205	0.4877	1.0000			
6	DrugAsp	0.0003	-0.4913	-0.2852	-0.3550	-0.3596	1.0000		
7	VarCov1	-0.2399	-0.0106	0.0080	-0.0239	-0.0040	-0.0027	1.0000	
									7
									VarCov1
7	VarCov1								1.0000

It is important to realize that these are not the correlations of the variables, but of the parameter estimates. These correlations can be used to assess collinearity problems in estimation.

3.3.2.4 3.3.2.4 Interpreting the results

Estimated outcomes for groups: unit-specific results

The expected number of headaches can be obtained in the following fashion. First, we substitute the estimated coefficients in the model formulation

$$\begin{aligned}
\log(\widehat{\text{Headache}}_{ij}) &= \hat{\beta}_0 + \hat{\beta}_1 \times \text{Period1}_{ij} + \hat{\beta}_2 \times \text{Period2}_{ij} \\
&\quad + \hat{\beta}_3 \times \text{Period3}_{ij} + \hat{\beta}_4 \times \text{Period4}_{ij} + \hat{\beta}_5 \times \text{DrugAsp}_{ij} \\
&= 0.24035 + 0.08031 \times \text{Period1}_{ij} + 0.03412 \times \text{Period2}_{ij} \\
&\quad - 0.22923 \times \text{Period3}_{ij} - 0.16071 \times \text{Period4}_{ij} + 0.21536 \times \text{DrugAsp}_{ij}.
\end{aligned}$$

or, after taking exponents on both sides, as

$$\begin{aligned}
\widehat{\text{Headache}}_{ij} &= \exp(0.24035 + 0.08031 \times \text{Period1}_{ij} + 0.03412 \times \text{Period2}_{ij} \\
&\quad - 0.22923 \times \text{Period3}_{ij} - 0.16071 \times \text{Period4}_{ij} + 0.21536 \times \text{DrugAsp}_{ij}).
\end{aligned}$$

As an example, we calculate the expected number of headaches for a typical patient to whom aspartame was administered ($\text{DrugAsp} = 1$).

During the first treatment period, we find that for such a patient

$$\begin{aligned}
\widehat{\text{Headache}}_{ij} &= \exp(0.24035 + 0.08031 + 0.21536) \\
&= 1.7092.
\end{aligned}$$

The expected numbers of headaches for a typical patient (again with $\text{DrugAsp} = 1$) for the second, third, and fourth treatment periods are calculated as

$$\begin{aligned}
\widehat{\text{Headache}}_{ij} &= \exp(0.24035 + 0.03412 + 0.21536) \\
&= 1.6320, \\
\widehat{\text{Headache}}_{ij} &= \exp(0.24035 - 0.22923 + 0.21536) \\
&= 1.2542,
\end{aligned}$$

and

$$\widehat{\text{Headache}}_{ij} = \exp(0.24035 - 0.16071 + 0.21536) \\ = 1.3431$$

respectively. Complete results for all groups are given in Table XXX.2.

Estimated outcomes for groups: population-average results

The latent response variable model,

$$y_{ij} = \mathbf{z}'_{(1)ij} \mathbf{b}_i + \mathbf{x}'_{(1)ij} \boldsymbol{\beta}_{(1)} + \varepsilon_{ij},$$

makes the assumption that $\varepsilon_{ij} \sim LID(0, \sigma^2)$. For a Poisson distribution it is assumed that $\sigma_\varepsilon^2 = 1$. Under the assumption that \mathbf{v}_i and ε_{ij} are independently distributed, it follows that

$$\sigma_{y_{ij}}^2 = \mathbf{z}'_{ij} \boldsymbol{\Phi}_{v_i} \mathbf{z}_{ij} + \sigma_\varepsilon^2.$$

The design effect d_{ij} is defined as

$$d_{ij} = \frac{\sigma_{y_{ij}}^2}{\sigma_\varepsilon^2},$$

which, for the current model, may be calculated as

$$d_{ij} = \frac{\sigma_{y_{ij}}^2}{\sigma_\varepsilon^2} = \frac{\text{var}(v_{i0}) + 1}{1} = 1.4132$$

where $\text{var}(v_{i0}) = (0.64277)^2 = 0.4132$, with v_{i0} denoting the random intercept coefficient.

