

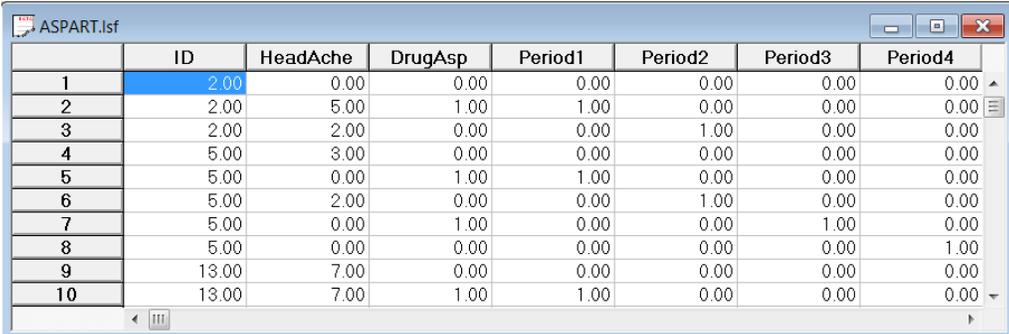
Models for count outcomes using ASPART data

Contents

1. The data.....	1
2. The model.....	2
3. Setting up the analysis	2
4. Discussion of results.....	5
5. Interpreting the results.....	7

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, contained in the file **aspart.isf** (**Multilevel Generalized Linear model Examples** folder) are shown below in the form of an LSF spreadsheet window.



	ID	HeadAche	DrugAsp	Period1	Period2	Period3	Period4
1	2.00	0.00	0.00	0.00	0.00	0.00	0.00
2	2.00	5.00	1.00	1.00	0.00	0.00	0.00
3	2.00	2.00	0.00	0.00	1.00	0.00	0.00
4	5.00	3.00	0.00	0.00	0.00	0.00	0.00
5	5.00	0.00	1.00	1.00	0.00	0.00	0.00
6	5.00	2.00	0.00	0.00	1.00	0.00	0.00
7	5.00	0.00	1.00	0.00	0.00	1.00	0.00
8	5.00	0.00	0.00	0.00	0.00	0.00	1.00
9	13.00	7.00	0.00	0.00	0.00	0.00	0.00
10	13.00	7.00	1.00	1.00	0.00	0.00	0.00

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).

2. The model

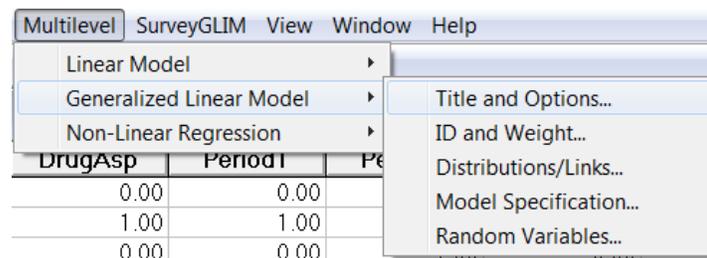
To model the relationship between the number of headaches during the week (Headache) and the treatment indicators (Period1 to Period4) and the type of drug administered (DrugAsp), the following Poisson regression model with a random intercept may be used:

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

where λ_{ij} denotes the 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. Setting up the analysis

Start by opening **aspart.lsf**. Select the **Multilevel, Generalized Linear Model** option on the main menu bar.



Using the **Titles and Options** dialog enter the title: 2-level Poisson log random intercept model, ASPART data. Select **Quadrature**, and set the number of quadrature points to 20.

Title and Options

Title: 2-Level Poisson Log random intercept model, ASPART data

Maximum Number of Iterations: 100

Convergence Criterion: 0.0001

Missing Data Value: -999999

Dependent Missing Value: -999999

Optimization Method

MAP Quadrature

Number of Quadrature Points: 10

Additional Output

Residual files No data summary

Asymptotic covariance

Next >> Cancel OK

To build syntax, proceed to the Random Variables screen and click the Finish button

Click the **Next** button to activate the **ID and Weight** dialog and select the variable ID, which defines the second level of the hierarchy, as the **Level-2 ID**. To proceed to the **Distributions/Links** dialog, click the **Next** button and enter the values shown below.

Distributions and Links

Distribution type: Poisson

Link function: Log

Include intercept?

