Generalized Mixed Linear Models
Practical 2

Dankmar Böhning

December 3, 2014

Prevalence of upper respiratory tract infection

The data below are taken from a survey on the prevalence of upper respiratory tract infection. The variable to be analysed is the number of swabs positive for pneumococcus during a certain period. Observations were made on 4 members in 18 families, i.e. on two adults and 2 children per family. Six families were a random selection of families living in "overcrowded" conditions, six were in "crowded" conditions and six were in "uncrowded" conditions.

		Family status			
		Adult		Child	
Crowding category	Family serial number	1	2	1	2
Overcrowded	1	5	7	25	19
	2	11	8	33	35
	3	3	12	6	21
	4	3	19	17	17
	5	10	9	11	17
	6	9	0	9	5
Crowded	7	11	7	15	13
	8	10	5	13	17
	9	5	4	18	10
	10	1	9	16	8
	11	5	5	16	20
	12	7	3	17	18
Uncrowded	13	6	3	17	18
	14	9	6	14	10
	15	2	2	15	8
	16	0	2	16	21
	17	3	2	3	14
	18	6	2	7	20

Questions of interest are**how the prevalence of upper respiratory tract infection is related to overcrowding conditions and to family status**. What are random and fixed effects here?

Solution

Family is considered as random effect whereas Family Status and Degree of Crowdedness are considered as fixed effects.

Since the outcome variable is a count (Number of Positive Swabs) we turn to **Mixed Poisson Regression**.

We start by considering Family and Degree of Crowdedness:

```
Mixed-effects Poisson regression                Number of obs      =        72
Group variable: family                          Number of groups   =        18

                                                Obs per group: min =         4
                                                               avg =       4.0
                                                               max =         4

Integration points =   1                        Wald chi2(2)       =      6.16
Log likelihood = -303.14966                     Prob > chi2        =    0.0460

------------------------------------------------------------------------------
   swaps_pos |        IRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    crowding |
          1  |   1.475074   .2310807     2.48   0.013     1.085094    2.005213
          2  |   1.237837   .1961639     1.35   0.178     .9073418    1.688714
             |
       _cons |    8.42314   .9625242    18.65   0.000     6.732959    10.53761
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
family: Identity             |
                   sd(_cons) |   .2202706   .0550381      .1349803    .3594537
------------------------------------------------------------------------------
LR test vs. Poisson regression:  chibar2(01) =    18.96 Prob>=chibar2 = 0.0000

Note: log-likelihood calculations are based on the Laplacian approximation.
```

We see that the random effect Family is needed and that the Overcrowded category has a significantly increased risk ratio (reference is Undercrowded).

We now include Family Status (child/adult):

```
Mixed-effects Poisson regression                Number of obs      =        72
Group variable: family                          Number of groups   =        18
```

```
                                                Obs per group: min =         4
                                                               avg =       4.0
                                                               max =         4

Integration points =   1                        Wald chi2(3)       =    151.57
Log likelihood = -221.58741                     Prob > chi2        =    0.0000

------------------------------------------------------------------------------
   swaps_pos |        IRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    crowding |
          1  |   1.475074   .2310808     2.48   0.013     1.085094    2.005214
          2  |   1.237837   .1961639     1.35   0.178     .9073418    1.688714
             |
       Child |   2.649289   .2140561    12.06   0.000      2.26128    3.103876
       _cons |   4.616318   .5929517    11.91   0.000     3.588901     5.93786
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
family: Identity             |
                   sd(_cons) |   .2202706   .0550381      .1349803    .3594537
------------------------------------------------------------------------------
LR test vs. Poisson regression:  chibar2(01) =    18.96 Prob>=chibar2 = 0.0000

Note: log-likelihood calculations are based on the Laplacian approximation.
```

The Family random effect is still needed and also the Overcrowded categroy remains still significnat. In addition, children show a significantly increased risk for upper respiratory infections if comapred to adults.

Post-operative sore throat study

The aim of a study carried out at the Royal Berkshire Hospital, Reading, in 2004 was to investigate the incidence of sore throat in patients who had undergone orthopaedic, gynaecological, genitourinary or general surgery. Of particular interest was whether the occurrence of a sore throat was affected by the method used to deliver anaesthetic gas, and patients were allocated to one of three types of airway device, namely the laryngeal mask airway (LMA), the endo-tracheal tube (ETT), and the traditional face mask (FM). The decision on which of the three types of device to use for a particular patient was made by the consultant anaesthetist, and there were 12 anaesthetists involved.

The response variable was binary and concerned whether or not a patient experienced a sore throat in the 24 hour period following the operation. The values of certain explanatory variables were also recorded, including the age and sex of the patient, the duration of surgery, and, for LMA and ETT, whether or not the throat was lubricated before the airway was inserted. The following

eight variables are contained in the datafile **sorethroat.dta**.

