
Statistical Inference & Model Checking for Poisson Regression Edpsy/Psych/Soc 589

Carolyn J. Anderson

Department of Educational Psychology



I L L I N O I S

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Outline

Overview

● Outline

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

SAS for these Models

Fitting GLMS

Statistic inference & the Likelihood function

“Deviance”

Summary

- Inference for model parameters (i.e., confidence intervals and hypothesis tests).
 - ◆ Wald Statistics
 - ◆ Likelihood ratio tests.
 - ◆ (Score tests).
- Assessing model fit.
 - ◆ Global fit statistics (X^2 , G^2).
 - ◆ Residuals.
 - ◆ Confidence intervals for fitted values
 - ◆ Overdispersion.
- Extended example: Bullies
- Zero Inflated Models (ZIPs)
- Fitting GLMs, which provides further tools and insights into inference and model assessment.

Much of the basic logic and concepts for Poisson regression are the same as those for logistic regression, but we'll consider logistic regression in more detail when we cover Chapter 5.

Inference for model parameters

Suppose that we've fit the Poisson regression model

$$\log(\hat{\mu}_i) = \hat{\alpha} + \hat{\beta}x_i$$

For the AIDS data we obtained

$$\log(\hat{\mu}_i) = -1.9442 + 2.1748x_i^*$$

where x_i^* is log(month period).

Typically, the kinds of hypothesis tests that we'd be interested in performing are whether our explanatory variable(s) had an effect on the response variable; that is,

$H_O : \beta = 0$ versus one of the following:

- $H_A : \beta \neq 0$ (2-tailed test).
- $H_A : \beta > 0$ (right tailed test).
- $H_A : \beta < 0$ (left tailed test).

Overview

Inference for model parameters

● Inference for model parameters

● Wald Statistics
● Wald Statistics for Hypothesis Testing

● Wald Statistic
● Likelihood Ratio Test
● Likelihood Ratio Test Statistic
● Example using Likelihood Ratio Test Statistic

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

SAS for these Models

Fitting GLMS

Statistic inference & the Likelihood function
Generalized Linear Models

"Deviance"

Wald Statistics

Fact: **Maximum likelihood estimates** are approximately normally distributed for large samples (i.e., MLEs are asymptotically normally distributed).

Which means that for “large” samples,

$$\hat{\beta} \approx \mathcal{N}(\beta, \sigma_{\hat{\beta}}^2)$$

We can use this to

- Construct confidence intervals for parameters.
- Test hypotheses.

For a $(1 - \alpha)100\%$ confidence interval estimate of β :

$$\hat{\beta} \pm z_{\alpha/2} ASE$$

where ASE is the asymptotic standard error of $\hat{\beta}$.

For the AIDs example, a 95% confidence interval for β is

$$2.1748 \pm 1.96(.2151) \longrightarrow (1.753, 2.596)$$

Overview

Inference for model parameters

● Inference for model parameters

● Wald Statistics

● Wald Statistics for Hypothesis Testing

● Wald Statistic

● Likelihood Ratio Test

● Likelihood Ratio Test Statistic

● Example using Likelihood Ratio Test Statistic

Ratio Test Statistic

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

SAS for these Models

Fitting GLMS

Statistic inference & the

Likelihood function

Generalized Linear Models

“Deviance”

Wald Statistics for Hypothesis Testing

For the hypothesis test of $H_0 : \beta = 0$ versus $H_A : \beta \neq 0$ (or $H_A : \beta > 0$ or $H_A : \beta < 0$),

$$z = \frac{\hat{\beta} - \beta_0}{ASE} = \frac{\hat{\beta}}{ASE}$$

where β_0 is the hypothesized value of β under the null hypothesis (i.e., $\beta_0 = 0$).

If the null hypothesis is *true*, then the statistic z is approximately standard normal (for large samples)

$$z = \frac{\hat{\beta}}{ASE} \approx \mathcal{N}(0, 1)$$

Overview

Inference for model parameters

● Inference for model parameters

● Wald Statistics
● Wald Statistics for Hypothesis Testing

● Wald Statistic
● Likelihood Ratio Test
● Likelihood Ratio Test Statistic
● Example using Likelihood Ratio Test Statistic

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

SAS for these Models

Fitting GLMS

Statistic inference & the Likelihood function
Generalized Linear Models

“Deviance”

Wald Statistic

An estimated parameter divided by its ASE and squared is a “Wald Statistic”.

$$z^2 = \left(\frac{\hat{\beta}}{ASE} \right)^2 \approx \chi_1^2$$

z^2 has (asymptotically) a chi-squared distribution with $df = 1$.

- Wald statistics are usually provided on computer output, along with p -values (see SAS output).
- Wald statistics can be used to test 2-sided alternatives, while z can be used to test 1-sided as well as 2-sided alternatives.

AIDS Example:

$$H_o : \beta = 0 \quad \text{versus} \quad H_o : \beta \neq 0$$

“Chi-square” = $(2.1748/.2151)^2 = (10.11)^2 = 102.23$, $df = 1$,
 $p < .01$.

Overview

Inference for model parameters

● Inference for model parameters

● Wald Statistics

● Wald Statistics for Hypothesis

Testing

● Wald Statistic

● Likelihood Ratio Test

● Likelihood Ratio Test Statistic

● Example using Likelihood

Ratio Test Statistic

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

SAS for these Models

Fitting GLMS

Statistic inference & the

Likelihood function

Generalized Linear Models

“Deviance”

Likelihood Ratio Test

For this test, we look at the ratio of

- The maximum value of the likelihood function over all possible parameter values assuming that the null hypothesis is true.
- The maximum value of the likelihood function over a larger set of possible parameter values (possible parameters for a “full” or more general model).

Suppose that we wish to test $H_O : \beta = 0$.

- Let l_1 = the maximum value of the likelihood function for the full model:

$$\log(\mu_i) = \alpha + \beta x_i$$

- Let l_0 = the maximum value of the likelihood function for the model when H_O is true:

$$\log(\mu_i) = \alpha$$

Model under H_O places more restrictions on the set of possible parameters, so

$$l_0 \leq l_1$$

Overview

Inference for model parameters

● Inference for model parameters

● Wald Statistics

● Wald Statistics for Hypothesis Testing

● Wald Statistic

● Likelihood Ratio Test

● Likelihood Ratio Test Statistic

● Example using Likelihood Ratio Test Statistic

Ratio Test Statistic

Ratio Test Statistic

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

SAS for these Models

Fitting GLMS

Statistic inference & the

Likelihood function

Generalized Linear Models

“Deviance”

Likelihood Ratio Test Statistic

The **Likelihood ratio test** statistic equals

$$\begin{aligned} -2 \log(l_0/l_1) &= -2\{\log(l_0) - \log(l_1)\} \\ &= -2(L_0 - L_1) \end{aligned}$$

where $L_0 = \log(l_0)$ (and $L_1 = \log(l_1)$) are the “maximized log-likelihood functions”.

- When $l_0 = l_1$ so $L_0 = L_1$, the test statistic equals 0.
- If H_0 is true, then the likelihood ratio test statistic is approximately chi-squared distributed with degrees of freedom equal to $df = 1$ (for $H_0 : \beta = 0$).
- When the null is false, $l_0 < l_1$ (so $L_0 < L_1$), the test statistic is > 0 . The larger the statistic, the greater the evidence against the null hypothesis being true.

Overview

Inference for model parameters

● Inference for model parameters

● **Wald Statistics**

● Wald Statistics for Hypothesis Testing

● Wald Statistic

● Likelihood Ratio Test

● Likelihood Ratio Test Statistic

● Example using Likelihood Ratio Test Statistic

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

SAS for these Models

Fitting GLMS

Statistic inference & the Likelihood function
Generalized Linear Models

“Deviance”

Example using Likelihood Ratio Test Statistic

AIDs example:

■ $H_O : \beta = 0$ versus $H_A : \beta \neq 0$.

■

$$-2(L_O - L_1) = -2(383.2532 - 478.3435) = 190.1806$$

■ which with $df = 1$ has a very small p -value.

Where to get these values come from?

■ Under “Criteria For Assessing Goodness of Fit”, SAS/GENMOD provides the maximized log-likelihood value (see “Log Likelihood”)

Criterion	DF	Value	Value/DF
Deviance	12	17.0917	1.4243
Pearson Chi-Square	12	15.9884	1.3324
Log Likelihood		478.3435	

■ Or using the “type3” option to the model statement:
`model count = lmonth / dist=count link=log type3;`

Overview

Inference for model parameters

● Inference for model parameters

● Wald Statistics

● Wald Statistics for Hypothesis Testing

● Wald Statistic

● Likelihood Ratio Test

● Likelihood Ratio Test Statistic

● Example using Likelihood Ratio Test Statistic

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

SAS for these Models

Fitting GLMS

Statistic inference & the

Likelihood function

Generalized Linear Models

“Deviance”

Assessing Model Fit

Overview

Inference for model parameters

Assessing Model Fit

● Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

SAS for these Models

Fitting GLMS

Statistic inference & the Likelihood function

“Deviance”

Summary

In assessing whether a model fits the data well (i.e., the model provides a “good” or accurate description of the data), we should examine/consider

- Global Fit Statistics.
- Residuals.
- Confidence Intervals
- Overdispersion.

Global Fit Statistics

The null hypothesis for these statistics is

H_0 : The model fits the data.

The lack of fit is not statistically large.

Let

- $i = 1, \dots, N$ index the levels of the explanatory variable.
- y_i be the observed count for the i th level of the explanatory variable.

The Pearson “goodness-of-fit” statistic is

$$X^2 = \sum_{i=1}^N \frac{(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i}$$

The Likelihood ratio “goodness-of-fit” statistic is

$$G^2 = 2 \sum_{i=1}^n y_i \log(y_i / \hat{\mu}_i)$$

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

- Global Fit Statistics
- Distribution of Global Fit

Statistics

- AIDs Example & Global Fit

Statistics

- Problem with the AIDs

Example

- Lung Cancer Example
- Ways to deal with Problem
- Grouping Observations
- Collapsing Horseshoe Crab

Data

- Method I: Fit Model then

Collapse

- Method II: Collapse then Fit

Model

- Method II: continued

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

Generalized Linear Models

SAS for these Models

Distribution of Global Fit Statistics

If

1. The null hypothesis is true,
2. The fitted values are large $\rightarrow \hat{\mu}_i \geq 5$,
3. The number of levels of the explanatory variable(s) is fixed,

Then X^2 and G^2 are approximately chi-squared distributed with degrees of freedom (or “*residual df*”) equal to

$$df = \text{number of counts} - \text{number of model parameters.}$$

If (1) is false (but (2) and (3) hold), then we expect X^2 and/or G^2 to be “large” (i.e., far out in the right tail of the proper chi-squared distribution) and this is evidence against the null hypothesis.

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

- Global Fit Statistics
- Distribution of Global Fit Statistics
- AIDs Example & Global Fit Statistics
- Problem with the AIDs Example
- Lung Cancer Example
- Ways to deal with Problem
- Grouping Observations
- Collapsing Horseshoe Crab Data
- Method I: Fit Model then Collapse
- Method II: Collapse then Fit Model
- Method II: continued

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

Generalized Linear Models
SAS for these Models

AIDs Example & Global Fit Statistics

For the AIDs example for the model $\log(\mu_i) = \alpha + \beta x_i$,

Statistic	df	Value
X^2	12	15.99
G^2	12	17.09

Note: $N = 14$, we have just 1 observation (count) for each quarter (3 month period), and the number of parameters = 2, so

$$df = 14 - 2 = 12$$

Wonder why SAS doesn't automatically provide p -values for these statistics?

