GLMM workshop 19 March 2021 University of Manitoba

Instructors: David Schneider, Victor Valdez, with assistance of Taurai Matengu

Module 5	10 AM	Online	GLM review. GzLM
Break			
Module 6	10:45 AM	Online	GLM extensions: PCA, CCA (Multivariate Analysis)
Break			
Module 7	1 PM	Online	Executing the analysis in R
Break			
Module 8	2 PM	Online	GLM extensions: MANOVA

GzLM- The second solution to heterogeneous errors - 2 examples (Poisson and Binomial)

The GLM applies to a ratio scale response variable Y with normal error ε_{Normal}

Count data usually violate the assumption of homogeneous residuals.

Ratio scale counts (counts in defined units, ranging from zero upward)

Use Poisson or Negative Binomial error model $\varepsilon_{Poisson}$ or $\varepsilon_{NegBinomial}$

So we use the Generalized Linear Model, which allows us to use a better error model.

The GzLM (which includes the GLM as a special case) has three components

1. The structural model consisting of linear predictors.

   For the GLM, the linear predictor is the sum of fixed factors and covariates.

   The linear predictor ANCOVA example (Brussard) was $\eta = \beta_o + \beta_{SP}SP + \beta_{Hsl}Hsl + \beta_{SP \cdot Hsl}SP \cdot Hsl$
2. A linkfunction, that links the linear predictor to the response variable.
3. The error

Model equation form: $Y = \beta_o + \beta_x X + \varepsilon_{Normal}$

Probability distribution form: $Y \sim Normal(\beta_o + \beta_x X, \sigma^2)$

This is read as : Y is normally distributed, given the parameters β_o , β_x , and σ^2 (the fixed variance)

The distributional assumption, given the parameters, can only be checked after estimating the parameters

Count data usually violate this assumption.-- > heterogeneous residuals

So we use a better error model (Generalized Linear Model)

Ratio scale counts (counts in defined units, ranging from zero upward)

Use Poisson or Negative Binomial error model $\varepsilon_{Poisson}$ or $\varepsilon_{NegBinomial}$

$$Count = e^{\eta} + \varepsilon_{Poisson}$$

$$\eta = \beta_o + \beta_{V1}V1 + \ ...$$

Corps	Deaths
Guard	16
First	16
Second	12
Third	12

8. Death by horsekick. The classic example of Poisson data is the number of deaths by horse kick for each of 16 corps in the Prussian army, from 1875 to 1894. Bortkiewicz (1898 *The Law of Small Numbers*) showed that the horsekick data fit a Poisson distribution.

Symbol for response variable ___________ and for explanatory variable __________

Write the model $Odds = e^{\eta} + \varepsilon_{Poisson}$ (Fit to 1:1:1:1 assumes Poisson error)

$\eta =$ ___________________________

Likelihood Ratio Test: ΔG = 1.147 df = 3 p = 0.7658 Therefore, cannot reject 1:1:1:1 fit

http://www.mun.ca/biology/schneider/b4605/LNotes/Pt5/Ch17_2.pdf

GzLM with fixed explanatory variables. - 2nd example

The GLM assumes a normal error with fixed (constant variance) $= \varepsilon_{Normal}$

Count data usually violate this assumption.-- > heterogeneous residuals

So we use a better error model (Generalized Linear Model)

Nominal scale counts (units scored Y or N) Use binomial error model $\varepsilon_{binomial}$

Yes/No = *Odds* $Odds = e^{\eta} + \varepsilon_{binomial}$

$$\eta = \beta_o + \beta_{V1}V1 + \ldots$$

The response variable, Odds, are calculated as p/(1-p), where p is the ratio of success to number of trials.

9. Example – Cancer in cigarette smokers. Data from Cornfield (1951) who established the mathematical basis for using case-control samples to estimate risk in a population.

	Lung Tumors		
	Present	Absent	Total
heavy smokers	27	99	126
light smokers	8	72	80

Odds of tumor for

Heavy smokers __________

Light smokers ____________

Odds ratio, heavy relative to light __________

Symbol for response variable ___________ and for explanatory variable __________

Write the model $Odds = e^{\eta} + \varepsilon_{binomial}$

(contingency test not correct. it assumes Poisson error instead of binomial)

$\eta =$ __________________________

The 95% confidence limits are 1.05 to 5.1.

The null hypothesis is OR = 1. Odds the same for light and heavy smokers)

Do the confidence limits exclude the null? ________________________

http://www.mun.ca/biology/schneider/b4605/LNotes/Pt5/Ch18_3.pdf