

SAA for many N level multilevel models

Welcome to an SAA for fitting many model types developed for Stat-JR v1.0.5

Input questions

Firstly on this page you will need to specify the dataset required from the list of available datasets.

Which dataset do you wish to use:

Submit

Next you need to choose many options including the response, estimation method, clustering variables and predictor variables (both continuous and categorical) from the chosen dataset. After choosing these variables the SAA will run and you will see a block of text describing how many observations are to be used at the bottom of this page. The rest of the analysis will appear in pages 2-12.

What estimation method do you want to use:

IGLS

What is the response variable:

obs

What distribution are you going to assume:

Poisson

Is there an offset:

Yes

Which column contains the offsets:

logexp

| | |
|--|---------------|
| Please enter your possible (nested) classifications / levels (lowest first, not including level-1): | region,nation |
| Are there any continuous predictors that need including in all models: | No |
| Are there any categorical predictors that need including in all models: | No |
| Do you want to include any continuous predictors as candidates for inclusion in the models: | Yes |
| Which continuous predictors do you want to consider: | uvbi |
| Do you want to include any categorical predictors as candidates for inclusion in the models: | Yes |
| Which categorical predictors do you want to consider: | nation |
| What selection type do you require: | Forward pass |
| Do you want to test for random slopes: | Yes |
| Do you want to test for interactions: | Yes |

The Analysis Assistant you are currently using is designed to work on complete datasets only and so as a pre-processing step we have to remove any rows that contain missing data in columns used in the analysis that follows. For now the list

of columns to be considered is: obs, logexp, region, nation, uvbi, nation. There are 0 (0.0%) rows that get deleted This results in a dataset of 354 rows.

On the next page we will look at the shape of the response and, in the case of normal responses, decide whether to log transform.

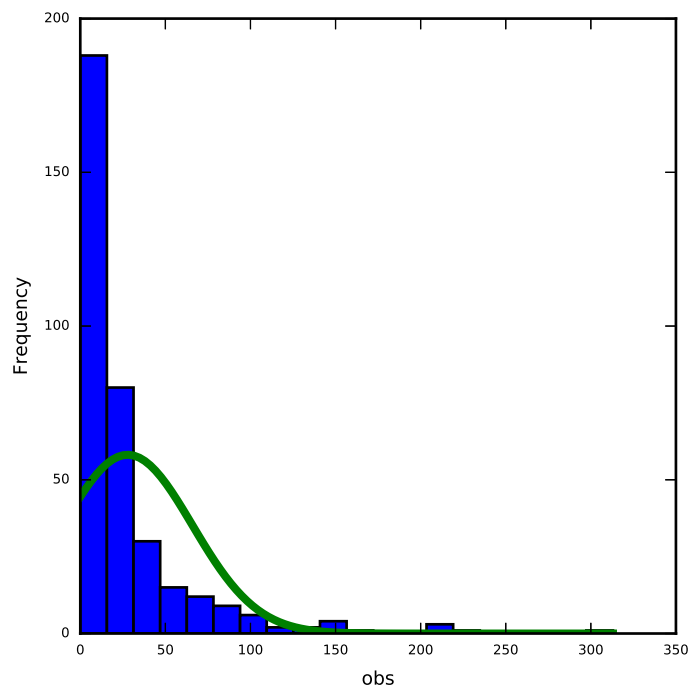
Exploring the response

We will begin our analysis of the dataset by doing some basic data exploration.

You have chosen obs as your response variable and so a first step is to take a look at this variable and assess its suitability for modelling. The summary statistics for the variable are in the table below:

| | |
|---------------------------|--------|
| Observations | 354 |
| Mean | 27.828 |
| Standard Deviation | 38.022 |
| Median | 14.5 |

We also look at a histogram of obs to see what it looks like - noting that for a Poisson model we would expect a skewed distribution as we will fit it with a log link.



Here the median is smaller than the mean and there is significant skew to the right. The skewness value is 3.369 and so this is practically large as well as statistically significant.

The values:

| Row | obs |
|------------|------------|
| 167 | 313 |

are greater than 6 sds from the mean. This might warrant investigating.

Exploring the predictors individually

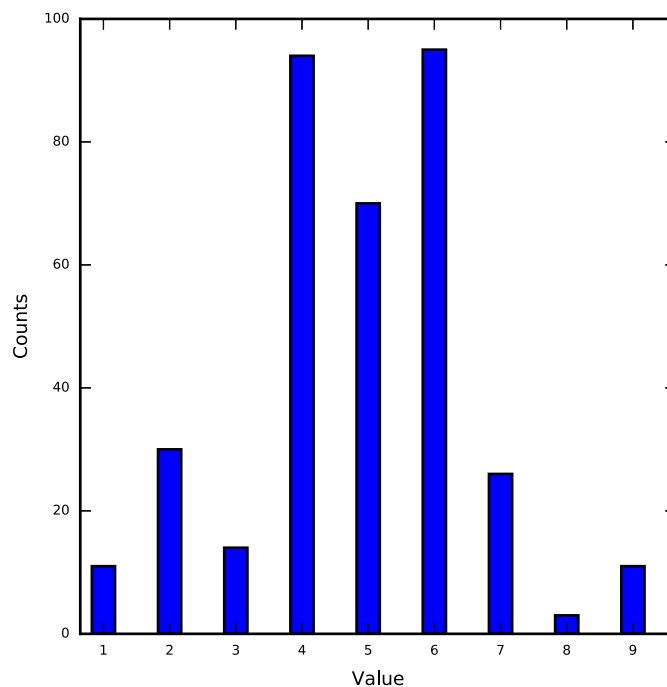
We can also look at each of the predictor variables in turn in isolation.