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

- Global Fit Statistics
- Distribution of Global Fit Statistics
- AIDs Example & Global Fit Statistics
- Problem with the AIDs Example

- Lung Cancer Example
- Ways to deal with Problem
- Grouping Observations
- Collapsing Horseshoe Crab Data
- Method I: Fit Model then Collapse
- Method II: Collapse then Fit Model
- Method II: continued

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

Generalized Linear Models
SAS for these Models

Problem with the AIDs Example

Problems with global fit statistics and Poisson regression:

1. Often there are (many) small “cell” counts.
 - In the AIDs data, 6 of the 14 counts are less than 5.
2. The chi-squared approximation for X^2 and G^2 is based on asymptotic theory, which means that as sample size increases for a fixed N (number of levels of the explanatory variable), the approximation becomes better.
 - If we add observations in the AIDs example, we also increase N , the number of levels of the explanatory variable.

Because of these two problems,

- X^2 and G^2 are not good measures of a model’s “badness-” or “lack-of-fit”
- The approximation of the distributions of X^2 and G^2 by the chi-squared distribution is bad.

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

- Global Fit Statistics
- Distribution of Global Fit Statistics
- AIDs Example & Global Fit Statistics
- Problem with the AIDs Example
- Lung Cancer Example
- Ways to deal with Problem
- Grouping Observations
- Collapsing Horseshoe Crab Data
- Method I: Fit Model then Collapse
- Method II: Collapse then Fit Model
- Method II: continued

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

Generalized Linear Models
SAS for these Models

Lung Cancer Example

- Table doesn't increase as add more observations, because the Explanatory Variables are :

- ◆ City in Denmark (Fredericia, Horsens, Kolding, Vejle).
- ◆ Age class (40–54, 55–59, 60–64, 65–69, 70–74, >75).

- Of the 24 cells, smallest ones are 2 and 4; that is, 92% are “large”.

Criterion	DF	Value	Value/DF	<i>p</i>
Deviance	20	26.2815	1.3141	.16
Pearson Chi-Square	20	24.2465	1.2123	.23

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

- Global Fit Statistics
- Distribution of Global Fit Statistics
- AIDs Example & Global Fit Statistics
- Problem with the AIDs Example
- Lung Cancer Example
- Ways to deal with Problem
- Grouping Observations
- Collapsing Horseshoe Crab Data
- Method I: Fit Model then Collapse
- Method II: Collapse then Fit Model
- Method II: continued

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

Generalized Linear Models
SAS for these Models

Ways to deal with Problem

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

- Global Fit Statistics
- Distribution of Global Fit Statistics
- AIDs Example & Global Fit Statistics
- Problem with the AIDs Example
- Lung Cancer Example
- Ways to deal with Problem
- Grouping Observations
- Collapsing Horseshoe Crab Data
- Method I: Fit Model then Collapse
- Method II: Collapse then Fit Model
- Method II: continued

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

Generalized Linear Models
SAS for these Models

1. Use small sample methods.
2. “Discretize” or collapse levels of the explanatory variable.

There are two variants of this latter strategy:

- (a) Sum observed counts and fitted values within the same category of the collapsed explanatory variable and recompute the test statistics.
- (b) Sum observed counts within the same category of the explanatory variable and re-fit the model to the data and using scores for each category of the explanatory variable (e.g., means, or other).

3. Bayesian methods

Grouping Observations

By grouping observations on the basis of the explanatory variable, X^2 and G^2 should be better approximated by chi-squared distributions.

By collapsing data, both problems are solved:

1. The observations/counts per level of the explanatory variable increases.
2. As you increase the number of observations per category/level of the explanatory variable, the number of levels of the explanatory variable is constant (i.e., N is fixed).

Will this strategy work for the AIDs example?

No, because of the nature of the study. If you collect more data, you necessarily increase N .

Collapsing strategies work with the horseshoe crab example (Agresti, 1996).

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

- Global Fit Statistics
- Distribution of Global Fit

Statistics

- AIDs Example & Global Fit

Statistics

- Problem with the AIDs

Example

- Lung Cancer Example
- Ways to deal with Problem
- Grouping Observations
- Collapsing Horseshoe Crab

Data

- Method I: Fit Model then

Collapse

- Method II: Collapse then Fit

Model

- Method II: continued

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

Generalized Linear Models

SAS for these Models

Collapsing Horseshoe Crab Data

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

- Global Fit Statistics
- Distribution of Global Fit Statistics
- AIDs Example & Global Fit Statistics
- Problem with the AIDs Example
- Lung Cancer Example
- Ways to deal with Problem
- Grouping Observations
- Collapsing Horseshoe Crab Data
- Method I: Fit Model then Collapse
- Method II: Collapse then Fit Model
- Method II: continued

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

Generalized Linear Models
SAS for these Models

Before collapsing, there are 66 different widths of female crabs (many with very small counts).

8 categories of widths were used and both of the two of the collapsing strategies were used.

Deciding on categories for explanatory variable:

- Often just take equal spacing (as done in horseshoe example, i.e., 1 cm) — easy, works well when observations are equally spread out over the range of the explanatory variable.
- Each of categories should have $\hat{\mu}_i \geq 5$.

Method I: Fit Model then Collapse

Both the observations and fitted values from the model (fit to the uncollapsed data) within the same width categories were summed.

Using the summed counts and fitted values, the model test statistics were recomputed:

Statistic	<i>df</i>	Value	<i>p</i> -value
X^2	6	6.5	.37
G^2	6	6.9	.33

Note:

$$df = (\# \text{ of categories}) - (\# \text{ of parameters}) = 8 - 2 = 6.$$

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

- Global Fit Statistics
- Distribution of Global Fit Statistics
- AIDs Example & Global Fit Statistics
- Problem with the AIDs Example
- Lung Cancer Example
- Ways to deal with Problem
- Grouping Observations
- Collapsing Horseshoe Crab Data
- Method I: Fit Model then Collapse
- Method II: Collapse then Fit Model
- Method II: continued

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

Generalized Linear Models
SAS for these Models

Method II: Collapse then Fit Model

Sum the observed counts within each of the width categories and re-fit the model to collapsed data using some scores for the width categories (mean width within category).

With this method, each observation within a category is treated as if it has the same width. (Method II is “easier” than Method I, but Method II is also “cruder” than Method I).

The Poisson regression with log link and the offset $t = \log(\text{number of cases per width category})$:

$$\log(\hat{\mu}_i/t) = -3.535 + .173x_i$$

and the estimated model from un-collapsed data was

$$\log(\hat{\mu}_i) = -3.305 + .164x_i$$

The fit statistics for the model fit to the collapsed data:

Statistic	<i>df</i>	Value	<i>p</i> -value
X^2	6	6.5	.37
G^2	6	6.9	.33

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

- Global Fit Statistics
- Distribution of Global Fit Statistics
- AIDs Example & Global Fit Statistics
- Problem with the AIDs Example
- Lung Cancer Example
- Ways to deal with Problem
- Grouping Observations
- Collapsing Horseshoe Crab Data
- Method I: Fit Model then Collapse
- Method II: Collapse then Fit Model
- Method II: continued

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

Generalized Linear Models

SAS for these Models

Method II: continued

fitted counts look pretty close to the observed counts

	Mean	Number	Number	Fitted
Width	Width	Cases	Counts	Count
< 23.25	22.69	14	14	20.5
23.25–24.25	23.84	14	20	25.1
24.25–25.25	24.77	28	67	58.9
25–25–26.25	25.84	39	105	98.6
26.25–27.25	26.79	22	63	65.6
27.25–28.25	27.74	24	93	84.3
28.25–29.25	28.67	18	71	74.2
> 29.25	30.41	14	72	77.9

Conclusion: Both Methods I & II yield indicate that the Poisson regression model looks pretty good for these data.

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

- Global Fit Statistics
- Distribution of Global Fit Statistics
- AIDs Example & Global Fit Statistics
- Problem with the AIDs Example
- Lung Cancer Example
- Ways to deal with Problem
- Grouping Observations
- Collapsing Horseshoe Crab Data
- Method I: Fit Model then Collapse
- Method II: Collapse then Fit Model
- Method II: continued

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

Generalized Linear Models
SAS for these Models

Residuals

Pearson Residuals are standardized differences between observed counts and fitted values:

$$\text{Pearson residual} = \frac{(\text{observed} - \text{fitted})}{\sqrt{\hat{\text{Var}}(\text{observed})}}$$

which for the Poisson generalized linear models for counts is,

$$e_i = \frac{y_i - \hat{\mu}_i}{\sqrt{\hat{\mu}_i}}$$

Notes:

- The estimated standard deviation of a fitted value in Poisson regression is the square root of the fitted value.
- $\sum_i e_i^2 = X^2$
- Observations with larger Pearson residuals make larger contributions to X^2 .

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

- Residuals
- Residuals from AIDs example
- Residuals from AIDs example (plot)
- Adjusted Residuals
- AIDs Adjusted Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

SAS for these Models

Fitting GLMS

Statistic inference & the Likelihood function

Deviance

Summary
Generalized Linear Models

Residuals from AIDs example

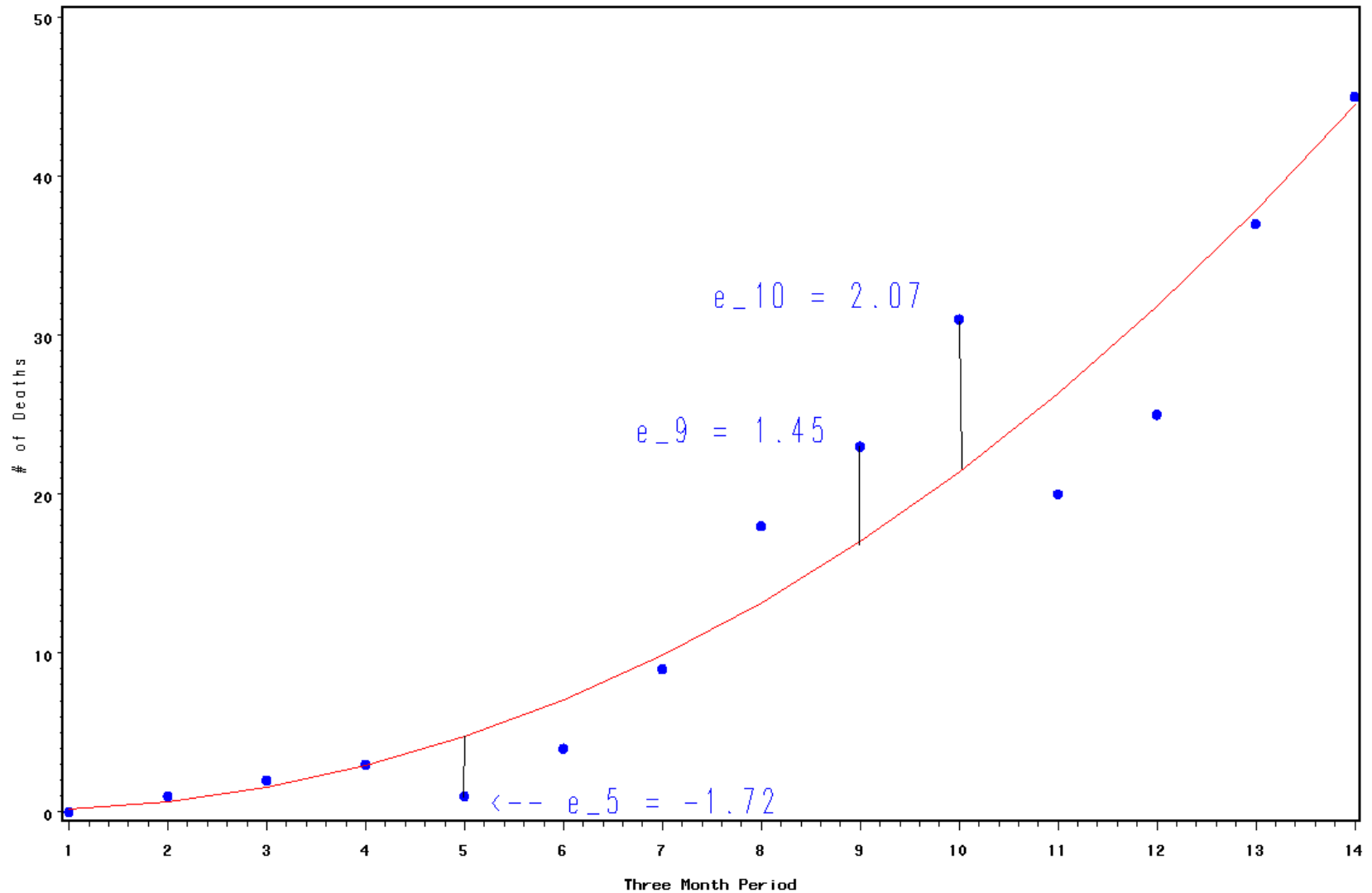
- Overview
- Inference for model parameters
- Assessing Model Fit
- Global Fit Statistics
- Residuals
 - Residuals
 - Residuals from AIDs example
 - Residuals from AIDs example (plot)
 - Adjusted Residuals
 - AIDs Adjusted Residuals
- Confidence intervals for fitted values
- Overdispersion
- Overview of Bully Example
- Zero Inflated Models
- SAS for these Models
- Fitting GLMS
- Statistic inference & the Likelihood function
- Deviance
- Summary
- Generalized Linear Models