The estimated population-average probabilities (Hedeker & Gibbons, 2006) are obtained in a similar fashion as the unit-specific probabilities, after replacing the exponent in the formula used for calculation of the estimated unit-specific probabilities with $\exp = \exp/\sqrt{d_{ij}}$ as shown below.

$$\widehat{\text{Headache}}_{ij} = \exp[(0.24035 + 0.08031 \times \text{Period1}_{ij} + 0.03412 \times \text{Period2}_{ij} - 0.22923 \times \text{Period3}_{ij} - 0.16071 \times \text{Period4}_{ij} + 0.21536 \times \text{DrugAsp}_{ij}) / \sqrt{1.4132}].$$

The expected unit-specific and population average probabilities are summarized in Table XXX.2. We see that there is very little difference in the estimated number of headaches. This result is to be expected as the design effect is $\sqrt{1.4132} = 1.1888$.

Table XXX.2: Estimated unit-specific and population average results for groups

DRUGASP	Period	Estimated headaches (unit-specific)	Estimated headaches (population-average)
0	1	1.3780	1.3096
0	2	1.3158	1.2597
0	3	1.0112	1.0094
0	4	1.0829	1.0693
1	1	1.7092	1.5697
1	2	1.6320	1.5099
1	3	1.2542	1.2099
1	4	1.3431	1.2817

3.3.3 A 2-level Poisson log model with an offset variable

3.3.3.1 The model

The previous analysis assumed that the counts were all observed for the same number of days. However, this was not the case since the number of treatment days in the period did vary to some degree. Most of the counts were based on the full seven days in the week; however, some observations were made only for 1 day in the given week. To take this into account, we need to specify a so-called OFFSET variable. The offset variable indicates the amount of time that each count is based on. If OFFSET = no is specified, as was the case in the previous example, SuperMix assumes that all counts are based on the same amount of time.

The offset variable is introduced into the Poisson model in the following way:

$$\log(\hat{\lambda}_{ij}) = \log(\text{offset variable}) + [\mathbf{x}_{ij}' \mathbf{b}_i]$$

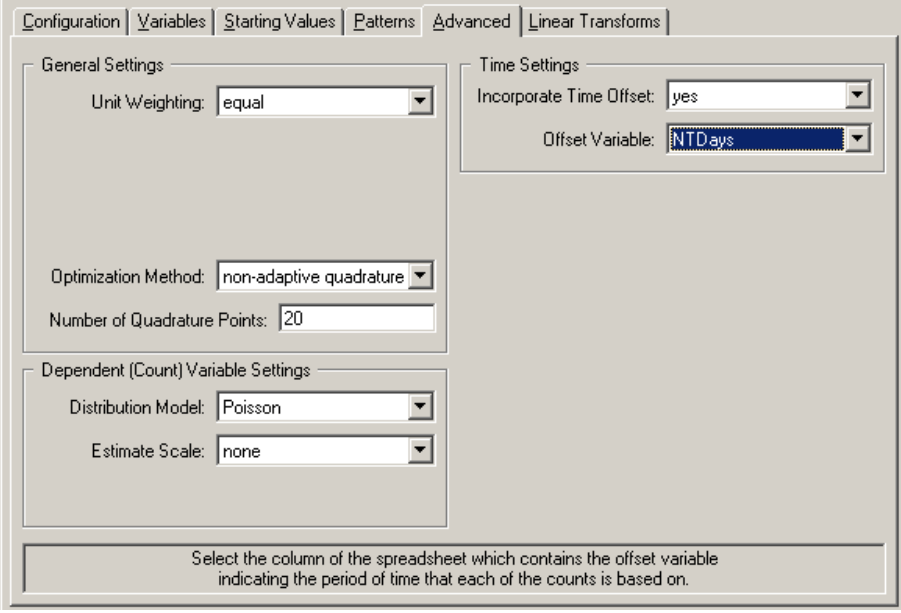
where \mathbf{x}_{ij} represent the values of the covariates corresponding to level-1 unit j nested within level-2 unit i and \mathbf{b}_i denotes the coefficient vector containing both fixed and random effects.