For categorical predictors we are looking at how common each category is in the dataset. In particular we are checking for rare categories which might cause difficulties in modelling and might therefore be usefully merged with other categories (though this would need to be done outside this SAA).

For predictor nation we see the following:

| nation | N | Percentage |
|---------------|----------|-------------------|
| 1 | 11 | 3.107 |
| 2 | 30 | 8.475 |
| 3 | 14 | 3.955 |
| 4 | 94 | 26.554 |
| 5 | 70 | 19.774 |
| 6 | 95 | 26.836 |
| 7 | 26 | 7.345 |
| 8 | 3 | 0.847 |
| 9 | 11 | 3.107 |
| Total | 354 | 100 |

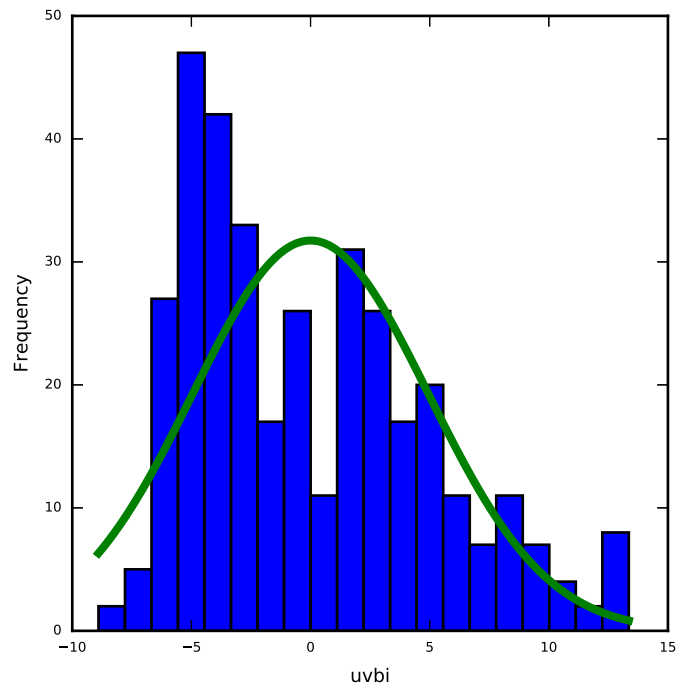
The values 8 are rare categories with fewer than 5 observations.



For continuous predictors we are interested in looking at summary statistics, the shape of the distribution and any unusual values. If the distribution is skewed then we might want to transform the variable before fitting it in the model although it is more important to consider transformations of the response variable and remember what is important is whether the relationship between the response and predictor is linear. If there are unusual values we will want to check that the unusual values are correct and not errors and also whether we may want to treat the variable differently. Another possibility for unusual shaped distributions is to instead categorise the variable into ranges of values.

For predictor uvbi we see the following:

| | |
|---------------------------|--------|
| Name | uvbi |
| Observations | 354 |
| Mean | 0.0 |
| Standard Deviation | 4.952 |
| Median | -0.886 |



Here the median is smaller than the mean and there is significant skew to the right. The skewness value is 0.658. Here the statistical significance may be to some degree due to the large sample size as from a practical perspective values of skew less than 2 in absolute magnitude are not considered too big a skew.

There are no obvious outliers in uvbi.

Assessing the relationship between the response and individual predictors

Once we are happy with our response variable and our set of predictors we now want to have a preliminary look at them together before progressing to the univariable modelling.

For the categorical predictors it is worth looking at the mean value of the response in each category to assess if there are differences. We can then formally test this with a t-test for binary predictors or an ANOVA for predictors with more than 2 categories.

Here is a tabulation of the response, obs for predictor nation with category 2 having the largest mean and category 7 the smallest.

| Category | N | Mean | Standard Deviation | Median |
|-----------------|----------|-------------|---------------------------|---------------|
| 1 | 11 | 40.82 | 28.52 | 33.0 |
| 2 | 30 | 98.3 | 52.94 | 87.5 |
| 3 | 14 | 48.64 | 48.04 | 36.0 |
| 4 | 94 | 15.9 | 13.08 | 12.0 |
| 5 | 70 | 31.13 | 39.34 | 23.5 |
| 6 | 95 | 15.39 | 18.78 | 10.0 |
| 7 | 26 | 2.577 | 4.03 | 1.5 |
| 8 | 3 | 7.667 | 8.014 | 2.0 |
| 9 | 11 | 49.64 | 41.72 | 32.0 |

The formal test is as follows:

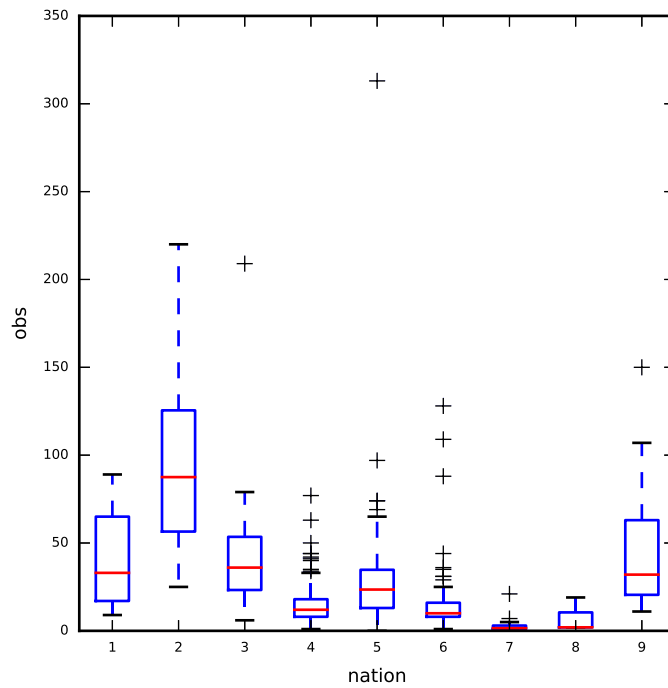
| | df | SS | MS | F |
|---|-----------|-----------|-----------|----------|
| Between groups | 8 | 208800.0 | 26100.0 | 29.71 |
| Within groups | 345 | 303000.0 | 878.3 | |
| Total | 353 | 511800.0 | 1450.0 | |
| Pooled within-group S.D. | 29.64 | | | |
| Between-group variance component | 710.4 | | | |