Month	# Deaths	Fitted Value	Pearson residual	Contribution to X^2
x_i	y_i			
1	0	.14	-.38	.14
2	1	.65	.44	.19
3	2	1.56	.35	.12
4	3	2.92	.05	.00
5	1	4.74	-1.72	2.95
6	4	7.05	-1.15	1.32
7	9	9.86	-.27	.07
8	18	13.17	1.33	1.77
9	23	17.02	1.45	2.10
10	31	21.40	2.07	4.30
11	20	26.33	-1.23	1.52
12	25	31.82	-1.29	1.46
13	37	37.87	-1.14	.02
14	45	44.49	.08	.01

Residuals from AIDs example (plot)

The largest residual $e_{10} = 2.07$, with the next largest $e_5 = -1.72$, which taken together contribute $4.30 + 2.95 = 7.25$ to X^2 for the model, which is about 45% of $X^2 = 15.99$.

Final Model for AIDs Data



Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

- Residuals
- Residuals from AIDs example
- Residuals from AIDs example (plot)
- Adjusted Residuals
- AIDs Adjusted Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

SAS for these Models

Fitting GLMS

Statistic inference & the Likelihood function

Deviance

Summary
Generalized Linear Models

Adjusted Residuals

When a model fits data, the Pearson residuals should be approximately normally distributed with mean 0 but the variance is slightly less than 1.

$(y_i - \hat{\mu}_i)$ tends to be smaller than $(y_i - \mu_i)$, because sample data are used to obtain $\hat{\mu}_i$.

$$\begin{aligned}\text{Adjusted residual} &= \frac{\text{Pearson residual}}{\text{Pearson residual's standard error}} \\ &= \frac{e_i}{\sqrt{(1 - h_i)}} = \frac{(y_i - \hat{\mu}_i)}{\sqrt{\hat{\mu}_i(1 - h_i)}}\end{aligned}$$

where h_i is “leverage”, which is a measure of how much an observation potentially influences the fit of the model.

Adjusted residuals are

- Approximately $\mathcal{N}(0, 1)$ when the model holds.
- Good for finding “large” residuals.

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

- Residuals
- Residuals from AIDs example
- Residuals from AIDs example (plot)
- Adjusted Residuals
- AIDs Adjusted Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

SAS for these Models

Fitting GLMS

Statistic inference & the Likelihood function

Deviance

Summary
Generalized Linear Models

AIDs Adjusted Residuals

- Overview
- Inference for model parameters
- Assessing Model Fit
- Global Fit Statistics
- Residuals
 - Residuals
 - Residuals from AIDs example
 - Residuals from AIDs example (plot)
 - Adjusted Residuals
 - AIDs Adjusted Residuals
- Confidence intervals for fitted values
- Overdispersion
- Overview of Bully Example
- Zero Inflated Models
- SAS for these Models
- Fitting GLMS
- Statistic inference & the Likelihood function
- Deviance
- Summary
- Generalized Linear Models

Month	# Deaths	Fitted Value	Pearson residual	Adjusted Residual
x_i	y_i			
1	0	.14	-.38	-.39
2	1	.65	.44	.46
3	2	1.56	.35	.38
4	3	2.92	.05	.05
5	1	4.74	-1.72	-1.86
6	4	7.05	-1.15	-1.24
7	9	9.86	-.27	-.29
8	18	13.17	1.33	1.41
9	23	17.02	1.45	1.53
10	31	21.40	2.07	2.19
11	20	26.33	-1.23	-1.32
12	25	31.82	-1.29	-1.33
13	37	37.87	-1.14	-1.16
14	45	44.49	.08	.10

Confidence intervals for fitted values

Just as in normal linear regression, we can put confidence intervals around our Poisson regression fitted values.

■ $\hat{\beta} \sim \mathcal{N}(\beta, \Sigma_{\hat{\beta}})$.

■ The linear predictor is a linear combination of $\hat{\beta}$:

$$\mathbf{x}'_i \hat{\beta}$$

e.g., For one explanatory variable,

$$\mathbf{x}'_i \hat{\beta} = (1, x_i) \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = \hat{\alpha} + \hat{\beta}x_i$$

Therefore the linear predictor is normally distributed,

$$\mathbf{x}'_i \hat{\beta} \sim \mathcal{N}(\mathbf{x}'_i \beta, \sigma^2)$$

where $\sigma^2 = \mathbf{x}'_i \Sigma_{\hat{\beta}} \mathbf{x}_i$.

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

● Confidence intervals for fitted values

● Confidence intervals for Regression

● AIDs Confidence Bands for Regression

Overdispersion

Overview of Bully Example

Zero Inflated Models

SAS for these Models

Fitting GLMS

Statistic inference & the Likelihood function

“Deviance”

Summary

Generalized Linear Models

Confidence intervals for Regression

- Since the linear predictor is normal, then a $(1 - \alpha)100\%$ confidence interval for $\log(\mu_i)$ is

$$\log(\hat{\mu}_i) \pm z_{\alpha/2} \sqrt{\sigma^2}$$

- A $(1 - \alpha)100\%$ confidence interval for μ_i is

$$\exp \left[\log(\hat{\mu}_i) \pm z_{\alpha/2} \sqrt{\sigma^2} \right]$$

- On the SAS/GENMOD output under the observation statistics, UPPER and LOWER correspond to the upper and lower ends of confidence intervals for μ_i .

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

● Confidence intervals for fitted values

● Confidence intervals for Regression

● AIDs Confidence Bands for Regression

Overdispersion

Overview of Bully Example

Zero Inflated Models

SAS for these Models

Fitting GLMS

Statistic inference & the Likelihood function

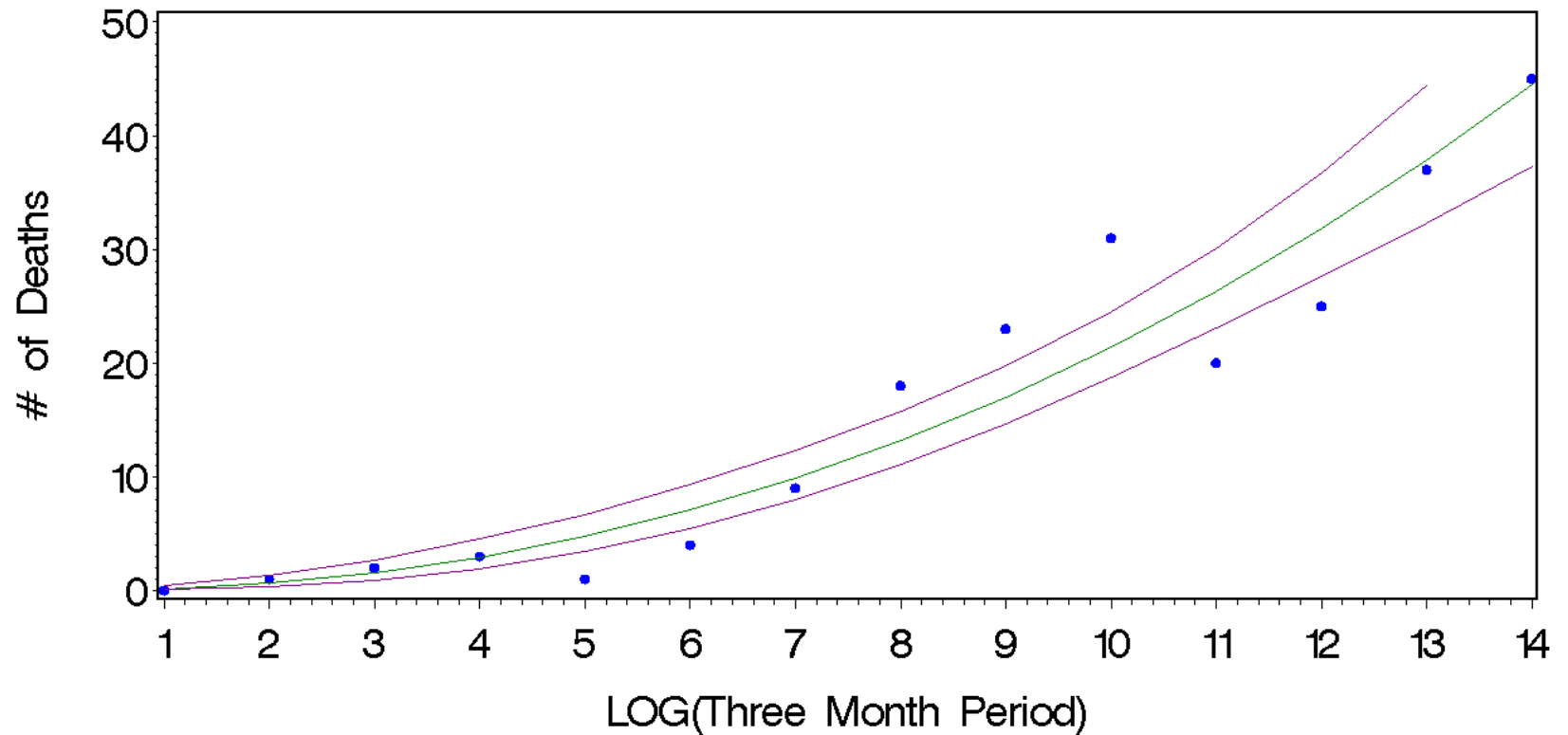
“Deviance”

Summary

Generalized Linear Models

AIDs Confidence Bands for Regression

95% Confidence Bands/Intervals for Mean



Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

● Confidence intervals for fitted values

● Confidence intervals for Regression

● AIDs Confidence Bands for Regression

Overdispersion

Overview of Bully Example

Zero Inflated Models

SAS for these Models

Fitting GLMS

Statistic inference & the Likelihood function

“Deviance”

Summary

Generalized Linear Models

Overdispersion

Observed count data often show greater variability than would be expected if the data were really from a Poisson (or binomial) distribution.