In the current situation, the variable NTDays is the appropriate choice as the OFFSET variable. The model to be fitted to the data now changes to:

$$\begin{aligned} \log(\text{Headache}_{ij}) = & \log(\text{NTDays}) + (\beta_0 + \beta_1 \times \text{Period1}_{ij} + \beta_2 \times \text{Period2}_{ij} \\ & + \beta_3 \times \text{Period3}_{ij} + \beta_4 \times \text{Period4}_{ij} + \beta_5 \times \text{DrugAsp}_{ij} + v_{i0}). \end{aligned}$$

3.3.3.2 Setting up the analysis

To create the model specifications for this model, start by opening **aspart.ss3** in a SuperMix spreadsheet window and using the **Open Existing Model Setup** option on the **File** menu to open the **Model Setup** window for **aspart1.mum**. On the **Configuration** screen, extend the title in the **Title 1** text box by adding the string "with Offset Variable." Next, click on the **Advanced** tab of the **Model Setup** window. Select **yes** from the **Incorporate Time Offset** drop-down list to activate the **Offset Variable** drop-down list box. Select the variable **NTDays** from the drop-down list of **Offset Variable** to produce the following **Advanced** tab.



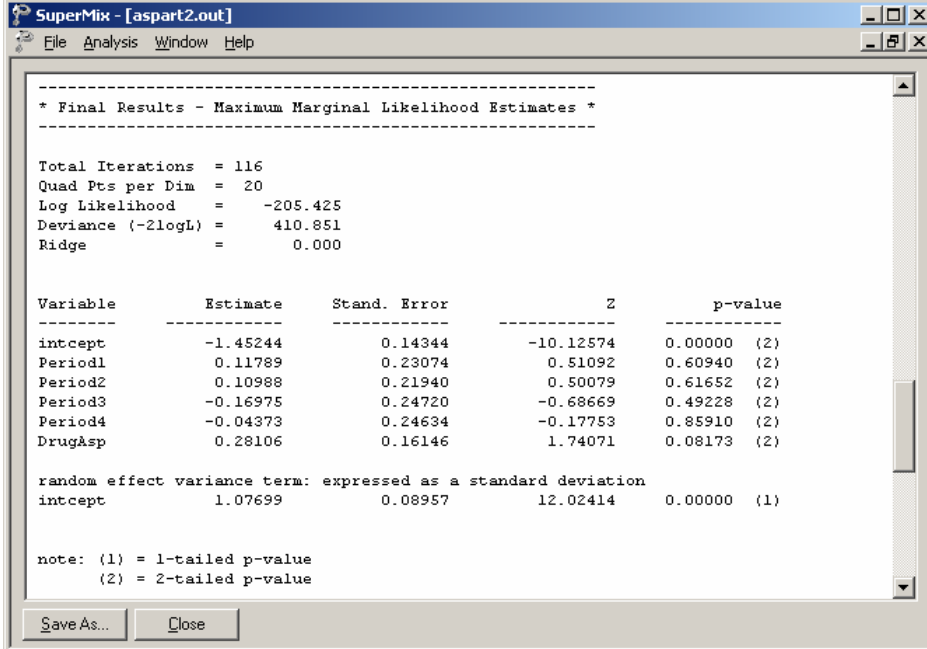
The screenshot shows the 'Advanced' tab of the 'Model Setup' window. The window has several tabs: Configuration, Variables, Starting Values, Patterns, Advanced (selected), and Linear Transforms. The 'Advanced' tab is divided into three main sections: 'General Settings', 'Time Settings', and 'Dependent (Count) Variable Settings'. In the 'General Settings' section, 'Unit Weighting' is set to 'equal', 'Optimization Method' is 'non-adaptive quadrature', and 'Number of Quadrature Points' is '20'. In the 'Time Settings' section, 'Incorporate Time Offset' is set to 'yes' and 'Offset Variable' is set to 'NTDays'. In the 'Dependent (Count) Variable Settings' section, 'Distribution Model' is 'Poisson' and 'Estimate Scale' is 'none'. At the bottom of the window, there is a note: 'Select the column of the spreadsheet which contains the offset variable indicating the period of time that each of the counts is based on.'