Yes No

Dispersion parameter

Yes Fixed value:

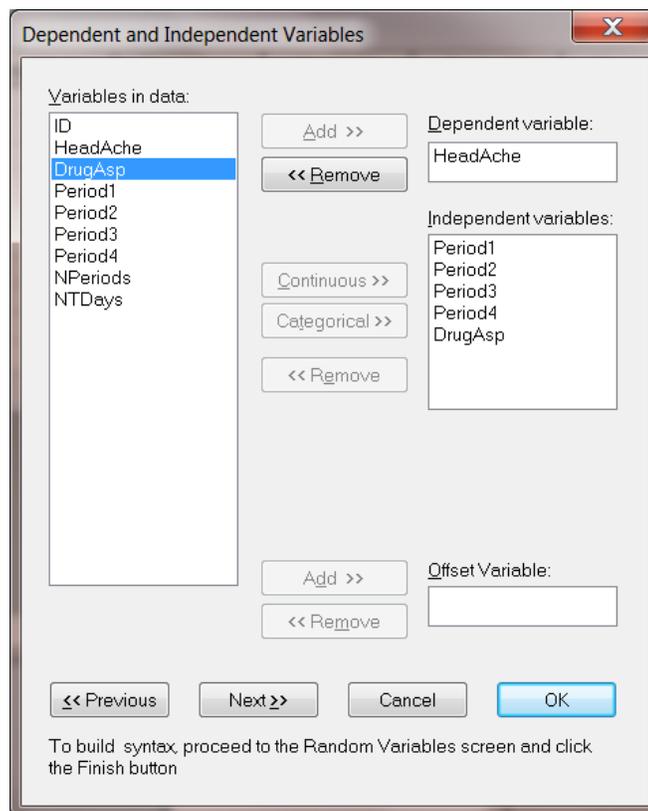
Estimate scale? None

<< Previous Next >> Cancel OK

To build syntax, proceed to the Random Variables screen and click the Finish button

Proceed to the **Dependent and Independent variable** dialog and select count outcome variable Headache as the dependent variable.

The variables Period1, Period2, Period3, Period4, and DrugAsp are specified as the independent variables of the model. By default, an intercept model is included in the fixed part of the model, along with a random intercept at level 2.



Click the **Finish** button on the **Random Variables** dialog to produce the syntax file shown below.

```
MGlimOptions Converge=0.0001 MaxIter=100 MissingCode=-999999 |
IterDetails=No Method=Quad NQUADPTS=20 Output = Residuals;
Title=2-level Poisson log random intercept model, ASPART data;
SY=aspart.lsf;
ID2=ID;
Distribution=POI;
Link=LOG;
Intercept=Yes;
Scale=None;
DepVar=HeadAche;
CoVars=DrugAsp Period1 Period2 Period3 Period4;
RANDOM2=intcept;
```

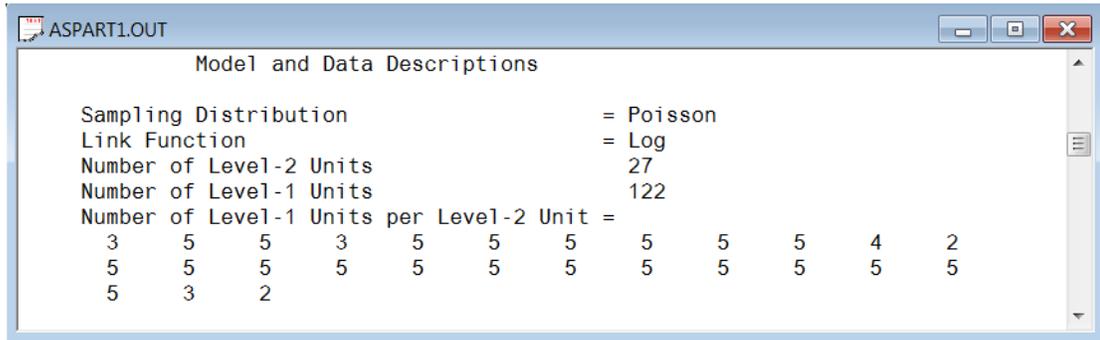
Before running the analysis, save the file as **aspart1.prl** and click the **RUN PRELIS** icon button to start the analysis.

4. Discussion of results

Portions of this output file are shown below.

Model and data description

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.