If data come from a Poisson distribution, then

$$\text{mean} = \text{variance}$$

But often we find that

$$\text{mean} < \text{variance}$$

This situation is referred to as **overdispersion**.

Most common causes of overdispersion: **Heterogeneity**.

In Poisson regression, we assume that the randomness of observations on individuals with the same value on the explanatory variable(s) can be described by the same Poisson distribution (i.e., the same mean).

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

- Overdispersion
- Example of Overdispersion
- Overdispersion: When to be Concerned
- Detecting Overdispersion
- Examples: Detecting Overdispersion
- Implication of Overdispersion
- Adjusting estimated standard errors
- Modeling the extra variability
- Gamma Distributions with $\mu = 1$
- Modeling the extra variability
- Negative Binomial Distribution
- vs Poisson Distribution
- Example: Modeling the extra variability

● Overdispersion and

SAS/GENMOD

● Overdispersion Example from Generalized Linear Models Lindsey

● Possible Model: Poisson

Example of Overdispersion

Consider the example of the number of violent incidences of individuals with mental illnesses who had been treated in ER of a psychiatric hospital.

If the Poisson model for counts is correct, then for each patient with the same age, concern score, and history of violent incidences, the expected count should equal

$$\hat{\mu}_i = t_i \exp\{\hat{\alpha} + \hat{\beta}_1 \text{age}_i + \hat{\beta}_2 \text{concern}_i + \hat{\beta}_3 \text{history}_i\}$$

where t_i is the offset (the number of days the individual was in the community during the 6 month period of the study).

And the estimated probability of observing counts equal to $y = 0, 1, 2, \dots$, should be given by

$$\hat{P}(Y_i = y) = \frac{e^{-\hat{\mu}_i} \hat{\mu}_i^y}{y!}$$

However, the observed variability is greater than $\hat{\mu}_i$.

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

- Overdispersion
- Example of Overdispersion
- Overdispersion: When to be Concerned
- Detecting Overdispersion
- Examples: Detecting Overdispersion
- Implication of Overdispersion
- Adjusting estimated standard errors
- Modeling the extra variability
- Gamma Distributions with $\mu = 1$
- Modeling the extra variability
- Negative Binomial Distribution
- vs Poisson Distribution
- Example: Modeling the extra variability

● Overdispersion and

SAS/GENMOD

● Overdispersion Example from Generalized Linear Models Lindsey

● Possible Model: Poisson

Overdispersion: When to be Concerned

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

- Overdispersion
- Example of Overdispersion
- Overdispersion: When to be Concerned
- Detecting Overdispersion
- Examples: Detecting Overdispersion
- Implication of Overdispersion
- Adjusting estimated standard errors
- Modeling the extra variability
- Gamma Distributions with $\mu = 1$
- Modeling the extra variability
- Negative Binomial Distribution
- vs Poisson Distribution
- Example: Modeling the extra variability

● Overdispersion and

SAS/GENMOD

● Overdispersion Example from Generalized Linear Models Lindsey

● Possible Model: Poisson

When the random component is a distribution where the mean and variance are related.

In particular, overdispersion is

- a concern with Poisson: $\sigma^2 = \mu_i$.
- a concern with Binomial: $\sigma^2 = N\mu_i(1 - \mu_i)$.
- not a concern with Normal: σ^2 is not a function of μ_i .

Detecting Overdispersion

For grouped data, you can compute the mean and sample variance of the counts. If the Poisson distribution is a good model for the data, then mean = variance, but if

sample mean < sample variance \longrightarrow overdispersion

For an example, see Agresti (1996). He illustrates this for the crab data.

For grouped (e.g., crab data grouped into 8 categories instead of 66 different values of width) or ungrouped data (e.g., AIDs and violent incidence examples), if the Poisson model is a good one, then Pearson's X^2 divided by df should equal 1, but if

$$X^2 / df > 1 \longrightarrow \text{overdispersion}$$

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

- Overdispersion
- Example of Overdispersion
- Overdispersion: When to be Concerned
- Detecting Overdispersion
- Examples: Detecting Overdispersion
- Implication of Overdispersion
- Adjusting estimated standard errors
- Modeling the extra variability
- Gamma Distributions with $\mu = 1$
- Modeling the extra variability
- Negative Binomial Distribution
- vs Poisson Distribution
- Example: Modeling the extra variability

● Overdispersion and

SAS/GENMOD

● Overdispersion Example from Generalized Linear Models Lindsey

● Possible Model: Poisson

Examples: Detecting Overdispersion

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

- Overdispersion
- Example of Overdispersion
- Overdispersion: When to be Concerned
- Detecting Overdispersion
- Examples: Detecting Overdispersion
- Implication of Overdispersion
- Adjusting estimated standard errors
- Modeling the extra variability
- Gamma Distributions with $\mu = 1$
- Modeling the extra variability
- Negative Binomial Distribution
- vs Poisson Distribution
- Example: Modeling the extra variability

● Overdispersion and

SAS/GENMOD

● Overdispersion Example from Generalized Linear Models Lindsey

● Possible Model: Poisson

$$X^2/df > 1 \longrightarrow \text{overdispersion}$$

Data Set	df	X^2	X^2/df
Horseshoe crabs	64	174.3	2.7
Deaths due to AIDs	12	15.99	1.33
Violent incidences	793	12711.79	16.03
Hodgkin's disease	38	9956.24	276.94

Implication of Overdispersion

& How to deal with it.

If there is extra variation in the data, then estimates of variances and standard errors for the estimated model parameters are too small. When estimated standard errors are too small, test statistics for testing hypotheses such as $H_o : \beta = 0$ are too big (i.e., “inflated”).

We’ll discuss two ways to deal with this.

- Adjust estimated standard errors — When you’re primarily concerned with testing hypotheses regarding parameter estimates.
- Use an alternative distribution as your random component (i.e., model the extra variability) — When you’re concerned with prediction.

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

- Overdispersion
- Example of Overdispersion
- Overdispersion: When to be Concerned
- Detecting Overdispersion
- Examples: Detecting Overdispersion
- Implication of Overdispersion
- Adjusting estimated standard errors
- Modeling the extra variability
- Gamma Distributions with $\mu = 1$
- Modeling the extra variability
- Negative Binomial Distribution
- vs Poisson Distribution
- Example: Modeling the extra variability

● Overdispersion and

SAS/GENMOD

● Overdispersion Example from Generalized Linear Models Lindsey

● Possible Model: Poisson

Adjusting estimated standard errors

An estimate of the extra variance is Pearson's X^2 for the model divided by its degrees of freedom.

$$X^2 / df$$

To adjust the ASE for parameter estimates we multiply them by

$$\text{Adjusted } ASE = \sqrt{X^2 / df} (ASE)$$

Example: Violent incidences — multiply ASE by $\sqrt{16.03} = 4.00$.

Coefficient	Estimated Parameter	Uncorrected			Corrected		
		ASE	z	p -value	ASE	z	p -value
Intercept	-3.410	.0690	-49.29	< .0001	.2800	-12.31	< .0001
Age	-.045	.0023	-19.69	< .0001	.0091	-4.92	< .0001
Concern	.083	.0075	11.20	< .0001	.0300	2.80	< .0051
History	.420	.0380	11.26	< .0001	.0150	2.81	< .0051

Same conclusion (the effects are very strong in this case).

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

- Overdispersion
- Example of Overdispersion
- Overdispersion: When to be Concerned
- Detecting Overdispersion
- Examples: Detecting Overdispersion
- Implication of Overdispersion
- Adjusting estimated standard errors
- Modeling the extra variability
- Gamma Distributions with $\mu = 1$
- Modeling the extra variability
- Negative Binomial Distribution
- vs Poisson Distribution
- Example: Modeling the extra variability

● Overdispersion and

SAS/GENMOD

● Overdispersion Example from Generalized Linear Models Lindsey

● Possible Model: Poisson

Modeling the extra variability

When you're concerned with prediction (as in the violent incidence of individuals with mental illnesses), simply adjusting ASE for hypothesis testing is not enough.

Model the extra variability: $\tilde{\mu}_i$ (the parameter of the Poisson distribution) is considered a random variable. Even after taking into account the linear predictor, there is still some variability in $\tilde{\mu}_i$ not accounted for, i.e.,

$$\tilde{\mu}_i = t_i \exp\left\{\alpha + \sum_j \beta_j x_{ij}\right\} \epsilon_i$$

where $\epsilon_i > 0$ is an unobserved random variable, and t_i is the offset (if there is one).

The probability distribution assumed for ϵ_i is usually a **Gamma** distribution (for mathematical convenience) with

$$E(\epsilon_i) = 1 \quad \text{and} \quad \text{Var}(\epsilon_i) = 1/\phi$$

So, . . .

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

- Overdispersion
- Example of Overdispersion
- Overdispersion: When to be Concerned
- Detecting Overdispersion
- Examples: Detecting Overdispersion
- Implication of Overdispersion
- Adjusting estimated standard errors
- Modeling the extra variability
- Gamma Distributions with $\mu = 1$
- Modeling the extra variability
- Negative Binomial Distribution
- vs Poisson Distribution
- Example: Modeling the extra variability