Save the changes to the file **aspart2.mum** by using the **Save As** option on the **File** menu and select the **Run** option on the **Analysis** menu to produce the output file **aspart2.out**.

3.3.3.3 Discussion of results

Fixed and random effect results

A portion of this output file is shown below.



The screenshot shows a window titled "SuperMix - [aspart2.out]" with a menu bar (File, Analysis, Window, Help). The main text area displays the following output:

```
-----  
* Final Results - Maximum Marginal Likelihood Estimates *  
-----  
  
Total Iterations = 116  
Quad Pts per Dim = 20  
Log Likelihood = -205.425  
Deviance (-2logL) = 410.851  
Ridge = 0.000  
  
Variable      Estimate      Stand. Error      Z      p-value  
-----  
intcept      -1.45244      0.14344      -10.12574      0.00000 (2)  
Period1       0.11789      0.23074       0.51092      0.60940 (2)  
Period2       0.10988      0.21940       0.50079      0.61652 (2)  
Period3      -0.16975      0.24720      -0.68669      0.49228 (2)  
Period4      -0.04373      0.24634      -0.17753      0.85910 (2)  
DrugAsp       0.28106      0.16146       1.74071      0.08173 (2)  
  
random effect variance term: expressed as a standard deviation  
intcept       1.07699      0.08957      12.02414      0.00000 (1)  
  
note: (1) = 1-tailed p-value  
      (2) = 2-tailed p-value  
  
Save As...  Close
```

Results for this model differ from those obtained for the model without offset variable discussed in the previous Section. While the overall trend in predictor coefficient estimates is similar, the intercept is now estimated as -1.45244, compared to 0.24035 previously. The variance in intercept over patients for this model is estimated as $(1.07699)^2 = 1.1599$ compared to $(0.6428)^2 = 0.4132$ previously.

3.3.3.4 3.3.3.4 Interpreting the results

Estimated outcomes for groups: unit-specific results

The expected number of headaches can be obtained in the following fashion. First, we substitute the estimated coefficients in the model formulation

$$\begin{aligned}\log(\widehat{\text{Headache}}_{ij}) &= \log(\text{NTDays}_{ij}) + (\hat{\beta}_0 + \hat{\beta}_1 \times \text{Period1}_{ij} + \hat{\beta}_2 \times \text{Period2}_{ij} \\ &\quad + \hat{\beta}_3 \times \text{Period3}_{ij} + \hat{\beta}_4 \times \text{Period4}_{ij} + \hat{\beta}_5 \times \text{DrugAsp}_{ij}) \\ &= \log(\text{NTDays}_{ij}) + (-1.45244 + 0.11789 \times \text{Period1}_{ij} + 0.10988 \times \text{Period2}_{ij} \\ &\quad - 0.16975 \times \text{Period3}_{ij} - 0.04373 \times \text{Period4}_{ij} + 0.28106 \times \text{DrugAsp}_{ij}),\end{aligned}$$

or, after taking exponents on both sides, as

$$\begin{aligned}\widehat{\text{Headache}}_{ij} &= \text{NTDays}_{ij} \times \exp(-1.45244 + 0.11789 \times \text{Period1}_{ij} + 0.10988 \times \text{Period2}_{ij} \\ &\quad - 0.16975 \times \text{Period3}_{ij} - 0.04373 \times \text{Period4}_{ij} + 0.28106 \times \text{DrugAsp}_{ij}).\end{aligned}$$

As most observations had a value of $\text{NTDays} = 7$, we start by considering typical patients with a full set of treatment days. We also assume that $\text{DrugAsp} = 1$, in other words, that aspartame rather than a placebo was administered.

