

# **Two-level Poisson model**

## **Contents**

3.1	Two	-LEVEL MODELS FOR COUNT OUTCOMES FROM ASPART DATA	. 1
3.1.1	1	The data	. 1
		A 2 level Poisson model with random intercept	

# 3.1 Two-level models for count outcomes from ASPART data

## 3.1.1 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**.

)			Apply						
	(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
•									Þ

The variables of interest are:

- ID is the patient ID (27 patients in total).
- $\circ$  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).

#### 3.1.2 A 2 level Poisson model with random intercept

#### 3.1.2.1 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:

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

where  $\lambda_{ij}$  denotes the mean number of headaches of patient *i* for treatment period *j*; Period1<sub>ij</sub>, Period2<sub>ij</sub>, Period3<sub>ij</sub> and Period4<sub>ij</sub> denote the values of the dummy variables Period1, Period2, Period3 and Period4 for patient *i* for treatment period *j* respectively; DrugAsp<sub>ij</sub> denotes the value of the DrugAsp for patient *i* for treatment period *j*;  $\beta_0$ ,  $\beta_1$ ,  $\beta_2$ ,  $\beta_3$ ,  $\beta_4$ and  $\beta_5$  denote unknown parameters; and  $v_{i0}$  denotes the random intercept for patient *i* for *i* = 1, 2, ..., 27 and *j* = 0, 1, 2, 3. This model is fitted to the data in **aspart.ss3** as described below.

#### 3.1.2.2 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.

🜠 Model Setup: aspart1.mum		<u> </u>
Configuration	anced 🛛 Linear Transforms 🗎	
Title 1: 2 level Poisson log random intercept model		
Title 2: ASPART data		
Dependent Variable Type: count	Level-2 IDs: ID	•
Dependent Variable: Headache	Level-3 IDs:	•
	Write Bayes Estimates: means & (co)variance	es 💌
	Convergence Criterion: 0.0001	
	Number of Iterations: 100	
Missing Values Present: false	Generate Table of Means: no	•
Use the arrow keys or click on the desired tab to	Output Type: standard	•

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.

Model Setup: aspart1.mum
Configuration Variables Starting Values Patterns Advanced Linear Transforms
Available       E       2         ID       ID       ID         Headache       ID       ID         Period1       IV       ID         Period2       IV       ID         Period3       IV       ID         Period3       IV       ID         Period4       IV       ID         DrugAsp       IV       ID         NPeriod3       IV       IN         NTDays       IV       IN         IV       IN       IV         IV       IV       IV

🜠 Model Setup: aspart1.mum	
Configuration   Variables   Starting Values   Patterns Ad	vanced Linear Transforms
General Settings Unit Weighting: equal	Time Settings
Optimization Method: non-adaptive quadrature  Number of Quadrature Points: 20 Dependent (Count) Variable Settings	
Distribution Model: Poisson	
Select an appropriat	e distribution model.

Finally, we click on the **Advanced** screen and keep all the default settings as shown above, except for those concering the method of estimation. Select **non-adaptive quadrature**, and set the quadrature points 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**.

#### 3.1.2.3 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.

1	SuperMix - [aspart1.out]	_ 🗆 🗙
2.40	Eile <u>A</u> nalysis <u>W</u> indow <u>H</u> elp	_ 8 ×
Γ	Model and Data Descriptions	
	Sampling Distribution = Poisson Link Function = Log	
	Number of Level-1 Units 27 Number of Level-1 Units 122	
L	Number of Level-1 Units per Level-2 Unit =	
	3 5 5 3 5 5 5 5 5 4 2 5 5 5 5 5 5 5 5 5 5 5	
	5 3 2	-
L		F
	Save As Close	

### **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.

Variable	Minimum	Maximum	Mean	Standard Deviation	
Headache	0.0000	7.0000	1.6803	1.8863	
intercept	1.0000	1.0000	1.0000	0.0000	
Periodl	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.

-			-	
	he model without any			
0			0	
	Goodness of fit	statistics		
Statistic		Value	DF	Ratio
Likelihood Ratio	Chi-smare	243.8257	116	2.1019
Pearson Chi-squa	-	253.8934	116	2.1817
-				
	Estimated regres	ssion weights		
	Dollmanca regre	-		
Parameter	Estimate	Standard Error	z Value	P Value
			z varue	F VAIUE
intercept	0.4654	0.1525	3.0516	0.0023
Periodl	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		0.4689
DrugAsp	0.2332	0.1596	1.4612	0.1440
Event Rate Rat	io and 95% Event Rat	ce Confidence Int	ervals	
			Bour	
Parameter	Estimate	Event Rate	Lower	Upper
intercept	0.4654	1.5926	1,1811	2.1474
intercept Periodl	0.4654	1.5926	0.7030	2.1474 1.7085
Period2	0.0316	1.0132	0.6486	1.5827
Period2 Period3	-0.2245	0.7989	0.4923	1.2967
Period3 Period4	-0.1840	0.8320	0.5057	1.3688
DrugAsp	0.2332	1.2626	0.9235	1.7263
	5.2002	1.2020		

