GLMM workshop 18 March 2021 University of Manitoba

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

First session	10 AM	Online	Writing the model
Break			
Second session	11 AM	Online	F-ratios from Expected Mean Squares
Break			
Third session	1 PM	Online	Executing the analysis in R

Goal of the first session – Writing Statistical Models

GLM	The General Linear Model	Fixed Effects + Normal Error
GzLM	The Generalized Linear model	Fixed Effects + Non-normal Errors
GLMM	The General Linear Mixed Model	Fixed + Random + Normal
GzLMM	The Generalized Linear Mixed Model	Fixed + Random + Non-normal

Goal of the second session - Writing out the expected mean squares
Forming unambiguous likelihood ratio tests (F, t, χ^2)

Goal of the third session - Executing a GLMM in a statistical package
Interpreting the output

First session **10 AM** **Writing the model**

Preliminaries

Definitions Nominal, Ordinal, Interval, and Ratio scale variables.

Definitions: GLM General Linear Model

GzLM Generalized Linear Model

GLMM General Linear Mixed Model

GzLMM Generalized Linear Mixed Model

Series of examples to work through.

Distinguish response from explanatory variables

Assign symbols to all variables

Notational conventions Nominal scale variable ALL UPPER CASE

Ratio scale variables Begin with upper case.

β for fixed effect coefficients (slopes and contrasts)

μ for random effect parameters

Write the model, calculate the df, complete the first 2 columns of the ANOVA table

GLM with a single fixed explanatory variable 3 examples.

Write the Fixed Effect GLM, calculate df, fill in the blank columns of the ANOVA table.

1. Pea section growth data, from Box 9.4 in Sokal and Rohlf (1995).
Does length depend on treatment (control versus 4 different sugars with auxin present) ?
10 measurements of length of pea section in each treatment group

Length	*Len*	Response variable, ratio scale
Treatment	*TRT*	Categorical explanatory variable

Sketch graph of response vs explanatory

Write the model $Len = \beta_o + \beta_{Trt}\, TRT + \varepsilon_{Normal}$

Calculate df $(10*5) = 1 + (5-1) + 45$

df total = ntot -1 TRT df = number of categories – 1

Source	df
TRT	
error	
total	49

Fill out first 2 columns of ANOVA table from model
http://www.mun.ca/biology/schneider/b4605/LNotes/Pt3/Ch10_3.pdf
http://www.mun.ca/biology/schneider/b4605/GLMMworkshop/Data/PeaSections.csv

2. Example 9.3.1 from Snedecor and Cochran (1989). Quantity of interest is the phosphorus content of corn (*Pcorn* in ppm), in relation to the phosphorus levels in samples of soils with experimentally fixed levels of phosphorus (*Psoil* in ppm). Does the phosphorus content of corn increase when organic soil phosphorus is increased ? *Pcorn* and *Psoil* are both ratio scale variables. 9 measurements of *Pcorn*, matched with 9 of *Psoil*

Model ______________________________

df ______________________________

Sketch graph of response vs explanatory

http://www.mun.ca/biology/schneider/b4605/LNotes/Pt3/Ch9_1.pdf

http://www.mun.ca/biology/schneider/b4605/GLMMworkshop/Data/PCorn.csv

Source	df
error	7
total	

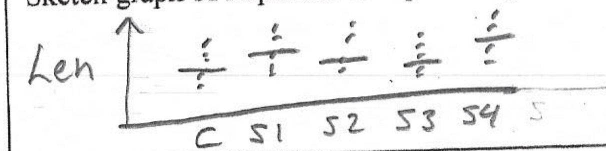
GLM with a single fixed explanatory variable 3 examples.

Write the Fixed Effect GLM, calculate df, fill in the blank columns of the ANOVA table.