During the first treatment period, we find that for such a patient

$$\begin{aligned}\widehat{\text{Headache}}_{ij} &= 7 \exp(-1.45244 + 0.11789 + 0.28106) \\ &= 7 \exp(-1.05349) \\ &= 2.4410.\end{aligned}$$

The expected numbers of headaches for a typical patient (again with $\text{NTDays} = 7$ and $\text{DrugAsp} = 1$) for the second, third, and fourth treatment periods are calculated as

$$\widehat{\text{Headache}}_{ij} = 7 \exp(-1.45244 + 0.10988 + 0.28106)$$

$$= 2.4216,$$

$$\widehat{\text{Headache}}_{ij} = 7 \exp(-1.45244 - 0.16975 + 0.28106)$$

$$= 1.8308,$$

and

$$\widehat{\text{Headache}}_{ij} = 7 \exp(-1.45244 - 0.04373 + 0.28106)$$

$$= 2.0767$$

respectively.

For a typical patient with only 5 treatment days, the expected numbers of headaches in each of the four treatment periods are 1.7436, 1.7297, 1.3077, and 1.4834 respectively.

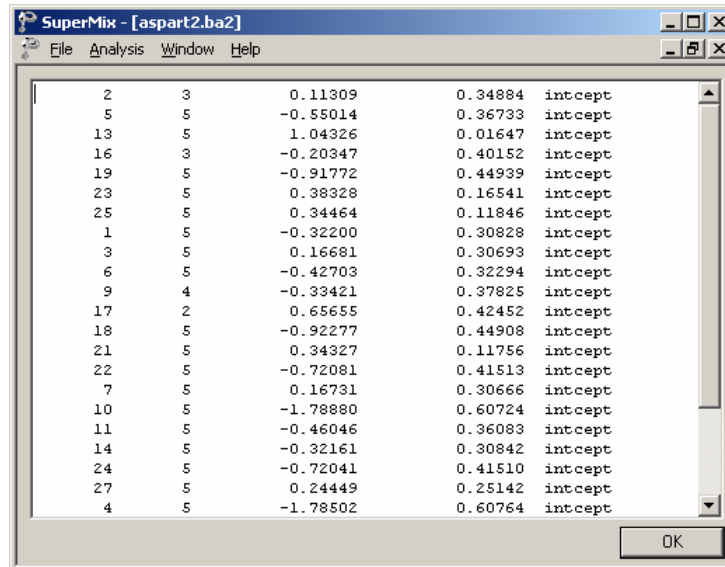
When the expected numbers of headaches for a typical patient receiving aspartame under the Poisson model without offset variable (see previous Section) and the Poisson model with offset variable are compared, we clearly see the impact of the inclusion of the offset variable on the estimated coefficients. These results are summarized in Table XXX.3.

Table XXX.3: Comparison of results for Poisson models

Period	Without offset variable	With offset variable (NTDays = 7)	With offset variable (NTDays = 5)
1	2.3553	2.4410	1.7436
2	1.6320	2.4216	1.7297
3	1.2542	1.8308	1.3077
4	1.3431	2.0767	1.4834

Level 2 Bayes results

As requested during the model specification stage, the empirical Bayes estimates of the random effects are written to the file **aspart2.ba2**. The first few lines of this file are shown below.



ID	Number of observations	Empirical Bayes estimate	Posterior standard deviation	Coefficient name
2	3	0.11309	0.34884	intcept
5	5	-0.55014	0.36733	intcept
13	5	1.04326	0.01647	intcept
16	3	-0.20347	0.40152	intcept
19	5	-0.91772	0.44939	intcept
23	5	0.38328	0.16541	intcept
25	5	0.34464	0.11846	intcept
1	5	-0.32200	0.30828	intcept
3	5	0.16681	0.30693	intcept
6	5	-0.42703	0.32294	intcept
9	4	-0.33421	0.37825	intcept
17	2	0.65655	0.42452	intcept
18	5	-0.92277	0.44908	intcept
21	5	0.34327	0.11756	intcept
22	5	-0.72081	0.41513	intcept
7	5	0.16731	0.30666	intcept
10	5	-1.78880	0.60724	intcept
11	5	-0.46046	0.36083	intcept
14	5	-0.32161	0.30842	intcept
24	5	-0.72041	0.41510	intcept
27	5	0.24449	0.25142	intcept
4	5	-1.78502	0.60764	intcept

The file **mixreg.ba2** contains five pieces of information per individual:

- the individual's ID,
- the number of repeated observations for that individual,
- the empirical Bayes estimate for that individual (which is the mean of the posterior distribution),
- the associated posterior standard deviation, and
- the name of the relevant random coefficient.

