

GzLM Intro. docx

Revision

GzLM
GzLMM

Generalized Linear Mixed Models
Generalized Linear Models

Special session

2:00 PM Tuesday 24 Oct 2017

Learning Goals

- 1 Analysis of Deviance
Likelihood ratio tests from any probability model
- 2 Probability models for count data
Binomial
Poisson --- > Overdispersed Poisson
--- > Negative Binomial
- 3 Link Functions "Canonical Link" (the default)
Science founded links

Needs examples

BF H² ✓
Table 2-4-2
binomial

Poisson—

p224
~~p333~~

p333?

~~Sat 7:30-9:30 Mariz~~

~~Sun night Olufemi solutions to DAA4 Kurt~~

NKS

updated Lab 08 to web

~~Saturday April~~

~~Blundell 2 ^{test} copies Monday~~

NO DEV The analysis of of deviance is used for any error structure normal as well as others

$$H = \beta_0 + \beta_E \text{Elev} + \epsilon$$

$$2R \frac{L(\beta_0 | \text{Data})}{L(\beta_0, \beta_E | \text{Data})}$$

We have already seen analysis of deviance for normal error model
<http://www.mun.ca/biology/schneider/b4605/LNotes/Pt2/Ch5.pdf>

Regression (Fly heterozygosity) Ch 5 page 9
 average(H) = 0.351429

Exp(H) 0.5796 + -0.12706 Elevation

	Elev H(%)	Elev(km)
850	0.59	0.26
3000	0.37	0.91
4600	0.41	1.4
6200	0.4	1.89
8000	0.31	2.44
8600	0.18	2.62
10000	0.2	3.05

Elevation H Hnull average(H) residuals

0.26 0.59
 0.91 0.37
 1.4 0.41
 1.89 0.4
 2.44 0.31
 2.62 0.18
 3.05 0.2

res^2

Alternative Exp(H) residuals

res^2

Sum

SS total =

SS res =

Improvement in fit

= SS regre

The improvement in fit is called the Deviance
 The improvement was due to one additional parameter (1 df)

The likelihood ratio is $(SS_{res}/SS_{total})^{(-n/2)}$

Calculate the likelihood ratio

LR =

The test statistics is $G = 2 \ln(LR)$

Calculate the G statistic

We use the cumulative chisquare distribution to obtain the p-value

1 - CHISQ.DIST(3.84,1,2) 0.05

Calculate the p-value for the G statistic

use
~~can~~
 Panagen
 t-test
 Instead of
 regression

ANODEV

The analysis of of deviance is used for any error structure normal as well as others

The analysis of deviance for a binomial distribjton is shown in chapter 16.2
http://www.mun.ca/biology/schneider/b4605/LNotes/Pt5/Ch16_2.pdf

See pages 4 & 5 for calculation of the G statistic

For the Mendel pea data $G =$

We use the cumulative chisquare distribution to obtain the p-value
 $1 - \text{CHISQ.DIST}(3.84, 1, 2) \quad 0.05$

Calculate the p-value for the G statistic

$$\text{Odds} = e^{\eta} + \epsilon$$

$$\eta = \beta_0 + \beta_m \cdot \text{Mendel}$$

$$L(\mu, \sigma^2 | \text{Data})$$
$$L(\mu, \mu_m | \text{Data})$$

2 Probability models for count data

Binomial

Poissor --- > Overdispersed Poisson
--- > Negative Binomial

For examples see Ch 16.1

http://www.mun.ca/biology/schneider/b4605/LNotes/Pt5/Ch16_1.pdf

Ch 16.1

3 Link Functions "Canonical Link" (the default) Science founded links

We start with the identity link, which is the canonical link for normal errors.

Here is the model for the analysis of the limpet data

$$VO_2 = \beta_0 + \beta_{sal} SAL + \beta_{sp} SP + \beta_{sal*Sp} SAL*Sp + \epsilon$$

We rewrite this as

$$VO_2 = v + \epsilon$$

where

$$v = \beta_0 + \beta_{sal} SAL + \beta_{sp} SP + \beta_{sal*Sp} SAL*Sp$$

For a logarithmic line we have

$$VO_2 = e^v + \epsilon$$

There is a canonical for each distribution.

Distribution	Link
normal	identity
Poisson	Log
Binomial	logit (log Odds)
Negative Binomial	log
Gamma	inverse

In some cases it makes sense to use a different link.

For example, we would prefer using the identity (or perhaps log) link for the Gamma distribution

The identity (additive effects) and log link (multiplicative effects) are more interpretable.

To gain experience with the GzLM try working the e Ch 17.2

Ch 18.3

Poisson Error

Binomial Error

In each case the data set is small. And the R-code is provided.