# Generalized Linear Mixed Effects Models

Generalized Linear Mixed Models (GLMMs) are an extension of linear mixed models to allow response variables from different distributions (such as binary or multi-nomial responses). Think of it as an extension of generalized linear models (e.g. logistic regression) to include both fixed and random effects.

- The general form of the model is: $\mathbf{y_{ij} = X_{ij}\,\beta + Z_{ij}\,b_i + \varepsilon}$
  - **y** (or sometimes η) is a N*1 column vector of the dependent (outcome) variable
  - **X** is a N*p of the p predictor variables
  - **Z** is the N*q design matrix for q random effects (the random complement to the fixed X)
  - **b** (or sometimes **u**) is a q*1 vector of the random effects
  - **ε** is an N*1 column of vector of the residuals (the part not explained by **Xβ + Zu**)

$$\overbrace{\mathbf{y}}^{\mathrm{N\,x\,1}} = \overbrace{\underbrace{\mathbf{X}}_{\mathrm{N\,x\,p}}\ \underbrace{\boldsymbol{\beta}}_{\mathrm{p\,x\,1}}}^{\mathrm{N\,x\,1}} + \overbrace{\underbrace{\mathbf{Z}}_{\mathrm{N\,x\,q}}\ \underbrace{\mathbf{u}}_{\mathrm{q\,x\,1}}}^{\mathrm{N\,x\,1}} + \overbrace{\boldsymbol{\varepsilon}}^{\mathrm{N\,x\,1}}$$

In classical statistics we do not actually estimate the vector of random effects; we nearly always assume that for the jth element of vector $u_j \sim N(0, G)$; Where G is the variance-covariance matrix of the random effects. Recall that the variance-covariance is always square, symmetric and positive semi-definite; This means for a q*q matrix there are q(q+1)/2 unique elements.

Because we directly estimated the fixed effects the random effect complements (Z) are modeled as deviations from the fixed effects with mean 0. The random effects are just deviations around the value in β (which is the mean). The only thing left to estimate is the variance. In a model with only a random intercept G is a 1*1 matrix (the variance of the random intercept). If we had a random intercept and a slope G would look like:

$$\mathbf{G} = \begin{bmatrix} \sigma^2_{int} & \sigma^2_{int,slope} \\ \sigma^2_{int,slope} & \sigma^2_{slope} \end{bmatrix}$$

So far everything we've covered applied to both linear mixed models and generalized linear mixed models. What's different about these is that in GLMM the response variables can come from different distributions besides Gaussian. Additionally, rather than modeling the responses directly some **link function** (referred toi as g(.) and its inverse as h()) is often applied, such as log link. For example, in count outcomes are a special case in which we use a log link function and the

probability mass function:

$$g(\cdot) = \cdot$$
$$h(\cdot) = \cdot$$
$$g(\cdot) = h(\cdot)$$
$$g(E(X)) = E(X) = \mu$$
$$g(Var(X)) = Var(X) = \Sigma^2$$
$$PDF(X) = \left(\frac{1}{\Sigma\sqrt{2\pi}}\right) e^{\frac{-(x-\mu)^2}{2\Sigma^2}}$$

# Interpretation

The interpretation of GLMMs is similar to GLMs. Often we want to transform our results to the original rather than the link function result;  this is where transformations complicate matters because they are nonlinear and so even random intercepts no longer. Consider the example below of a mixed effects logistic model predicting remission:

*Mixed Effects Logistic for Remission Status*

| Parameter | Est. | SE | p-value | OR |
|---|---|---|---|---|
| Intercept | 1.467 | .274 | <.001 | 4.335 |
| Age | -.056 | .005 | <.001 | .946 |
| Married (yes v no) | .26 | .064 | <.001 | 1.297 |
| IL6 | -.053 | .011 | <.001 | .948 |
| $\Sigma^2_{intercept}$ | 3.979 | | | |
| $N_{patients}$ = 8,525 | | $N_{doctors}$ = 407 | | |

The estimates can be interpreted as always, where the estimated effect of the parameter is interpreted as a change in the log odds. However, we run into issues trying to interpret the odds ratios. Odds ratios take on a more complex meaning when there are mixed effects, as in a regular logistic model we assume all other effects are fixed. Thus we must interpret the odds ratio here as a conditional odds ratio for holding the remaining factors constant.

# Estimation

For parameter estimation, there are no closed form solutions for GLMMs you must use some approximation

- **Gaussian quadrature** is a particularly well-suited method to numerically evaluate integrals against probability measures.
  - Suppose we have a probability density function p(x) and the function to be integrated against it f(x), the quadrature rule:
    $$\int_{-\infty}^{\infty} f(x)p(x) \approx \sum_{i=1}^{N} w_i f(x_i)$$
    where N denotes the number of quadrature points where w is the quadrature weights and x is the abscissas (locations).