● Overdispersion and

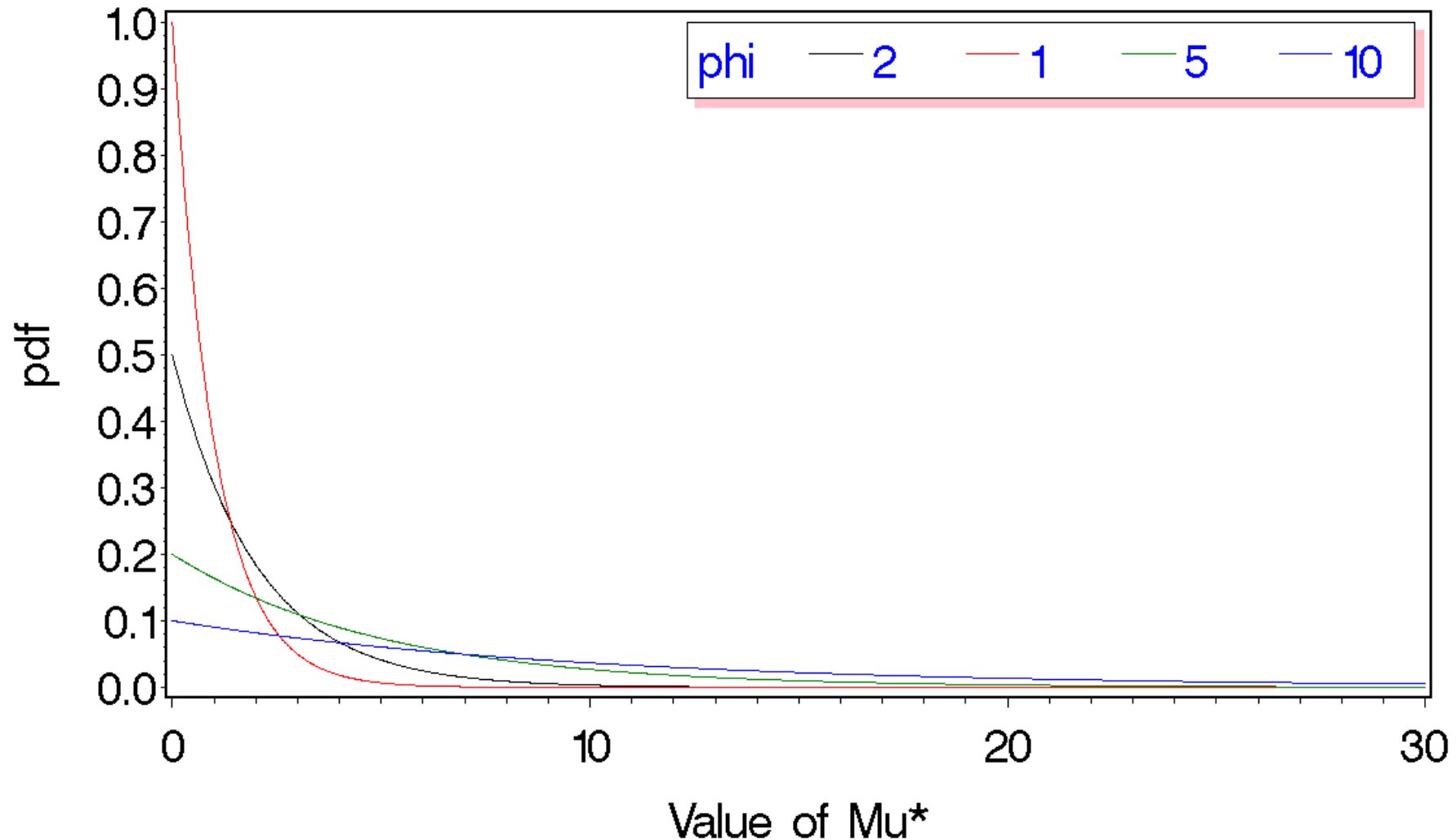
SAS/GENMOD

● Overdispersion Example from Generalized Linear Models Lindsey

● Possible Model: Poisson

Gamma Distributions with $\mu = 1$

Gamma Distribution for Error
 $E(\mu^*) = 1$ & different values of phi



Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

- Overdispersion
- Example of Overdispersion
- Overdispersion: When to be Concerned
- Detecting Overdispersion
- Examples: Detecting Overdispersion
- Implication of Overdispersion
- Adjusting estimated standard errors
- Modeling the extra variability
- Gamma Distributions with $\mu = 1$
- Modeling the extra variability
- Negative Binomial Distribution
- vs Poisson Distribution
- Example: Modeling the extra variability

● Overdispersion and SAS/GENMOD

● Overdispersion Example from Generalized Linear Models Lindsey

● Possible Model: Poisson

Modeling the extra variability

$$E(\tilde{\mu}_i) = E(t_i \exp\{\alpha + \sum_j \beta_j x_{ij}\} \epsilon_i) = E(t_i \exp\{\alpha + \sum_j \beta_j x_{ij}\})$$

and

$$\text{Var}(\tilde{\mu}_i) = \mu_i^2 / \phi$$

and the variance of observed counts $\mu_i + \mu_i^2 / \phi$.

If $\tilde{\mu}_i$ is known, then the distribution of counts y_i would be Poisson with parameter $\tilde{\mu}_i$.

Since $\tilde{\mu}_i$ is not known and is a random variable, the distribution for y_i is a **Negative Binomial Distribution**.

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

- Overdispersion
- Example of Overdispersion
- Overdispersion: When to be Concerned
- Detecting Overdispersion
- Examples: Detecting Overdispersion
- Implication of Overdispersion
- Adjusting estimated standard errors
- Modeling the extra variability
- Gamma Distributions with $\mu = 1$
- Modeling the extra variability
- Negative Binomial Distribution
- vs Poisson Distribution
- Example: Modeling the extra variability

● Overdispersion and

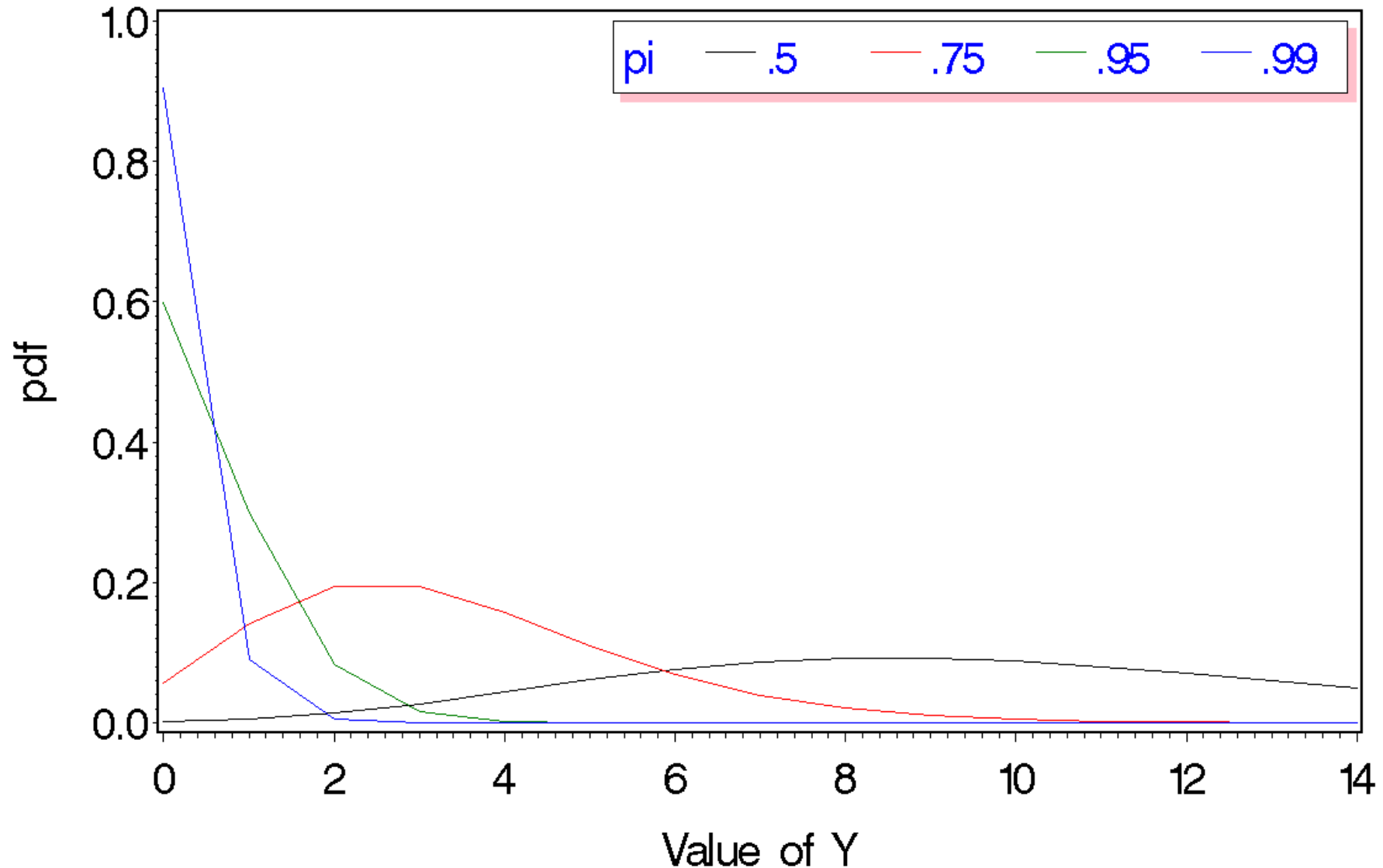
SAS/GENMOD

● Overdispersion Example from Generalized Linear Models Lindsey

● Possible Model: Poisson

Negative Binomial Distribution

Negative Binominal Distribution (n= 10)



Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

- Overdispersion
- Example of Overdispersion
- Overdispersion: When to be Concerned
- Detecting Overdispersion
- Examples: Detecting Overdispersion
- Implication of Overdispersion
- Adjusting estimated standard errors
- Modeling the extra variability
- Gamma Distributions with $\mu = 1$
- Modeling the extra variability
- Negative Binomial Distribution
- vs Poisson Distribution
- Example: Modeling the extra variability

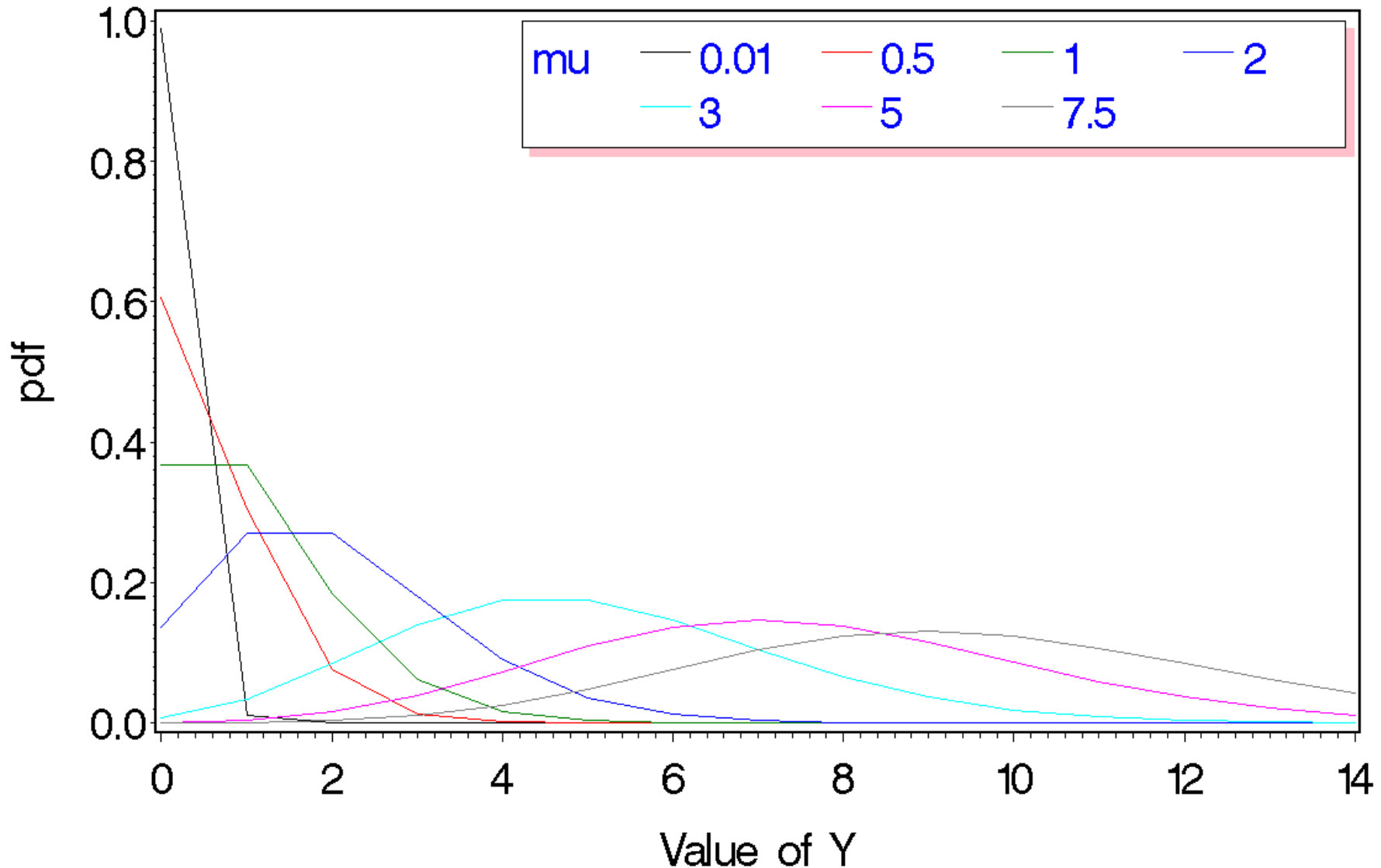
● Overdispersion and SAS/GENMOD

● Overdispersion Example from Generalized Linear Models Lindsey

● Possible Model: Poisson

vs Poisson Distribution

Poisson Distribution



Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

- Overdispersion
- Example of Overdispersion
- Overdispersion: When to be Concerned
- Detecting Overdispersion
- Examples: Detecting Overdispersion
- Implication of Overdispersion
- Adjusting estimated standard errors
- Modeling the extra variability
- Gamma Distributions with $\mu = 1$
- Modeling the extra variability
- Negative Binomial Distribution
- vs Poisson Distribution
- Example: Modeling the extra variability

● Overdispersion and SAS/GENMOD

● Overdispersion Example from Generalized Linear Models Lindsey

● Possible Model: Poisson

Example: Modeling the extra variability

For the violent incidents example, when the negative binomial distribution is used as the random component of the GLM, the following estimated parameters are obtained:

Coefficient	Estimate	ASE	z	p -value
Intercept	-3.5500	.26200	-13.55	< .0001
Age	-.0459	.00799	-5.74	< .0001
Concern	.0962	.03090	3.12	.0029
History	.5360	.11550	3.45	.0008

which is similar to before., but the probability distribution is much better approximated by the Negative Binomial.