### 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 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.

<sup>®</sup> <u>Fi</u> le <u>A</u> nalysis <u>W</u> indow	Help				_ 8
0==========					
Optimization	Method: Non-Adaptiv	ve Quadrature			
0=========					
Number of quadra	ature points =	20			
Number of free p	arameters =	7			
Number of iterat	ions used =	4			
-21nL (deviance	statistic) =	406.34905			
Akaike Informati	ion Criterion	420.34905			
Schwarz Criterio	n	439.97720			
	Estimated regre	ssion weights			
	Estimated regre	-			
Parameter	Estimated regre Estimate	ssion weights Standard Error	z Value	P Value	
	Estimate	Standard Brror			
intercept	Estimate 0.2572	Standard Error  0.2024	1.2705	0.2039	
intercept Periodl	Estimate  0.2572 0.0807	Standard Error  0.2024 0.2349	1.2705 0.3434	0.2039	
intercept Periodl Period2	Estimate 0.2572 0.0807 0.0345	Standard Error 0.2024 0.2349 0.2237	1.2705 0.3434 0.1542	0.2039 0.7313 0.8775	
intercept Periodl Period2 Period3	Estimate  0.2572 0.0307 0.0345 -0.2267	Standard Error 0.2024 0.2349 0.2237 0.2545	1.2705 0.3434 0.1542 -0.8909	0.2039 0.7313 0.8775 0.3730	
intercept Periodl Period2 Period3 Period4	Estimate 0.2572 0.0807 0.0345 -0.2267 -0.1592	Standard Error 0.2024 0.2349 0.2237 0.22545 0.2545	1.2705 0.3434 0.1542 -0.8909 -0.6295	0.2039 0.7313 0.8775 0.3730 0.5290	
intercept Periodl Period2 Period3	Estimate  0.2572 0.0307 0.0345 -0.2267	Standard Error 0.2024 0.2349 0.2237 0.2545	1.2705 0.3434 0.1542 -0.8909 -0.6295	0.2039 0.7313 0.8775 0.3730	
intercept Periodl Period2 Period3 Period4	Estimate 0.2572 0.0807 0.0345 -0.2267 -0.1592	Standard Error 0.2024 0.2349 0.2237 0.22545 0.2545	1.2705 0.3434 0.1542 -0.8909 -0.6295	0.2039 0.7313 0.8775 0.3730 0.5290	

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.

Event Rate Ratio	and 95% Event Ra	te Confidence Int	tervals		_
			Bou	nds	
Parameter	Estimate	Event Rate	Lower	Upper	
intercept	0.2572	1.2933	0.8697	1.9231	
Periodl	0.0807	1.0840	0.6840	1.7180	
Period2	0.0345	1.0351	0.6677	1.6046	
Period3	-0.2267	0.7971	0.4841	1.3127	
Period4	-0.1592	0.8528	0.5195	1.4000	
DrugAsp	0.2151	1.2400	0.8994	1.7096	
Est	imated level 2 va	riances and cova:	riances		
		Stan	dard		
Parameter	Est	imate Erro:	r zV	alue P Val	ıe
intercept/intercep		.4290 0.3	1715 2	5024 0.01	22

The random-effect variance is estimated as 0.429, 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.

### 3.1.2.4 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

$$\log\left(\operatorname{Headache}_{ij}\right) = \hat{\beta}_{0} + \hat{\beta}_{1} \times \operatorname{Period1}_{ij} + \hat{\beta}_{2} \times \operatorname{Period2}_{ij}$$
$$+ \hat{\beta}_{3} \times \operatorname{Period3}_{ij} + \hat{\beta}_{4} \times \operatorname{Period4}_{ij} + \hat{\beta}_{5} \times \operatorname{DrugAsp}_{ij}$$
$$= 0.2572 + 0.0807 \times \operatorname{Period1}_{ij} + 0.0345 \times \operatorname{Period2}_{ij}$$
$$- 0.2267 \times \operatorname{Period3}_{ij} - 0.1592 \times \operatorname{Period4}_{ij} + 0.2151 \times \operatorname{DrugAsp}_{ij}.$$

or, after taking exponents on both sides, as

$$\hat{\text{Headache}}_{ij} = \exp(0.2572 + 0.0807 \times \text{Period1}_{ij} + 0.0345 \times \text{Period2}_{ij}) - 0.2267 \times \text{Period3}_{ij} - 0.1592 \times \text{Period4}_{ij} + 0.2151 \times \text{DrugAsp}_{ij}).$$

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

Headache<sub>ij</sub> = exp
$$(0.2572 + 0.0807 + 0.2151)$$
  
= 1.7385.

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

$$\hat{\text{Headache}}_{ij} = \exp(0.2572 + 0.0345 + 0.2151)$$
  
= 1.6600,  
$$\hat{\text{Headache}}_{ij} = \exp(0.2572 - 0.2267 + 0.2151)$$
  
= 1.2784,

and

$$Headache_{ii} = \exp(0.2752 - 0.1592 + 0.2151)$$
  
= 1.3677

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

#### Estimated outcomes for groups: population-average results

The latent response variable model,

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

makes the assumption that  $e_{ij}$ :  $LID(0, \sigma_e^2)$ . 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

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

The design effect  $d_{ii}$  is defined as

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

which, for the current model, may be calculated as

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

where  $var(v_{i0}) = 0.4290$ , 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  $exp = exp/\sqrt{d_{ii}}$  as shown below.

$$Headache_{ii} = \exp[(0.2572 + 0.0807 \times \text{Period1}_{ii} + 0.0345 \times \text{Period2}_{ii} - 0.2267 \times \text{Period3}_{ii} - 0.1592 \times \text{Period4}_{ii} + 0.2151 \times \text{DrugAsp}_{ii}) / \sqrt{1.4290}].$$

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  $\sqrt{1.4290} = 1.1954$ .

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

 Table 5.3: Estimated unit-specific and population average results for groups