Since they are estimates of b_{i0} for each individual, the empirical Bayes estimates are expressed on the standard normal scale. Inspection of these estimates indicates

that subject 13 has a very high score. This person's estimate of 1.043 (with standard deviation .016) suggests a very high level of headaches. This agrees well with the raw data, which reveals that this person recorded 7 headaches on four occasions and 6 on the only other occasion.

Graphical displays

Figure XXX.5 is a comparison (represented by a dotted line) of the predicted average number of headaches reported by each patient when taking a placebo (left axis) as opposed to the predicted average number when the treatment is aspartame (right axis). From the graphical display, it appears as if all of the lines (each representing a patient) have a positive slope. The slopes become steeper as the number of headaches increases. This suggests an increase, albeit small, in the expected average number of headaches when aspartame is used. Note that patient 13, who reported a consistently high number of headaches at all occasions, was excluded from this graph.

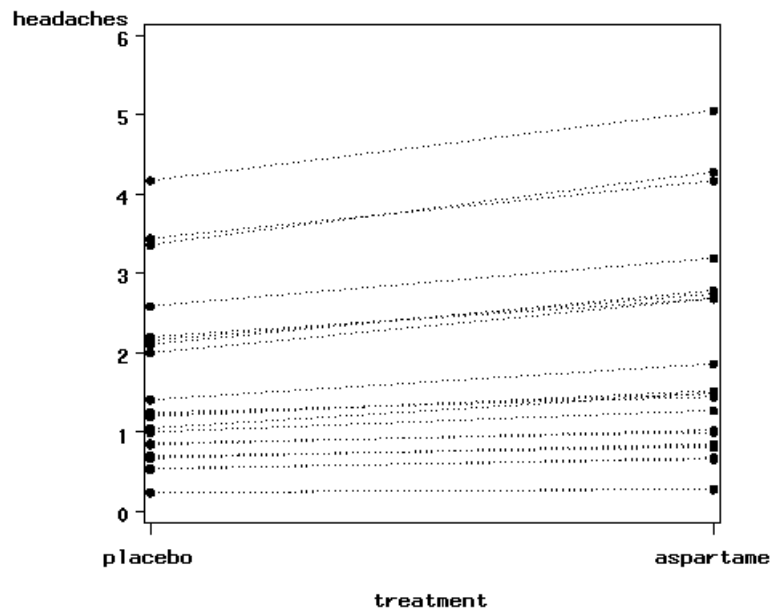
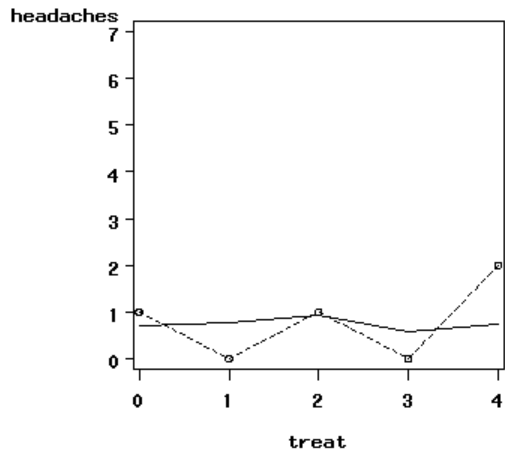
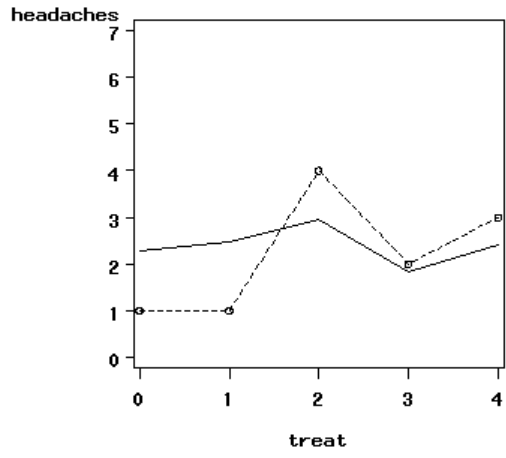