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

- Overdispersion
- Example of Overdispersion
- Overdispersion: When to be Concerned
- Detecting Overdispersion
- Examples: Detecting Overdispersion
- Implication of Overdispersion
- Adjusting estimated standard errors
- Modeling the extra variability
- Gamma Distributions with $\mu = 1$
- Modeling the extra variability
- Negative Binomial Distribution
- vs Poisson Distribution
- Example: Modeling the extra variability

● Overdispersion and

SAS/GENMOD

● Overdispersion Example from Generalized Linear Models Lindsey

● Possible Model: Poisson

Overdispersion and SAS/GENMOD

■ Built in MODEL options

From the SAS/GENMOD documentation:

- ◆ **SCALE** = number
- ◆ **PSCALE** sets scaling parameter equal to 1 during estimation but standard errors and statistics are adjusted using Pearson's X^2/df .
- ◆ **DSCALE** same as PSCALE but uses G^2/df .

■ Using Negative Binomial distribution as the random component.

model count = < linear predictor > / **link**=log **dist**=Negbin ;

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

- Overdispersion
- Example of Overdispersion
- Overdispersion: When to be Concerned
- Detecting Overdispersion
- Examples: Detecting Overdispersion
- Implication of Overdispersion
- Adjusting estimated standard errors
- Modeling the extra variability
- Gamma Distributions with $\mu = 1$
- Modeling the extra variability
- Negative Binomial Distribution
- vs Poisson Distribution
- Example: Modeling the extra variability

● Overdispersion and

SAS/GENMOD

● Overdispersion Example from Generalized Linear Models Lindsey

● Possible Model: Poisson

Overdispersion Example from Lindsey

(1997); *Applying Generalized Linear Models*

The data consist of counts of T_4 cells/mm in blood samples from 20 patients in remission from Hodgkin's disease and 20 other patients in remission from disseminated malignancies:

Hodgkin's Disease		Non-Hodgkin's Disease	
396	568	375	375
1212	171	752	208
554	1104	151	116
257	435	736	192
295	397	315	1252
288	1004	657	700
431	795	440	771
1621	1378	688	426
902	958	410	979
1283	2415	377	503

Question: Is the average count of T_4 cells/mm the same or different for patients in remission from Hodgkin's disease as the average count from those in remission from disseminated malignancies?

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

- Overdispersion
- Example of Overdispersion
- Overdispersion: When to be Concerned
- Detecting Overdispersion
- Examples: Detecting Overdispersion
- Implication of Overdispersion
- Adjusting estimated standard errors
- Modeling the extra variability
- Gamma Distributions with $\mu = 1$
- Modeling the extra variability
- Negative Binomial Distribution
- vs Poisson Distribution
- Example: Modeling the extra variability

● Overdispersion and

SAS/GENMOD

● Overdispersion Example from Generalized Linear Models Lindsey

● Possible Model: Poisson

Possible Model: Poisson regression

Hodgkin's disease example (continued)

- **Random component:** Y = number of T_4 cells/mm.
Assume Poisson and if this doesn't fit, we'll try Negative Binomial.
- **Systematic component:** In this example, the explanatory/predictor variables is discrete so we'll define

$$X = \begin{cases} 0 & \text{if non-Hodgkin's} \\ 1 & \text{if Hodgkin's disease} \end{cases}$$

So linear predictor is

$$\alpha + \beta x$$

- **Link function:** \log .

$$\log(\mu_y) = \begin{cases} \alpha & \text{if non-Hodgkin's} \\ \alpha + \beta & \text{if Hodgkin's disease} \end{cases}$$

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

- Overdispersion
- Example of Overdispersion
- Overdispersion: When to be Concerned
- Detecting Overdispersion
- Examples: Detecting Overdispersion
- Implication of Overdispersion
- Adjusting estimated standard errors
- Modeling the extra variability
- Gamma Distributions with $\mu = 1$
- Modeling the extra variability
- Negative Binomial Distribution
- vs Poisson Distribution
- Example: Modeling the extra variability

● Overdispersion and

SAS/GENMOD

● Overdispersion Example from Generalized Linear Models Lindsey

● Possible Model: Poisson

Possible Model: Poisson regression

Hodgkin's disease example (continued)

Fit of this model to data yields

Model	df	G^2	p -value	X^2	p -value
Poisson regression	38	9956.23	< .001	10523.75	< .001
Negative binomial	38	42.41	.29	40.26	.37

Parameter Estimates from the two models:

Parameter	df	Poisson Distribution			Negative Binomial		
		Estimate	ASE	Wald	Estimate	ASE	Wald
α	1	6.2560	.0098	407935	6.2560	.1365	2101.53
β	1	.4572	.0125	1333.90	.4572	.1929	5.62

Note: dispersion parameter = .3706 and ASE = .0787 (95% CI: 0.2163 0.5248). This was estimated by MLE.

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

- Overdispersion
- Example of Overdispersion
- Overdispersion: When to be Concerned
- Detecting Overdispersion
- Examples: Detecting Overdispersion
- Implication of Overdispersion
- Adjusting estimated standard errors
- Modeling the extra variability
- Gamma Distributions with $\mu = 1$
- Modeling the extra variability
- Negative Binomial Distribution
- vs Poisson Distribution
- Example: Modeling the extra variability

● Overdispersion and

SAS/GENMOD

● Overdispersion Example from Generalized Linear Models Lindsey

● Possible Model: Poisson

Outline of Bully Example

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

- Outline of Bully Example
- The Problem
- Exploratory Analysis
- The Predictor Variable
- The Predictor Variable
- SAS for these
- SAS for Distribution of Peer Nominations
- SAS for Bully Scale
- Relationship Between the Measures
- We Know So Far That. . .
- Initial Modeling
- Fit of Poisson Regression Model
- Fit of Marginal Distribution

- Revised Model
- Fit Statistics & Parameter

- Estimates
- Generalized Linear Models
 - Fit of Negative Binomial

Model to Data

- The Problem
- A little exploratory analysis.
- Initial Modeling
- Revised models
- Conclusion
- Zero inflated models (something new)

The Problem

- The data are from:

Espelage, D.L., Holt, M.K., & Henkel, R.R. (2004). Examination of peer-group contextual effects on aggression during early adolescence. *Child Development*, 74, 205–220.

- Two ways to measure bullying

- ◆ **Self Report:** 9 item Illinois Bully Scale (Espelage & Holt, 2001).
- ◆ **Peer nominations:** Kids list everyone who they view as a bully. The total number of nominations a child receives is a measure of bullying that child's bullying.

- Peer nominations more “objective” than self report and it's getting harder to obtain IRB approval of peer nominations.

- Model peer nominations (a count) with self report measure (bully scale) as a predictor variable... ignoring clustering...

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

- Outline of Bully Example
- The Problem
- Exploratory Analysis
- The Predictor Variable
- The Predictor Variable
- SAS for these
- SAS for Distribution of Peer Nominations
- SAS for Bully Scale
- Relationship Between the Measures
- We Know So Far That...
- Initial Modeling
- Fit of Poisson Regression Model
- Fit of Marginal Distribution

- Revised Model
- Fit Statistics & Parameter

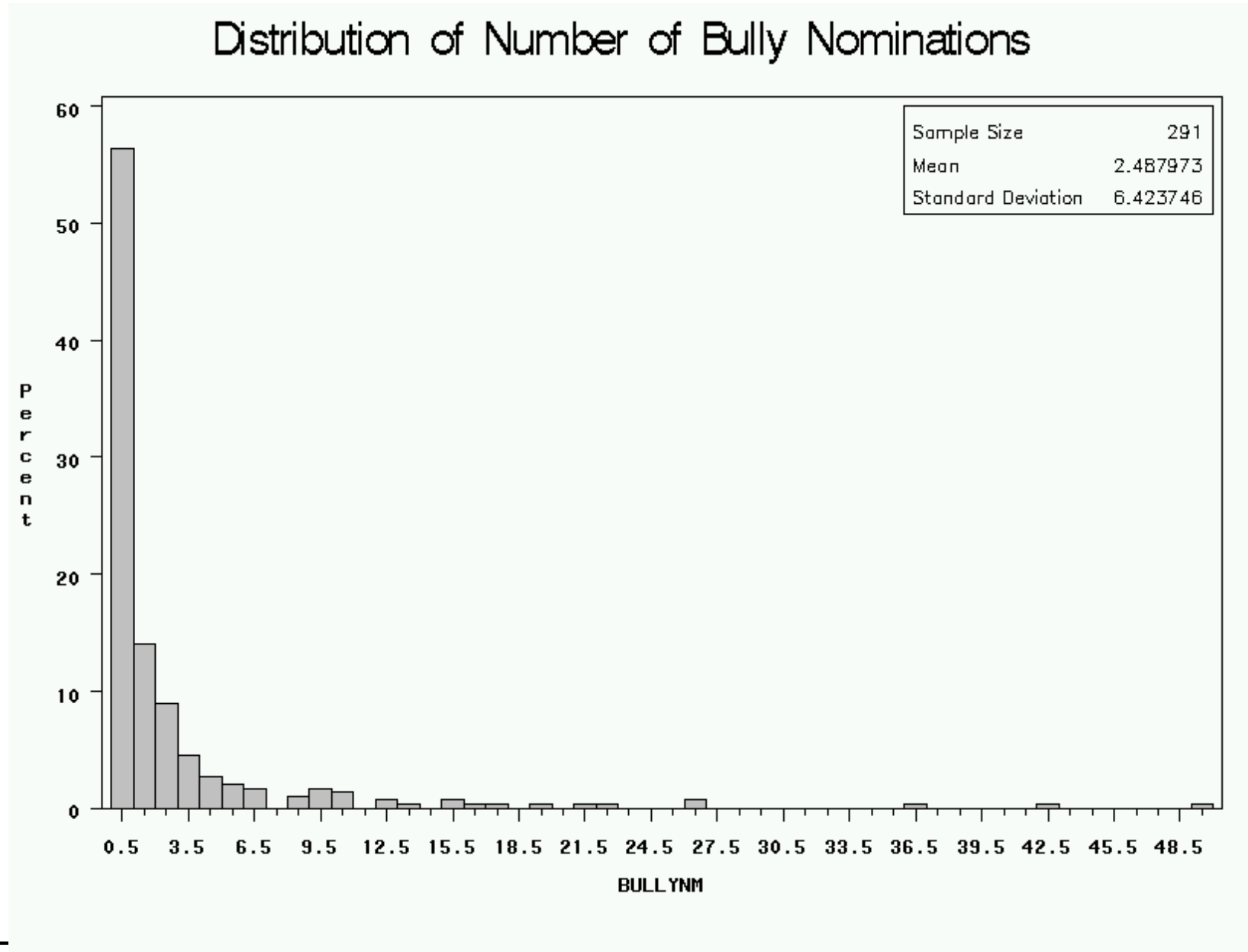
- Estimates
- Generalized Linear Models
 - Fit of Negative Binomial