For the ANOVA we are testing whether there are differences in the means of the response variable between the different groups. As shown in the table above this is done by constructing an ANOVA table that compares how much of the variability in the data is within the groups compared to between the groups. This results in a test statistic that follows an F distribution with 8 and 345 degrees of freedom. This F table has values of 2.229 for $p=0.05$ and 2.809 for $p=0.01$.

As $29.713 > 2.809$ our p value is less than 0.01 and we have strong evidence to reject the null hypothesis (at the $p=0.01$ level).

The p-value is in fact less than 0.0001.

| Category | N | Mean | S.E.M. |
|-----------------|----------|-------------|---------------|
| 1 | 11 | 40.82 | 8.935 |
| 2 | 30 | 98.3 | 5.411 |
| 3 | 14 | 48.64 | 7.92 |
| 4 | 94 | 15.9 | 3.057 |
| 5 | 70 | 31.13 | 3.542 |
| 6 | 95 | 15.39 | 3.041 |
| 7 | 26 | 2.577 | 5.812 |
| 8 | 3 | 7.667 | 17.11 |
| 9 | 11 | 49.64 | 8.935 |

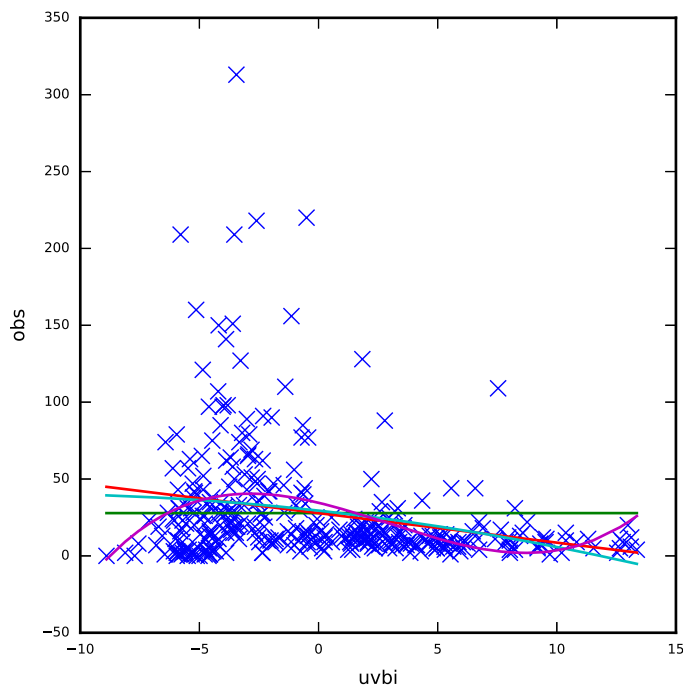


For the continuous predictors we can look at correlations with the response and scatterplots to see if there is a linear relationship.

Predictor : uvbi

The Pearson correlation between obs and uvbi is -0.25 (p value < 0.001).

The Spearman rank correlation between obs and uvbi is -0.23 (p value < 0.001).



The graph includes best fitting curves for a constant, linear, quadratic and cubic relationship between obs and uvbi. In this case a linear relationship is most appropriate.

Choosing appropriate random classifications

We begin this section by deciding which of the possible random classifications to include in the modelling.

This is done by fitting combinations in turn and picking more complicated models if they make a significant improvement via a Wald test. All models are displayed along with their chi-squared test statistic in the table below:

| Higher-level classifications | Significance |
|-------------------------------------|---------------------|
| region | < 0.001 |
| region,nation | 0.05 |

The best model based on the Likelihood has levels: region

As this is a multilevel modelling SAA we will also want to look at how the response is distributed across the levels of the model.

For this we will use the best model chosen above and look at how the variance is distributed across levels.

| Variable | Coefficient | SE |
|------------------------|--------------------|-----------|
| Intercept | 0.11 | 0.151 |
| region Variance | 0.0454 | 0.00964 |

It is difficult to calculate the VPC for Poisson models so we have not done this here.