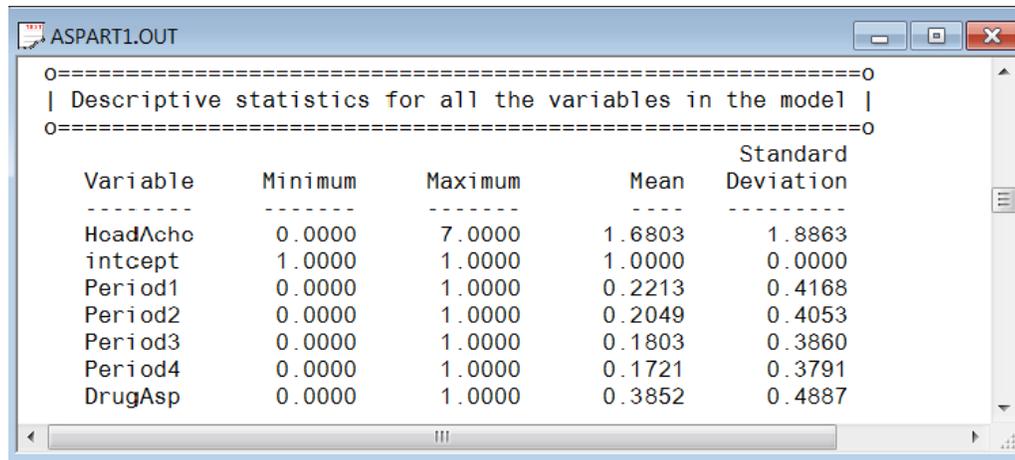


The screenshot shows a window titled 'ASPART1.OUT' with the subtitle 'Model and Data Descriptions'. It displays the following information:

```
Sampling Distribution           = Poisson
Link Function                   = Log
Number of Level-2 Units        = 27
Number of Level-1 Units        = 122
Number of Level-1 Units per Level-2 Unit =
3   5   5   3   5   5   5   5   5   5   5   4   2
5   5   5   5   5   5   5   5   5   5   5   5   5
5   3   2
```

Descriptive statistics

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. This can be verified by fitting a negative binomial model with a small dispersion parameter.



The screenshot shows a window titled 'ASPART1.OUT' displaying descriptive statistics for all variables in the model. The output is as follows:

```
=====0
| Descriptive statistics for all the variables in the model |
=====0
```

Variable	Minimum	Maximum	Mean	Standard Deviation
HeadAche	0.0000	7.0000	1.6803	1.8863
intcept	1.0000	1.0000	1.0000	0.0000
Period1	0.0000	1.0000	0.2213	0.4168
Period2	0.0000	1.0000	0.2049	0.4053
Period3	0.0000	1.0000	0.1803	0.3860
Period4	0.0000	1.0000	0.1721	0.3791
DrugAsp	0.0000	1.0000	0.3852	0.4887

Results for the model without any random effects

The results for the model without any random effects are shown below. In this section the goodness of fit statistics, estimated regression weights and event rate ratio and 95% event rate confidence intervals are included.

Statistic	Value	DF	Ratio
Likelihood Ratio Chi-square	243.8257	116	2.1019
Pearson Chi-square	253.8934	116	2.1887

Estimated regression weights				
Parameter	Estimate	Standard Error	z Value	P Value
intcept	0.4654	0.1525	3.0516	0.0023
Period1	0.0916	0.2265	0.4043	0.6860
Period2	0.0131	0.2276	0.0575	0.9542
Period3	-0.2245	0.2471	-0.9084	0.3637
Period4	-0.1840	0.2540	-0.7242	0.4689
DrugAsp	0.2332	0.1596	1.4612	0.1440

Event Rate Ratio and 95% Event Rate Confidence Intervals				
Parameter	Estimate	Event Rate	Bounds	
			Lower	Upper
intcept	0.4654	1.5926	1.1811	2.1474
Period1	0.0916	1.0959	0.7030	1.7085
Period2	0.0131	1.0132	0.6486	1.5827
Period3	-0.2245	0.7989	0.4923	1.2967
Period4	-0.1840	0.8320	0.5057	1.3688
DrugAsp	0.2332	1.2626	0.9235	1.7263

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.

The random-effect variance component is estimated as .4327 which is significant at the 5% level. Regarding the regression coefficients, all effects are non-significant. The results indicate that the model does not fit the data very well.

The event ratio and 95% event rate confidence interval and estimated level-2 variances and covariances are shown next to the estimated regression weights. The event ratios are the exponents ($e^{\hat{\beta}}$) of the estimated regression coefficients.