Model to Data

Exploratory Analysis

- Overview
- Inference for model parameters
- Assessing Model Fit
- Global Fit Statistics
- Residuals
- Confidence intervals for fitted values
- Overdispersion
- Overview of Bully Example
 - Outline of Bully Example
 - The Problem
 - Exploratory Analysis
 - The Predictor Variable
 - The Predictor Variable
 - SAS for these
 - SAS for Distribution of Peer Nominations
 - SAS for Bully Scale
 - Relationship Between the Measures
 - We Know So Far That...
 - Initial Modeling
 - Fit of Poisson Regression Model
 - Fit of Marginal Distribution

- Revised Model
- Fit Statistics & Parameter Estimates
- Generalized Linear Models
- Fit of Negative Binomial Model to Data



The Predictor Variable

- Overview

- Inference for model parameters

- Assessing Model Fit

- Global Fit Statistics

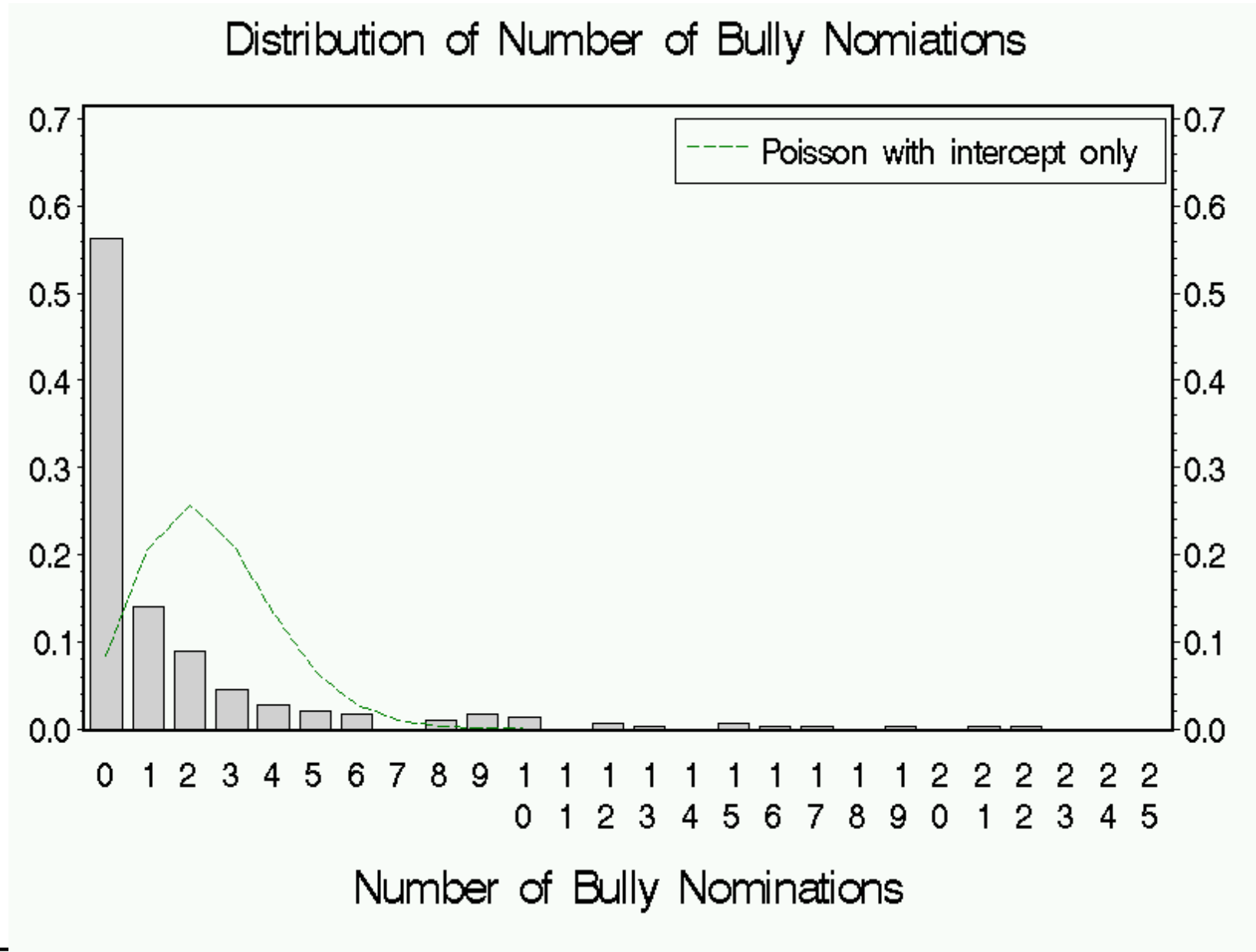
- Residuals

- Confidence intervals for fitted values

- Overdispersion

- Overview of Bully Example

 - Outline of Bully Example
 - The Problem
 - Exploratory Analysis
 - The Predictor Variable
 - The Predictor Variable
 - SAS for these
 - SAS for Distribution of Peer Nominations
 - SAS for Bully Scale
 - Relationship Between the Measures
 - We Know So Far That...
 - Initial Modeling
 - Fit of Poisson Regression Model
 - Fit of Marginal Distribution

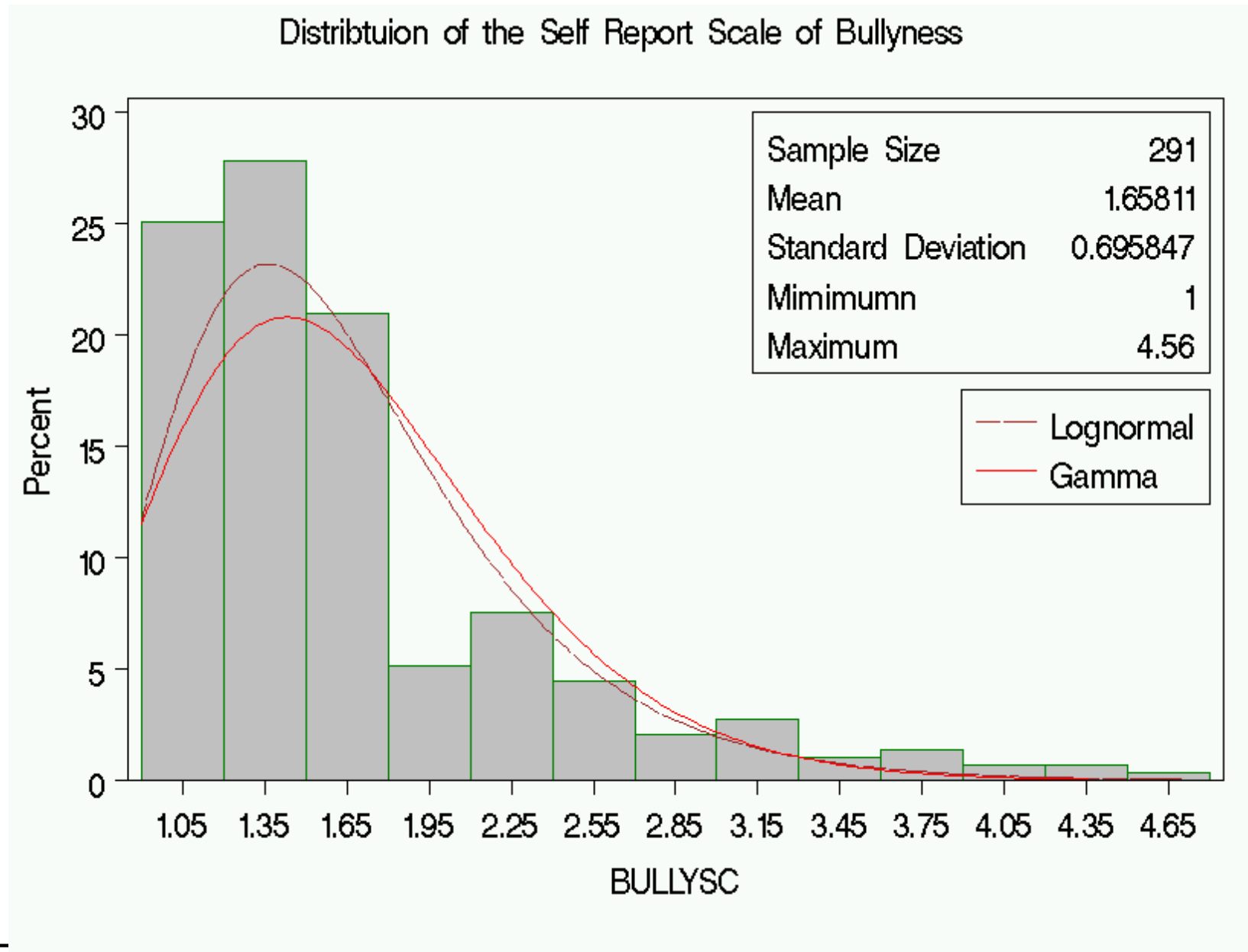


- Revised Model
- Fit Statistics & Parameter
- Estimates

 - Generalized Linear Models
 - Fit of Negative Binomial

- Model to Data

The Predictor Variable



Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

- Outline of Bully Example
- The Problem
- Exploratory Analysis
- The Predictor Variable
- The Predictor Variable
- SAS for these
- SAS for Distribution of Peer Nominations
- SAS for Bully Scale
- Relationship Between the Measures
- We Know So Far That...
- Initial Modeling
- Fit of Poisson Regression Model
- Fit of Marginal Distribution

- Revised Model
- Fit Statistics & Parameter

- Estimates
- Generalized Linear Models
 - Fit of Negative Binomial

Model to Data

SAS for these

```

data bullynom;
  input          BULLYSC          BULLYNM;
  datalines;
    1.56          0.00
    1.56          0.00
    1.11          0.00
    1.56          0.00
    1.22          4.00
    . . .
    3.22          0.00
    1.22          0.00
    1.89          0.00
run;

```

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

- Outline of Bully Example
- The Problem
- Exploratory Analysis
- The Predictor Variable
- The Predictor Variable
- SAS for these
- SAS for Distribution of Peer Nominations
- SAS for Bully Scale
- Relationship Between the Measures
- We Know So Far That...
- Initial Modeling
- Fit of Poisson Regression Model
- Fit of Marginal Distribution

- Revised Model
- Fit Statistics & Parameter

- Estimates
- Generalized Linear Models
 - Fit of Negative Binomial

Model to Data

SAS for Distribution of Peer Nominations

```
options nogstyle; ods select Quantiles MyHist;
proc univariate data=bullynom;
    var BULLYNM;
    histogram BULLYNM/ cfill=ltgray
        midpoints = .5 1.5 2.5 3.5 4.5
        5.5 6.5 7.5 8.5 9.5 10.5
        11.5 12.5 13.5 14.5 15.5
        16.5
        name='MyHist'
        ;
    inset n='Sample Size' mean='Mean'
        std='Standard Deviation'
        / position=ne;
    Title 'Distribution of Number of Bully Nominations';
run;
options gstyle;
```

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

- Outline of Bully Example
- The Problem
- Exploratory Analysis
- The Predictor Variable
- The Predictor Variable
- SAS for these
- SAS for Distribution of Peer Nominations
- SAS for Bully Scale
- Relationship Between the Measures
- We Know So Far That...
- Initial Modeling
- Fit of Poisson Regression Model
- Fit of Marginal Distribution