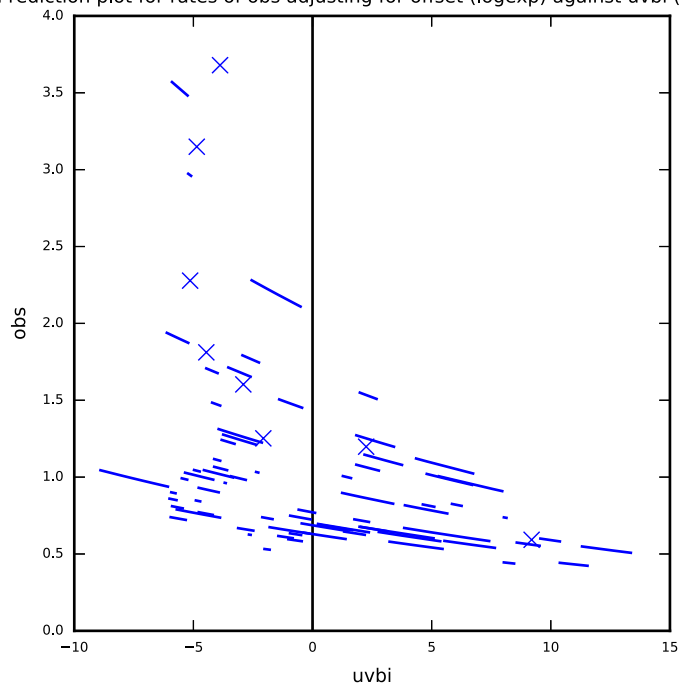
Performing univariable modelling

Our next step in modelling now that we have a set of potential predictors is to consider models for each predictor in turn along with a random intercept at each chosen classification from the best model in the last section. In the fixed part these models simply contain an intercept and the particular predictor and so for continuous predictors will be multilevel linear regressions and for categorical predictors will be multilevel generalisations of ANOVAs. In the table below we summarise the modelling by showing the coefficients for each predictor along with the p value comparing the model with that predictor with a Null model. This Univariable modelling step will identify a set of candidate predictors to be taken forward into the next stage of modelling.

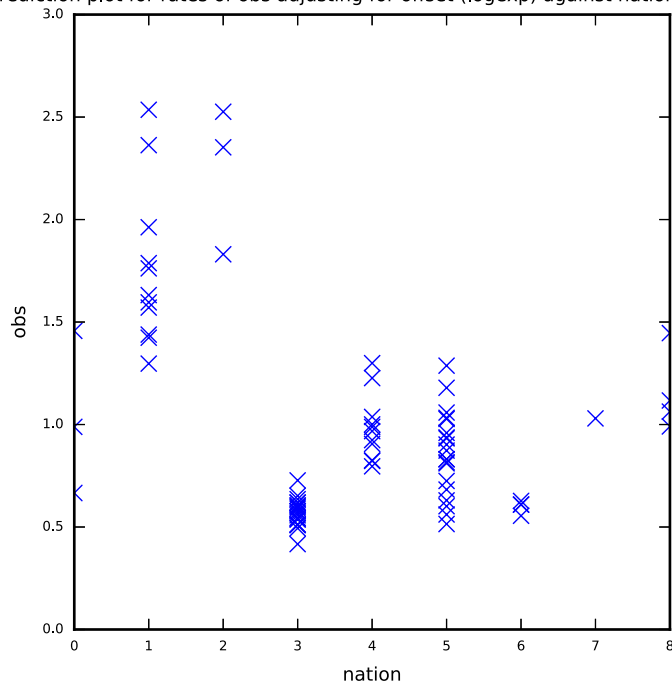
| Variable | Coefficient | SE | p value | Significance |
|-----------------|--------------------|-----------|----------------|---------------------|
| uvbi | -0.038 | 0.00995 | < 0.001 | *** |
| nation_2 | 0.559 | 0.15 | < 0.001 | *** |
| nation_3 | 0.81 | 0.186 | | |
| nation_4 | -0.553 | 0.144 | | |
| nation_5 | -0.0165 | 0.15 | | |
| nation_6 | -0.165 | 0.145 | | |
| nation_7 | -0.498 | 0.218 | | |
| nation_8 | 0.0434 | 0.325 | | |
| nation_9 | 0.147 | 0.177 | | |

Which predictors we consider for the next stage of analysis will depend on their significance in the above table (but may in practice also depend on the size the effect and substantive interest of the variable though this is hard to automate). We will use a threshold on the p values associated with the predictors to decide whether to include the predictors in the next stage. Here we are currently using a threshold of 0.05. so the predictors to carry forward are: uvbi, and nation.

Prediction plot for rates of obs adjusting for offset (logexp) against uvbi (region)



Prediction plot for rates of obs adjusting for offset (logexp) against nation (region)



Looking at correlations between predictors

Our next step is to check that none of the correlations between the predictor variables are too great as this could cause estimation problems when we add the predictors to the model together. To do this we look at all correlations between the predictor variables that have been identified as significant univariably and are thus candidates to be added to the model.

The correlations are as follows:

| Variables | Correlation |
|----------------------|--------------------|
| (nation_2, uvbi) | -0.172 |
| (nation_3, uvbi) | -0.231 |
| (nation_3, nation_2) | -0.062 |
| (nation_4, uvbi) | 0.194 |
| (nation_4, nation_2) | -0.183 |
| (nation_4, nation_3) | -0.122 |
| (nation_5, uvbi) | -0.469 |
| (nation_5, nation_2) | -0.151 |
| (nation_5, nation_3) | -0.101 |
| (nation_5, nation_4) | -0.299 |
| (nation_6, uvbi) | 0.716 |
| (nation_6, nation_2) | -0.184 |
| (nation_6, nation_3) | -0.123 |
| (nation_6, nation_4) | -0.364 |
| (nation_6, nation_5) | -0.301 |
| (nation_7, uvbi) | -0.287 |
| (nation_7, nation_2) | -0.086 |
| (nation_7, nation_3) | -0.057 |
| (nation_7, nation_4) | -0.169 |
| (nation_7, nation_5) | -0.14 |
| (nation_7, nation_6) | -0.171 |
| (nation_8, uvbi) | -0.044 |

| Variables | Correlation |
|----------------------|--------------------|
| (nation_8, nation_2) | -0.028 |
| (nation_8, nation_3) | -0.019 |
| (nation_8, nation_4) | -0.056 |
| (nation_8, nation_5) | -0.046 |
| (nation_8, nation_6) | -0.056 |
| (nation_8, nation_7) | -0.026 |
| (nation_9, uvbi) | -0.152 |
| (nation_9, nation_2) | -0.054 |
| (nation_9, nation_3) | -0.036 |
| (nation_9, nation_4) | -0.108 |
| (nation_9, nation_5) | -0.089 |
| (nation_9, nation_6) | -0.108 |
| (nation_9, nation_7) | -0.05 |
| (nation_9, nation_8) | -0.017 |

Correlations greater than 0.8 (in magnitude) are worth looking at as they may result in model fitting problems when both predictors are included.