1. Pea section growth data, from Box 9.4 in Sokal and Rohlf (1995).
Does length depend on treatment (control versus 4 different sugars with auxin present) ?
10 measurements of length of pea section in each treatment group

Length	*Len*	Response variable, ratio scale
Treatment	*TRT*	Categorical explanatory variable

Sketch graph of response vs explanatory



Write the model $Len = \beta_o + \beta_{Trt}\,TRT + \varepsilon_{Normal}$

Calculate df $(10*5) = 1 + (5-1) + 45$

df total = ntot -1 TRT df = number of categories – 1

Source	df
TRT	4
error	45
total	49

Fill out first 2 columns of ANOVA table from model

http://www.mun.ca/biology/schneider/b4605/LNotes/Pt3/Ch10_3.pdf

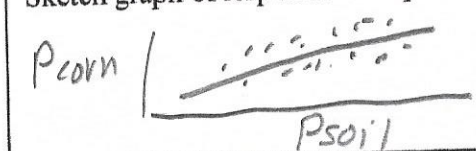
http://www.mun.ca/biology/schneider/b4605/GLMMworkshop/Data/PeaSections.csv

2. Example 9.3.1 from Snedecor and Cochran (1989). Quantity of interest is the phosphorus content of corn (*Pcorn* in ppm), in relation to the phosphorus levels in samples of soils with experimentally fixed levels of phosphorus (*Psoil* in ppm). Does the phosphorus content of corn increase when organic soil phosphorus is increased ? *Pcorn* and *Psoil* are both ratio scale variables. 9 measurements of *Pcorn*, matched with 9 of *Psoil*

Model $Pcorn = B_0 + B_{ps}\,Psoil + \varepsilon$

df $8 = 1 + 7$

Sketch graph of response vs explanatory



Source	df
Psoil	1
error	7
total	8

http://www.mun.ca/biology/schneider/b4605/LNotes/Pt3/Ch9_1.pdf

http://www.mun.ca/biology/schneider/b4605/GLMMworkshop/Data/PCorn.csv

GLM with a single fixed explanatory variable 3rd example.

3. Does inversion heterozygosity (HZYG) change with elevation above sea level (Hsl) in *Drosophila pseudoobscura*). Data are from Dobzhansky (1948) as reported in Brussard (1984).
One measurement of HZYG at each of 7 different elevations.

Source	df
	5

Response variable with symbol ____________________

Explanatory variable with symbol ____________________

Model ______________________________

df ______________________________

GLM with a single fixed explanatory variable Review

Definition of fixed effects:

1. *TRT* is a fixed effect because we are interested in the contrast among the 5 means.
   β_{TRT} is a set of unknown fixed effect contrasts.
2. *Psoil* is a fixed effect because we are interested in rate of increase in *Pcorn* with increase in *Psoil.*
   β_{Psoil} is the unknown rate.
3. *Hsl* is a fixed effect because we are interested in the whether *Hzyg* changes with elevation (altitude above sea level)
   β_{Hsl} is the fixed effect rate. $\hat{\beta}_{Hsl}$ is an estimate of β_{Hsl}

GLM with a single fixed explanatory variable 3rd example.

3. Does inversion heterozygosity (HZYG) change with elevation above sea level (Hsl) in *Drosophila pseudoobscura*). Data are from Dobzhansky (1948) as reported in Brussard (1984).
One measurement of HZYG at each of 7 different elevations.

Source	df
Hsl	1
error	5
total	6

Response variable with symbol Hzyg

Explanatory variable with symbol Hsl

Model $Hzyg = \beta_0 + \beta_H Hsl + \varepsilon$

df $6 = 1 + 5$

GLM with a single fixed explanatory variable Review

Definition of fixed effects:

1. *TRT* is a fixed effect because we are interested in the contrast among the 5 means.
   β_{TRT} is a set of unknown fixed effect contrasts.
2. *Psoil* is a fixed effect because we are interested in rate of increase in *Pcorn* with increase in *Psoil.*
   β_{Psoil} is the unknown rate.
3. *Hsl* is a fixed effect because we are interested in the whether *Hzyg* changes with elevation (altitude above sea level)
   β_{Hsl} is the fixed effect rate. $\hat{\beta}_{Hsl}$ is an estimate of β_{Hsl}