Figure XXX.5: Predicted average number of headaches for placebo and aspartame

Fitted line and observed trajectory
patient=24



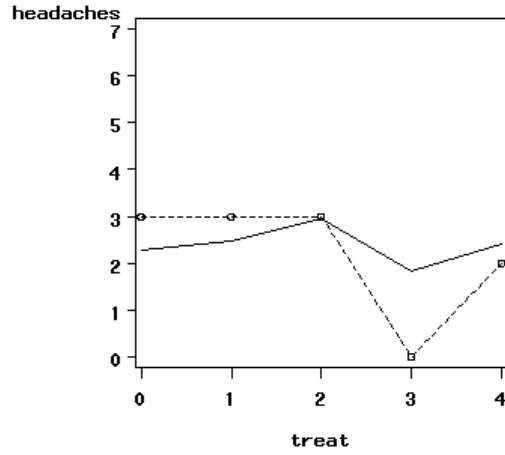
Fitted line and observed trajectory
patient=7



Fitted line and observed trajectory
patient=3



Fitted line and observed trajectory
patient=27



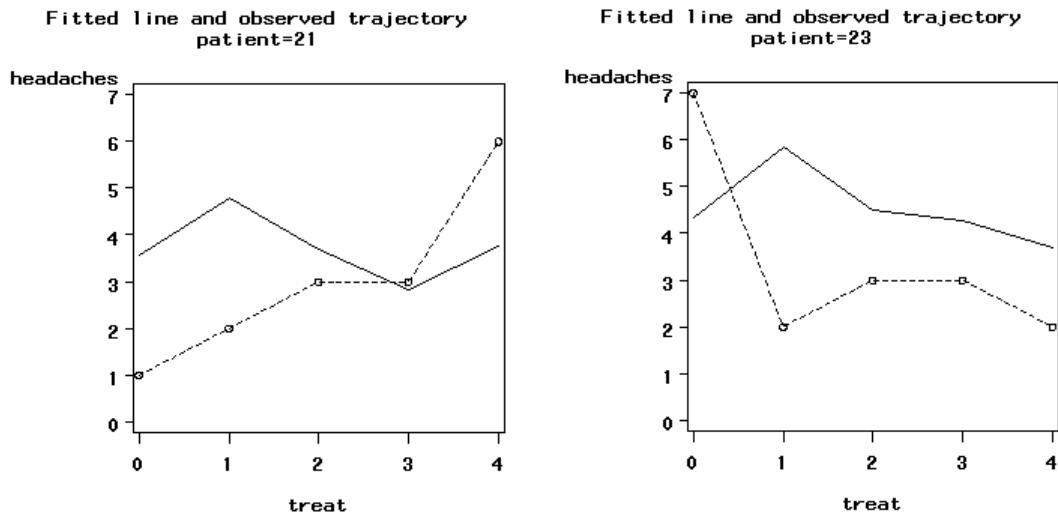


Figure XXX.6: Fitted and observed trajectories

Figure XXX.6 is a graphical display of the fitted trajectory (solid line) and observed trajectory (dotted line) for a sample of 6 patients. These displays are ordered from a patient who reported a relatively small number of headaches at the different treatment occasions to one who reported a relatively high number of headaches at the treatment occasions. A study of the fitted and observed trajectories reveals that, in general, the model fit is best when the number of headaches is smaller and becomes less accurate as the number of headaches increases. For patient 13, who is not represented in the graphical display, the number of predicted headaches is almost twice the number observed.

The fitted lines were obtained as

$$\widehat{\text{Headache}}_{ij} = \text{NTDays}_{ij} \times \exp(-1.45244 + 0.11789 \times \text{Period1}_{ij} + 0.10988 \times \text{Period2}_{ij} - 0.16975 \times \text{Period3}_{ij} - 0.04373 \times \text{Period4}_{ij} + 0.28106 \times \text{DrugAsp}_{ij}) + \hat{b}_{i0}.$$

where \hat{b}_{i0} is obtained from the **aspart2.ba2** file, shown previously in this section.

References

McKnight and Van Den Eeden (1993)