Performing multivariable model selection - random intercept models

In this next stage we will look at the best random intercepts model using only main effects for the variables to be considered. You have chosen to perform forward pass which is a quicker method than full forward selection. It may therefore not explore as many possible models. The predictor variables are considered in turn based on their significance in the univariable analysis and each is added to the current model. If the resulting model is a significant improvement then the predictor is kept in the model otherwise it is removed. Attention then moves on to the next predictor until all predictors are considered.

You have chosen to use Wald tests to compare models. These work by looking at estimates and standard error matrices for each predictor to assess significance and run quicker than the alternative methods as they do not need to run submodels.

The most significant predictor in the univariable analysis was nation so our starting point in multivariable modelling is the model:

$$\text{obs}_i \sim \text{Poisson}(p_i), \log(p_i) = \log \exp + \beta_0 \text{nation_2}_i + \beta_1 \text{nation_3}_i + \beta_2 \text{nation_4}_i + \beta_3 \text{nation_5}_i + \beta_4 \text{nation_6}_i + \beta_5 \text{nation_7}_i + \beta_6 \text{nation_8}_i + \beta_7 \text{nation_9}_i + \beta_8 \text{intercept}_i + u_{0, \text{region}_i}^{(2)}$$

| | Variable | Coefficient | SE | p value | Significance |
|--|--------------------------------|--------------------|-----------|----------------|---------------------|
| | nation_2 | 0.559 | 0.15 | < 0.001 | *** |
| | nation_3 | 0.81 | 0.186 | | |
| | nation_4 | -0.553 | 0.144 | | |
| | nation_5 | -0.0165 | 0.15 | | |
| | nation_6 | -0.165 | 0.145 | | |
| | nation_7 | -0.498 | 0.218 | | |
| | nation_8 | 0.0434 | 0.325 | | |
| | nation_9 | 0.147 | 0.177 | | |
| | Intercept | -0.014 | 0.133 | | |
| | Between region Variance | 0.0444 | 0.00937 | | |

Adding variable nation was a significant improvement and so we retain it in the model.

Our next step is to consider adding variable uvbi to the current model.

$$\text{obs}_i \sim \text{Poisson}(p_i), \log(p_i) = \log \exp + \beta_0 \text{nation_2}_i + \beta_1 \text{nation_3}_i + \beta_2 \text{nation_4}_i + \beta_3 \text{nation_5}_i + \beta_4 \text{nation_6}_i + \beta_5 \text{nation_7}_i + \beta_6 \text{nation_8}_i + \beta_7 \text{nation_9}_i + \beta_8 \text{uvbi}_i + \beta_9 \text{intercept}_i + u_{0, \text{region}_i}^{(2)}$$

| Variable | Coefficient | SE | p value | Significance |
|-------------------------|-------------|---------|---------|--------------|
| nation_2 | 0.553 | 0.141 | < 0.001 | *** |
| nation_3 | 0.741 | 0.177 | | |
| nation_4 | -0.437 | 0.142 | | |
| nation_5 | -0.0541 | 0.142 | | |
| nation_6 | 0.0658 | 0.167 | | |
| nation_7 | -0.545 | 0.21 | | |
| nation_8 | 0.0658 | 0.312 | | |
| nation_9 | 0.12 | 0.167 | | |
| uvbi | -0.0273 | 0.0109 | | |
| Intercept | -0.0992 | 0.129 | | |
| Between region Variance | 0.0383 | 0.00831 | | |

Adding variable uvbi was a significant improvement and so we retain it in the model.

This is our final model.

Choosing interactions

In this section we add to the best random intercepts model with main effects found in the last section. Here we consider all possible pairwise interactions between the significant predictors already found including quadratic terms for predictors. The model selection methods used are as for the previous best random intercepts models.

$$\begin{aligned} \text{obs}_i \sim & \text{Poisson}(p_i), \log(p_i) = \log\text{exp} + \beta_0 \text{nation_2}_i + \beta_1 \text{nation_3}_i \\ & + \beta_2 \text{nation_4}_i + \beta_3 \text{nation_5}_i + \beta_4 \text{nation_6}_i + \beta_5 \text{nation_7}_i + \beta_6 \text{nation_8}_i \\ & + \beta_7 \text{nation_9}_i + \beta_8 \text{uvbi}_i + \beta_9 \text{uvbi_X_nation_2}_i + \beta_{10} \text{uvbi_X_nation_3}_i \\ & + \beta_{11} \text{uvbi_X_nation_4}_i + \beta_{12} \text{uvbi_X_nation_5}_i + \beta_{13} \text{uvbi_X_nation_6}_i \\ & + \beta_{14} \text{uvbi_X_nation_7}_i + \beta_{15} \text{uvbi_X_nation_8}_i + \beta_{16} \text{uvbi_X_nation_9}_i \\ & + \beta_{17} \text{intercept}_i + u_{0, \text{region}_i}^{(2)} \end{aligned}$$