Parameter	Estimate	Standard Error	z Value	P Value
intcept	0.2507	0.2051	1.2222	0.2216
Period1	0.0806	0.2350	0.3429	0.7317
Period2	0.0345	0.2236	0.1541	0.8775
Period3	-0.2273	0.2546	-0.8927	0.3720
Period4	-0.1594	0.2528	-0.6306	0.5283
DrugAsp	0.2151	0.1639	1.3123	0.1894

Parameter	Estimate	Event Rate	Lower	Upper
intcept	0.2507	1.2849	0.8596	1.9208
Period1	0.0806	1.0839	0.6838	1.7181
Period2	0.0345	1.0351	0.6677	1.6045
Period3	-0.2273	0.7967	0.4837	1.3122
Period4	-0.1594	0.8526	0.5194	1.3995
DrugAsp	0.2151	1.2399	0.8993	1.7096

Parameter	Estimate	Standard Error	z Value	P Value
intcept/intcept	0.4327	0.1737	2.4917	0.0127

5. 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

$$\begin{aligned}
 \log\left(\widehat{Headache}_{ij}\right) &= \hat{\beta}_0 + \hat{\beta}_1 \times \text{Period1}_{ij} + \hat{\beta}_2 \times \text{Period2}_{ij} \\
 &\quad + \hat{\beta}_3 \times \text{Period3}_{ij} + \hat{\beta}_4 \times \text{Period4}_{ij} + \hat{\beta}_5 \times \text{DrugAsp}_{ij} \\
 &= 0.2572 + 0.0807 \times \text{Period1}_{ij} + 0.0345 \times \text{Period2}_{ij} \\
 &\quad - 0.2267 \times \text{Period3}_{ij} - 0.1592 \times \text{Period4}_{ij} + 0.2151 \times \text{DrugAsp}_{ij}.
 \end{aligned}$$

or, after taking exponents on both sides, as

$$\begin{aligned}
 \widehat{Headache}_{ij} &= \exp(0.2572 + 0.0807 \times \text{Period1}_{ij} + 0.0345 \times \text{Period2}_{ij} \\
 &\quad - 0.2267 \times \text{Period3}_{ij} - 0.1592 \times \text{Period4}_{ij} + 0.2151 \times \text{DrugAsp}_{ij}).
 \end{aligned}$$

As an example, we calculate the expected number of headaches for a typical patient to whom aspartame was administered (DrugAsp = 1). During the first treatment period, we find that for such a patient

$$\begin{aligned}\widehat{Headache}_{ij} &= \exp(0.2572 + 0.0807 + 0.2151) \\ &= 1.7385.\end{aligned}$$

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

$$\begin{aligned}\widehat{Headache}_{ij} &= \exp(0.2572 + 0.0345 + 0.2151) \\ &= 1.6600,\end{aligned}$$

$$\begin{aligned}\widehat{Headache}_{ij} &= \exp(0.2572 - 0.2267 + 0.2151) \\ &= 1.2784,\end{aligned}$$

and

$$\begin{aligned}\widehat{Headache}_{ij} &= \exp(0.2752 - 0.1592 + 0.2151) \\ &= 1.3677\end{aligned}$$

respectively. Complete results for all groups are given in Table 5.2.

Estimated outcomes for groups: population-average results

The latent response variable model,

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

makes the assumption that $e_{ij} \sim LID(0, \sigma_e^2)$. For a Poisson distribution it is assumed that $\sigma_e^2 = 1$.

Under the assumption that \mathbf{v}_i and e_{ij} are independently distributed, it follows that

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

The design effect d_{ij} is defined as

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

which, for the current model, may be calculated as

$$d_{ij} = \frac{\sigma_{y_{ij}}^2}{\sigma_e^2} = \frac{\text{var}(v_{i0}) + 1}{1} = 1.4290$$

where $\text{var}(v_{i0}) = 0.4290$, 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.

$$\begin{aligned}\widehat{Headache}_{ij} &= \exp[(0.2572 + 0.0807 \times \text{Period1}_{ij} + 0.0345 \times \text{Period2}_{ij} - 0.2267 \times \text{Period3}_{ij} \\ &\quad - 0.1592 \times \text{Period4}_{ij} + 0.2151 \times \text{DrugAsp}_{ij}) / \sqrt{1.4290}].\end{aligned}$$

The expected unit-specific and population average probabilities are summarized in Table 5.3. 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 $\sqrt{1.4290} = 1.1954$.

Estimated unit-specific and population average results for groups

DRUGASP	Period	Estimated headaches (unit-specific)	Estimated headaches (population-average)
0	1	1.4020	1.1728
0	2	1.3387	1.1199
0	3	1.0310	0.8624
0	4	1.1030	0.9227
1	1	1.7385	1.4543
1	2	1.6600	1.3886
1	3	1.2784	1.0694
1	4	1.3677	1.1441