PATIENT	Patient number (1 - 947)
AGE	Age of patient in years
SEX	Sex of patient (0 = male, 1 = female)
DURATION	Duration of surgery in minutes
AIRWAY	Type of airway used (0 = LMA, 2 = ETT or 1 = FM)
LUBRIC	Lubrication used in inserting mask (0 = no, 1 = yes, . = n/a)
CONSULT	Consultant anaesthetist (1 - 12)
SORE	Occurrence of sore throat (0 = no, 1 = yes)

How do the three types of airway compare in terms of the incidence of post-operative sore throat?

Is there any evidence that the probability that a consultant selects the face mask (FM) is dependent upon the age and sex of the patient or the duration of surgery?

Solution

Evidently, we need to evaluate the risks of FM(1), LMA(0) and ETT(2). We choose ETT as reference (arbitrary). We are now able to give a more satisfactory answer as we can include Consultant as a random effect. We get the following. Clearly, FM has the highest preventive effect.

```
Mixed-effects logistic regression               Number of obs      =       947
Group variable: consultant                      Number of groups   =        12

                                                Obs per group: min =         5
                                                               avg =      78.9
                                                               max =       133

Integration points =   1                        Wald chi2(2)       =     13.74
Log likelihood = -392.27112                     Prob > chi2        =    0.0010

------------------------------------------------------------------------------
        sore | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      airway |
          0  |    .594771    .1759699    -1.76   0.079     .3330519    1.062154
          1  |   .0599768    .0459854    -3.67   0.000      .013346    .2695355
             |
       _cons |   .3269232    .0913365    -4.00   0.000     .1890755    .5652703
```

```
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
consultant: Identity         |
                   sd(_cons) |   2.94e-10   .2642975            0           .
------------------------------------------------------------------------------
LR test vs. logistic regression: chibar2(01) =     0.00 Prob>=chibar2 = 1.0000

Note: log-likelihood calculations are based on the Laplacian approximation.
```

The consultant effect is not significant whereas FM has a high preventive effect where ETT is borderline.

But how is this influenced by other covariate such as gender and age?

```
Mixed-effects logistic regression               Number of obs      =       947
Group variable: consultant                      Number of groups   =        12

                                                Obs per group: min =         5
                                                               avg =      78.9
                                                               max =       133

Integration points =   1                        Wald chi2(4)       =     32.53
Log likelihood = -381.62451                     Prob > chi2        =    0.0000

------------------------------------------------------------------------------
        sore | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      airway |
          0  |   .5463571   .1650394    -2.00   0.045     .3022401    .9876453
          1  |   .0514535   .0396385    -3.85   0.000     .0113676    .2328946
             |
         age |   1.000994   .0061108     0.16   0.871     .9890885    1.013043
         sex |   2.538882   .5415278     4.37   0.000     1.671424    3.856546
       _cons |   .1788752   .0784383    -3.92   0.000      .075734    .4224831
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
consultant: Identity         |
                   sd(_cons) |   1.40e-09   .1551432            0           .
------------------------------------------------------------------------------
LR test vs. logistic regression: chibar2(01) =     0.00 Prob>=chibar2 = 1.0000

Note: log-likelihood calculations are based on the Laplacian approximation.
```

We see that gender is important, but not age. Further analysis shows that also duration is not needed. The final analysis below shows that FM has a high preventive effect whereas ETT is borderline. Women have a significantly increased risk for a sore throat. Further analysis could look for an airway-gender interaction.

```
Logistic regression                               Number of obs   =        947
                                                  LR chi2(3)      =      47.50
                                                  Prob > chi2     =     0.0000
Log likelihood = -381.63776                       Pseudo R2       =     0.0586

------------------------------------------------------------------------------
        sore | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      airway |
         0   |   .5455191   .1647065    -2.01   0.045      .301863     .985848
         1   |   .0514472   .0396343    -3.85   0.000      .011366    .2328718
             |
         sex |   2.535186   .5402145     4.37   0.000     1.669668    3.849369
       _cons |   .1878986   .0595563    -5.27   0.000     .1009548    .3497198
------------------------------------------------------------------------------
```

Finally, we look at the question of FM selection and how this is affected by age and gender. We include consultant as a random effect.

```
Mixed-effects logistic regression               Number of obs      =       947
Group variable: consultant                      Number of groups   =        12

                                                Obs per group: min =         5
                                                               avg =      78.9
                                                               max =       133

Integration points =   1                        Wald chi2(2)       =      8.04
Log likelihood = -259.25083                     Prob > chi2        =    0.0179

------------------------------------------------------------------------------
          FM | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         sex |   .7744962   .2041577    -0.97   0.332         .462    1.298364
         age |   1.020736   .0082147     2.55   0.011     1.004762    1.036965
       _cons |   .0110163   .0095899    -5.18   0.000     .0020001    .0606769
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
consultant: Identity         |
                   sd(_cons) |   2.004827   .6486995      1.063297    3.780065
------------------------------------------------------------------------------
LR test vs. logistic regression: chibar2(01) =   133.69 Prob>=chibar2 = 0.0000

Note: log-likelihood calculations are based on the Laplacian approximation.
```

In contrast to our previous analysis in Practical 1, there is a significant consultant. Also, the age of the patient influences the decision for using a FM but gender does not.