| Variable | Coefficient | SE | p value | Significance |
|--------------------------------|--------------------|-----------|----------------|---------------------|
| nation_2 | -0.246 | 0.681 | < 0.001 | *** |
| nation_3 | -0.316 | 1.052 | | |
| nation_4 | -1.289 | 0.674 | | |
| nation_5 | -0.109 | 0.699 | | |
| nation_6 | -0.412 | 0.679 | | |
| nation_7 | -1.19 | 1.429 | | |
| nation_8 | 14.02 | 15.41 | | |
| nation_9 | -1.023 | 1.075 | | |
| uvbi | 0.252 | 0.227 | | |
| uvbi_X_nation_2 | -0.275 | 0.229 | < 0.001 | *** |
| uvbi_X_nation_3 | -0.323 | 0.268 | | |
| uvbi_X_nation_4 | -0.241 | 0.228 | | |
| uvbi_X_nation_5 | -0.115 | 0.231 | | |
| uvbi_X_nation_6 | -0.339 | 0.228 | | |
| uvbi_X_nation_7 | -0.246 | 0.341 | | |
| uvbi_X_nation_8 | 6.156 | 6.723 | | |
| uvbi_X_nation_9 | -0.361 | 0.304 | | |
| Intercept | 0.709 | 0.672 | | |
| Between region Variance | 0.0256 | 0.00612 | | |

Adding variable uvbi_X_nation significantly improved the model and so is retained in the model.

Our next step is to consider adding variable uvbi_X_uvbi to the current model.

$$\begin{aligned}
\text{obs}_i \sim & \text{Poisson}(p_i), \log(p_i) = \text{logexp} + \beta_0 \text{nation_2}_i + \beta_1 \text{nation_3}_i \\
& + \beta_2 \text{nation_4}_i + \beta_3 \text{nation_5}_i + \beta_4 \text{nation_6}_i + \beta_5 \text{nation_7}_i + \beta_6 \text{nation_8}_i \\
& + \beta_7 \text{nation_9}_i + \beta_8 \text{uvbi}_i + \beta_9 \text{uvbi_X_nation_2}_i + \beta_{10} \text{uvbi_X_nation_3}_i \\
& + \beta_{11} \text{uvbi_X_nation_4}_i + \beta_{12} \text{uvbi_X_nation_5}_i + \beta_{13} \text{uvbi_X_nation_6}_i \\
& + \beta_{14} \text{uvbi_X_nation_7}_i + \beta_{15} \text{uvbi_X_nation_8}_i + \beta_{16} \text{uvbi_X_nation_9}_i \\
& + \beta_{17} \text{uvbi_X_uvbi}_i + \beta_{18} \text{intercept}_i + u_{0, \text{region}_i}^{(2)} \Big|
\end{aligned}$$

| Variable | Coefficient | SE | p value | Significance |
|-------------------------|-------------|---------|---------|--------------|
| nation_2 | -0.246 | 0.681 | < 0.001 | *** |
| nation_3 | -0.316 | 1.054 | | |
| nation_4 | -1.289 | 0.675 | | |
| nation_5 | -0.109 | 0.7 | | |
| nation_6 | -0.412 | 0.682 | | |
| nation_7 | -1.19 | 1.429 | | |
| nation_8 | 14.02 | 15.41 | | |
| nation_9 | -1.023 | 1.076 | | |
| uvbi | 0.252 | 0.228 | | |
| uvbi_X_nation_2 | -0.275 | 0.229 | 0.011 | * |
| uvbi_X_nation_3 | -0.323 | 0.269 | | |
| uvbi_X_nation_4 | -0.241 | 0.23 | | |
| uvbi_X_nation_5 | -0.115 | 0.231 | | |
| uvbi_X_nation_6 | -0.339 | 0.236 | | |
| uvbi_X_nation_7 | -0.247 | 0.342 | | |
| uvbi_X_nation_8 | 6.156 | 6.723 | | |
| uvbi_X_nation_9 | -0.361 | 0.304 | | |
| uvbi_X_uvbi | -1.17e-05 | 0.00319 | | |
| Intercept | 0.709 | 0.673 | | |
| Between region Variance | 0.0256 | 0.00612 | | |

Adding variable uvbi_X_uvbi did not significantly improve the model, so we remove it from the model.

We have considered all interaction variables so now run our final model.

$$\begin{aligned}
\text{obs}_i &\sim \text{Poisson}(p_i), \log(p_i) = \text{logexp} + \beta_0 \text{nation_2}_i + \beta_1 \text{nation_3}_i \\
&+ \beta_2 \text{nation_4}_i + \beta_3 \text{nation_5}_i + \beta_4 \text{nation_6}_i + \beta_5 \text{nation_7}_i + \beta_6 \text{nation_8}_i \\
&+ \beta_7 \text{nation_9}_i + \beta_8 \text{uvbi}_i + \beta_9 \text{uvbi_X_nation_2}_i + \beta_{10} \text{uvbi_X_nation_3}_i \\
&+ \beta_{11} \text{uvbi_X_nation_4}_i + \beta_{12} \text{uvbi_X_nation_5}_i + \beta_{13} \text{uvbi_X_nation_6}_i \\
&+ \beta_{14} \text{uvbi_X_nation_7}_i + \beta_{15} \text{uvbi_X_nation_8}_i + \beta_{16} \text{uvbi_X_nation_9}_i \\
&+ \beta_{17} \text{intercept}_i + u_{0,region_i}^{(2)}
\end{aligned}$$