- The Gaussian quadrature chooses abscissas in areas of high density, and if it is continuous the quadrille rule is exact is f(x) is a polynomial of up to degree 2R - 1
- In the generalized linear mixed model the roles of f(x) and p(x) are played by the conditional distribution of the data given by the random effects and random effects distribution
- Quadrature abscissas and weights are those of the standard Gauss-Hermite Quadrature

- **Gauss-Hermite Quadrature** is appropriate when the density is:

$$\int_{-\infty}^{+\infty} e^{-x^2} f(x)\,dx \approx \sum_{i=1}^{n} w_i f(x_i)$$

  - First we change the variables of integration to independent standard normal random effects v using the Cholesky decomposition L such that G = LL' so that $b_i = Lv_i$; We re-write the log likelihood:

$$\ell(\boldsymbol{\beta}, \mathsf{G}) = \ln \prod_{i=1}^{N} \int_{-\infty}^{\infty} \phi(v_{iq}) \ldots \left[ \int_{-\infty}^{\infty} \phi(v_{i1}) \prod_{i=1}^{n_i} f(y_{ij}|v_i) dv_{i1} \right] \ldots dv_{iq}$$

    where phi(.) is the univariate standard normal density.
  - Each uni-dimensional integral can be approximated by Gauss-Hermite quadrature as:

$$\int_{-\infty}^{\infty} \phi(v_{i1}) \prod_{i=1}^{n_i} f(y_{ij}|v_i) dv_{i1} \approx \sum_{r=1}^{R} w_r \prod_{j=1}^{n_i} f(y_{ij}|a_r, v_{i2}, \ldots, v_{iq}))$$

    where sqrt(pi)*wr and ar/sqrt(2) are the weights and abscissas of the (2R-1)th degree Gauss-Hermite quadrature rule
  - The approximation becomes more accurate as more quadrature points R are used.
  - This can be computationally expensive since the number of terms required to evaluate each multivariate integral is $R^q$

Gauss-Hermite quadrature has limitations; The number of function evaluations required grows exponentially as the number of dimensions increases. A random intercept is one dimension, adding a random slope would be two. For three level models with random intercepts and slopes, it is easy to create problems that are intractable with Gaussian quadrature. Consequently, it is a useful method when a high degree of accuracy is desired but performs poorly in high dimensional spaces, for large datasets, or if speed is a concern. Additionally, if the outcome is skewed there can also be problems with the random effects.

# SAS Code

```
libname S857 'Z:\';


data amenorrhea;
set s857.amenorrhea;
t=time;
time2=time**2;
run;
```

```
proc sort data=amenorrhea;
by descending trt;
run;
title1 'Marginal Models';
title2 'Clinical Trial of Contracepting Women';

proc genmod data=amenorrhea descending;
class id trt (ref='0') t/param=ref;
model y = time time2 trt*time trt*time2 /dist=binomial link=logit type3 wald;
    repeated subject=id / withinsubject=t logor=fullclust;
store p1;
run;
ods graphics on;
ods html style=journal;
proc plm source=p1;
  score data = amenorrhea out=pred /ilink;
run;
proc sort data = pred;
  by trt time;
run;
proc sgplot data = pred;
  series x = time y = predicted /group=trt;
run;
ods graphics off;

title1 'Generalized Linear Mixed Effects Models';

proc glimmix  data=amenorrhea  method=quad(qpoints=5) empirical order=data ;
class id trt ;
model y = time time2 trt*time trt*time2 /dist=binomial link=logit s oddsratio;
random intercept time/subject=id type=un;
run;

title1 'Generalized Linear Mixed Effects Models';
title2 'Clinical Trial of AntiEpileptic-Drug';
data epilepsy;
set s857.epilepsy;
time=0;y=y0;output;
time=1;y=y1;output;
time=2;y=y3;output;
```

```
time=3;y=y3;output;
time=4;y=y4;output;
drop y0 y1 y2 y3 y4;
run;
data epilepsy;
set epilepsy;
if time=0 then t=log(8);
else t=log(2);
run;
proc sort data=epilepsy;
by descending trt;
run;
proc glimmix  data=epilepsy  method=quad(qpoints=50) empirical order=data ;
class id trt ;
model y = time trt trt*time /dist=poisson link=log s offset=t;
random intercept time/subject=id type=un;
run;
ods rtf close;
```

---

Revision #4
Created 7 March 2023 21:51:32 by Elkip
Updated 8 March 2023 15:11:28 by Elkip