| Variable | Coefficient | SE | p value | Significance |
|--------------------------------|--------------------|-----------|----------------|---------------------|
| nation_2 | -0.246 | 0.681 | < 0.001 | *** |
| nation_3 | -0.316 | 1.052 | | |
| nation_4 | -1.289 | 0.674 | | |
| nation_5 | -0.109 | 0.699 | | |
| nation_6 | -0.412 | 0.679 | | |
| nation_7 | -1.19 | 1.429 | | |
| nation_8 | 14.02 | 15.41 | | |
| nation_9 | -1.023 | 1.075 | | |
| uvbi | 0.252 | 0.227 | | |
| uvbi_X_nation_2 | -0.275 | 0.229 | < 0.001 | *** |
| uvbi_X_nation_3 | -0.323 | 0.268 | | |
| uvbi_X_nation_4 | -0.241 | 0.228 | | |
| uvbi_X_nation_5 | -0.115 | 0.231 | | |
| uvbi_X_nation_6 | -0.339 | 0.228 | | |
| uvbi_X_nation_7 | -0.246 | 0.341 | | |
| uvbi_X_nation_8 | 6.156 | 6.723 | | |
| uvbi_X_nation_9 | -0.361 | 0.304 | | |
| Intercept | 0.709 | 0.672 | | |
| Between region Variance | 0.0256 | 0.00612 | | |

This is our final model.

Adding random slopes

Having found a best model that only includes random intercepts we now investigate random slopes for significant predictor variables in the model. Here we use a simple forward pass method to look at each possible random slope in turn using the same comparison method as chosen for earlier models.

The most significant predictor in the univariable analysis was nation so our starting point in adding in random slopes is the model:

nation does not vary at the region level, so we will not attempt to add a random slope for it.

Our next step is to consider adding random slopes for the variable uvbi at the region level to the current model.

$$\begin{aligned} \text{obs}_i \sim \text{Poisson}(p_i), \log(p_i) = & \log\text{exp} + \beta_0 \text{nation_2}_i + \beta_1 \text{nation_3}_i \\ & + \beta_2 \text{nation_4}_i + \beta_3 \text{nation_5}_i + \beta_4 \text{nation_6}_i + \beta_5 \text{nation_7}_i + \beta_6 \text{nation_8}_i \\ & + \beta_7 \text{nation_9}_i + \beta_8 \text{uvbi}_i + \beta_9 \text{uvbi_X_nation_2}_i + \beta_{10} \text{uvbi_X_nation_3}_i \\ & + \beta_{11} \text{uvbi_X_nation_4}_i + \beta_{12} \text{uvbi_X_nation_5}_i + \beta_{13} \text{uvbi_X_nation_6}_i \\ & + \beta_{14} \text{uvbi_X_nation_7}_i + \beta_{15} \text{uvbi_X_nation_8}_i + \beta_{16} \text{uvbi_X_nation_9}_i \\ & + \beta_{17} \text{intercept}_i + u_{0,region_i}^{(2)} + u_{1,region_i}^{(2)} \text{uvbi}_i \end{aligned}$$

| Variable | Coefficient | SD | p value | Significance |
|--|--------------------|-----------|----------------|---------------------|
| Intercept | 0.651 | 0.671 | | |
| nation_2 | -0.179 | 0.678 | | |
| nation_3 | -0.406 | 1.067 | | |
| nation_4 | -1.226 | 0.672 | | |
| nation_5 | 0.00799 | 0.698 | | |
| nation_6 | -0.357 | 0.676 | | |
| nation_7 | -1.169 | 1.459 | | |
| nation_8 | 14.08 | 15.41 | | |
| nation_9 | -1.037 | 1.121 | | |
| uvbi | 0.232 | 0.228 | | |
| uvbi_X_nation_2 | -0.251 | 0.23 | | |
| uvbi_X_nation_3 | -0.331 | 0.271 | | |
| uvbi_X_nation_4 | -0.222 | 0.228 | | |
| uvbi_X_nation_5 | -0.0816 | 0.232 | | |
| uvbi_X_nation_6 | -0.318 | 0.228 | | |
| uvbi_X_nation_7 | -0.234 | 0.347 | | |
| uvbi_X_nation_8 | 6.176 | 6.723 | | |
| uvbi_X_nation_9 | -0.358 | 0.315 | | |
| region Variance(intercept) | 0.0172 | 0.00738 | | |
| region Covariance(intercept,uvbi) | -0.000976 | 0.000828 | | |
| region Variance(uvbi) | 0.000393 | 0.000326 | 0.363 | N/S |

| Variable | Coefficient | SD | p value | Significance |
|------------------|-------------|----------|---------|--------------|
| Level 1 Variance | 1.0 | 9.35e-18 | | |

Variable uvbi did not show a significant random slope, so we remove it from the random part of the model.

We have considered all predictor variables so now run our final random slopes model.

$$\begin{aligned}
\text{obs}_i \sim \text{Poisson}(p_i), \log(p_i) = & \log\text{exp} + \beta_0 \text{nation_2}_i + \beta_1 \text{nation_3}_i \\
& + \beta_2 \text{nation_4}_i + \beta_3 \text{nation_5}_i + \beta_4 \text{nation_6}_i + \beta_5 \text{nation_7}_i + \beta_6 \text{nation_8}_i \\
& + \beta_7 \text{nation_9}_i + \beta_8 \text{uvbi}_i + \beta_9 \text{uvbi_X_nation_2}_i + \beta_{10} \text{uvbi_X_nation_3}_i \\
& + \beta_{11} \text{uvbi_X_nation_4}_i + \beta_{12} \text{uvbi_X_nation_5}_i + \beta_{13} \text{uvbi_X_nation_6}_i \\
& + \beta_{14} \text{uvbi_X_nation_7}_i + \beta_{15} \text{uvbi_X_nation_8}_i + \beta_{16} \text{uvbi_X_nation_9}_i \\
& + \beta_{17} \text{intercept}_i + u_{0, \text{region}_i}^{(2)}
\end{aligned}$$

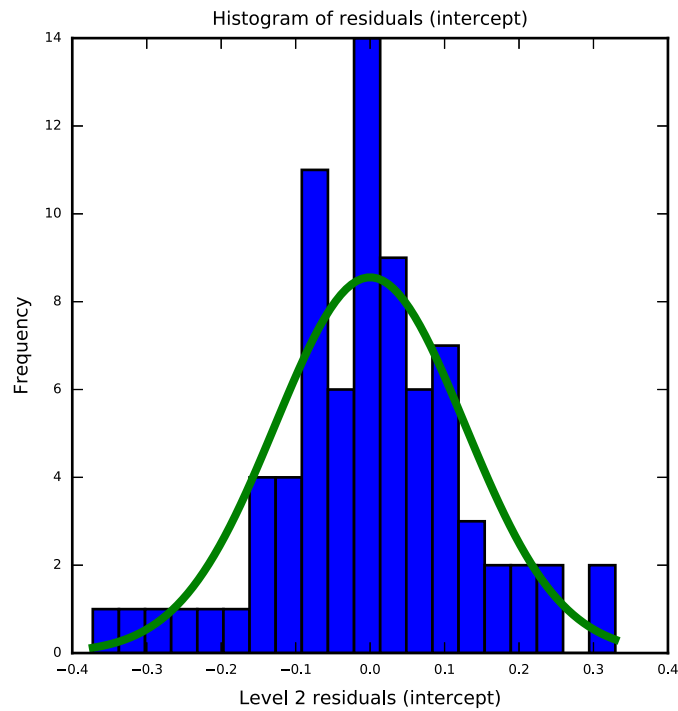
| Variable | Coefficient | SD | p value | Significance |
|-----------------------------------|--------------------|-----------|----------------|---------------------|
| Intercept | 0.709 | 0.672 | | |
| nation_2 | -0.246 | 0.681 | | |
| nation_3 | -0.316 | 1.052 | | |
| nation_4 | -1.289 | 0.674 | | |
| nation_5 | -0.109 | 0.699 | | |
| nation_6 | -0.412 | 0.679 | | |
| nation_7 | -1.19 | 1.429 | | |
| nation_8 | 14.02 | 15.41 | | |
| nation_9 | -1.023 | 1.075 | | |
| uvbi | 0.252 | 0.227 | | |
| uvbi_X_nation_2 | -0.275 | 0.229 | | |
| uvbi_X_nation_3 | -0.323 | 0.268 | | |
| uvbi_X_nation_4 | -0.241 | 0.228 | | |
| uvbi_X_nation_5 | -0.115 | 0.231 | | |
| uvbi_X_nation_6 | -0.339 | 0.228 | | |
| uvbi_X_nation_7 | -0.246 | 0.341 | | |
| uvbi_X_nation_8 | 6.156 | 6.723 | | |
| uvbi_X_nation_9 | -0.361 | 0.304 | | |
| region Variance(intercept) | 0.0256 | 0.00612 | | |
| Level 1 Variance | 1.0 | 0.0 | | |

This is our final random slopes model.

Analysing the residuals

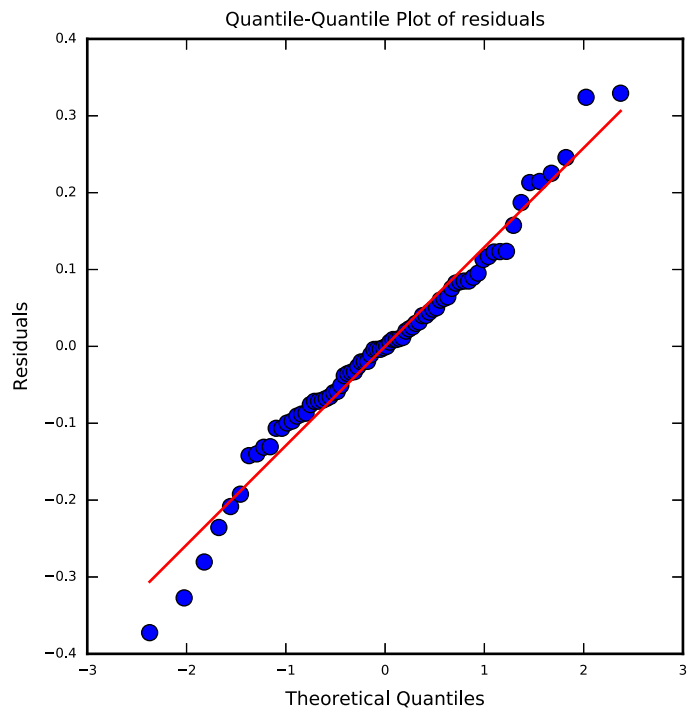
Here we look at the residuals from the model and plot them in various ways.

Next the level 2 residuals for intercept:



Here the distribution is reasonably symmetric with skewness value -0.105.

There are no obvious outliers in the residuals.



If the residuals are fairly normally distributed then the points in this graph should be close to the red line.

Looking at predictions

Having fitted a model with several predictors we might like to represent this model graphically. This is more difficult than when we have only one predictor and so for now we consider each predictor in turn and set all other predictors to their mean values.