- Revised Model
- Fit Statistics & Parameter

- Estimates
- Generalized Linear Models
- Fit of Negative Binomial

Model to Data

SAS for Bully Scale

```
options nogstyle;
ods select Quantiles BullyScaleHist;
proc univariate data=bullynom;
  var bullysc;
  histogram bullysc / lognormal
                    gamma
                    cfill=ltgray
                    name='BullyScaleHist';
  inset n='Sample Size' mean='Mean'
        std='Standard Deviation' min='Mimimumn'
        max='Maximum' / position=ne;
  inset lognormal gamma / position=e;
  Title 'Distribtuion of the Self Report Scale of
        Bullyness';
run;
options gstyle;
```

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

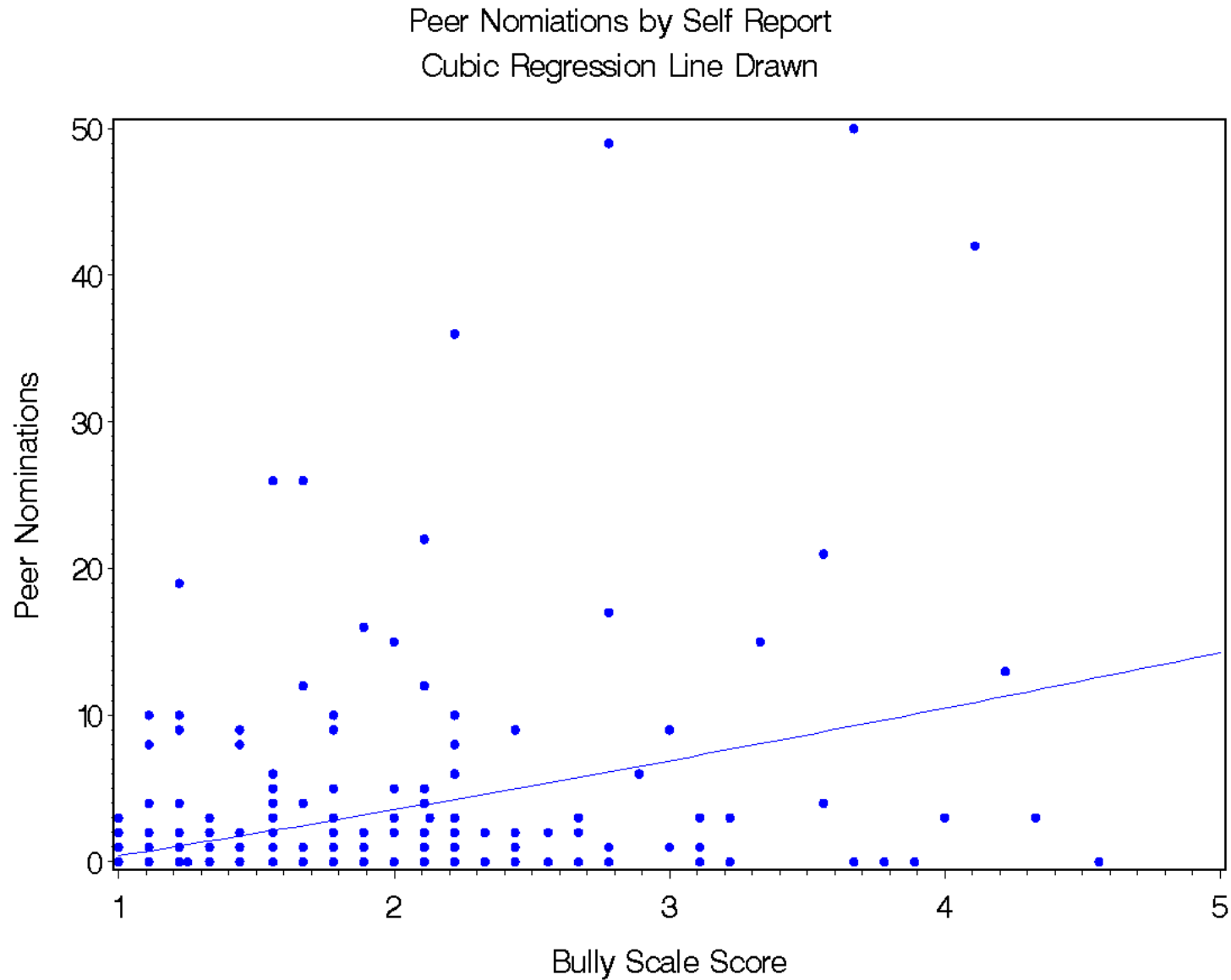
- Outline of Bully Example
- The Problem
- Exploratory Analysis
- The Predictor Variable
- The Predictor Variable
- SAS for these
- SAS for Distribution of Peer Nominations
- SAS for Bully Scale
- Relationship Between the Measures
- We Know So Far That...
- Initial Modeling
- Fit of Poisson Regression Model
- Fit of Marginal Distribution

- Revised Model
- Fit Statistics & Parameter

- Estimates
- Generalized Linear Models
 - Fit of Negative Binomial

Model to Data

Relationship Between the Measures



Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

- Outline of Bully Example
- The Problem
- Exploratory Analysis
- The Predictor Variable
- The Predictor Variable
- SAS for these
- SAS for Distribution of Peer Nominations
- SAS for Bully Scale
- Relationship Between the Measures
- We Know So Far That...
- Initial Modeling
- Fit of Poisson Regression Model
- Fit of Marginal Distribution

- Revised Model
- Fit Statistics & Parameter

Estimates

- Generalized Linear Models
- Fit of Negative Binomial

Model to Data

We Know So Far That...

- Both variables are highly positively skewed.
- There are a lot of kids who did not receive any peer nominations.
- There does appear to be a relationship between peer nominations and scale score.
- Mean peer nominations is much smaller than the variance:

$$2.49 < 41.26$$

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

- Outline of Bully Example
- The Problem
- Exploratory Analysis
- The Predictor Variable
- The Predictor Variable
- SAS for these
- SAS for Distribution of Peer Nominations
- SAS for Bully Scale
- Relationship Between the Measures
- We Know So Far That...
- Initial Modeling
- Fit of Poisson Regression Model
- Fit of Marginal Distribution

- Revised Model
- Fit Statistics & Parameter

- Estimates
- Generalized Linear Models
- Fit of Negative Binomial

Model to Data

Initial Modeling

Starting Model:

Random Component:

- Y_{ij} = the number of nominations received by kid i in peer group j .
- **Poisson distribution.**

Linear Predictor:

$$\beta_0 + \beta_1(\text{bullysc})_{ij} = \beta_0 + \beta_1 x_{ij}$$

The Link is the **Log**, the canonical link.

The initial model is a standard Poisson regression model

$$E(Y_{ij}) = \mu_{ij} = \exp[\beta_0 + \beta_1 x_{ij}]$$

where

$$P(Y_{ij} = y) = \frac{e^{-\mu_{ij}} \mu_{ij}^y}{y_{ij}!}$$

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

- Outline of Bully Example
- The Problem
- Exploratory Analysis
- The Predictor Variable
- The Predictor Variable
- SAS for these
- SAS for Distribution of Peer Nominations
- SAS for Bully Scale
- Relationship Between the Measures
- We Know So Far That...
- Initial Modeling
- Fit of Poisson Regression Model
- Fit of Marginal Distribution

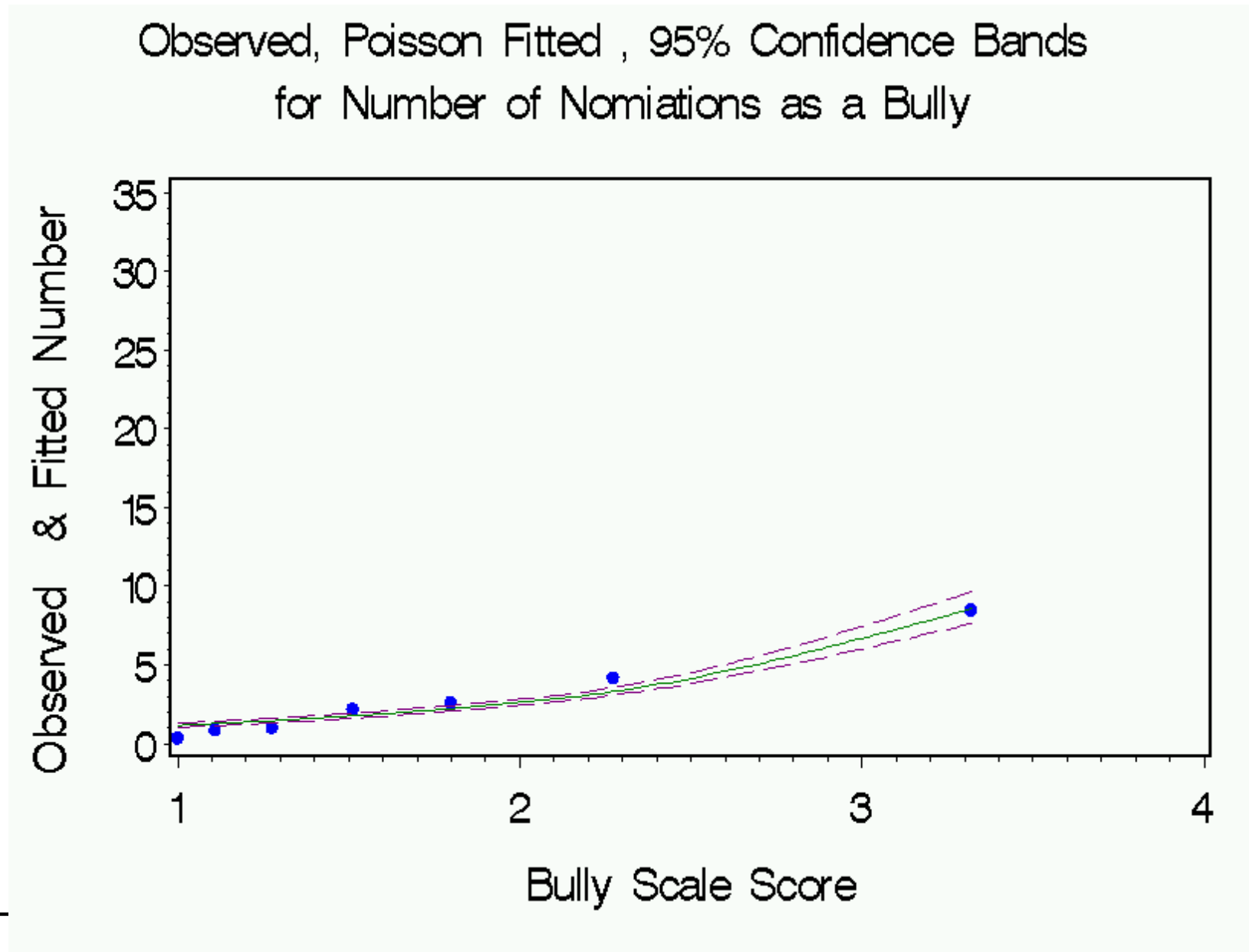
- Revised Model
- Fit Statistics & Parameter

- Estimates
- Generalized Linear Models
- Fit of Negative Binomial

Model to Data

Fit of Poisson Regression Model

(model fit and then grouped to “look” at fit)



Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

- Outline of Bully Example
- The Problem
- Exploratory Analysis
- The Predictor Variable
- The Predictor Variable
- SAS for these
- SAS for Distribution of Peer Nominations
- SAS for Bully Scale
- Relationship Between the Measures
- We Know So Far That ...
- Initial Modeling
- Fit of Poisson Regression Model
- Fit of Marginal Distribution

- Revised Model
- Fit Statistics & Parameter

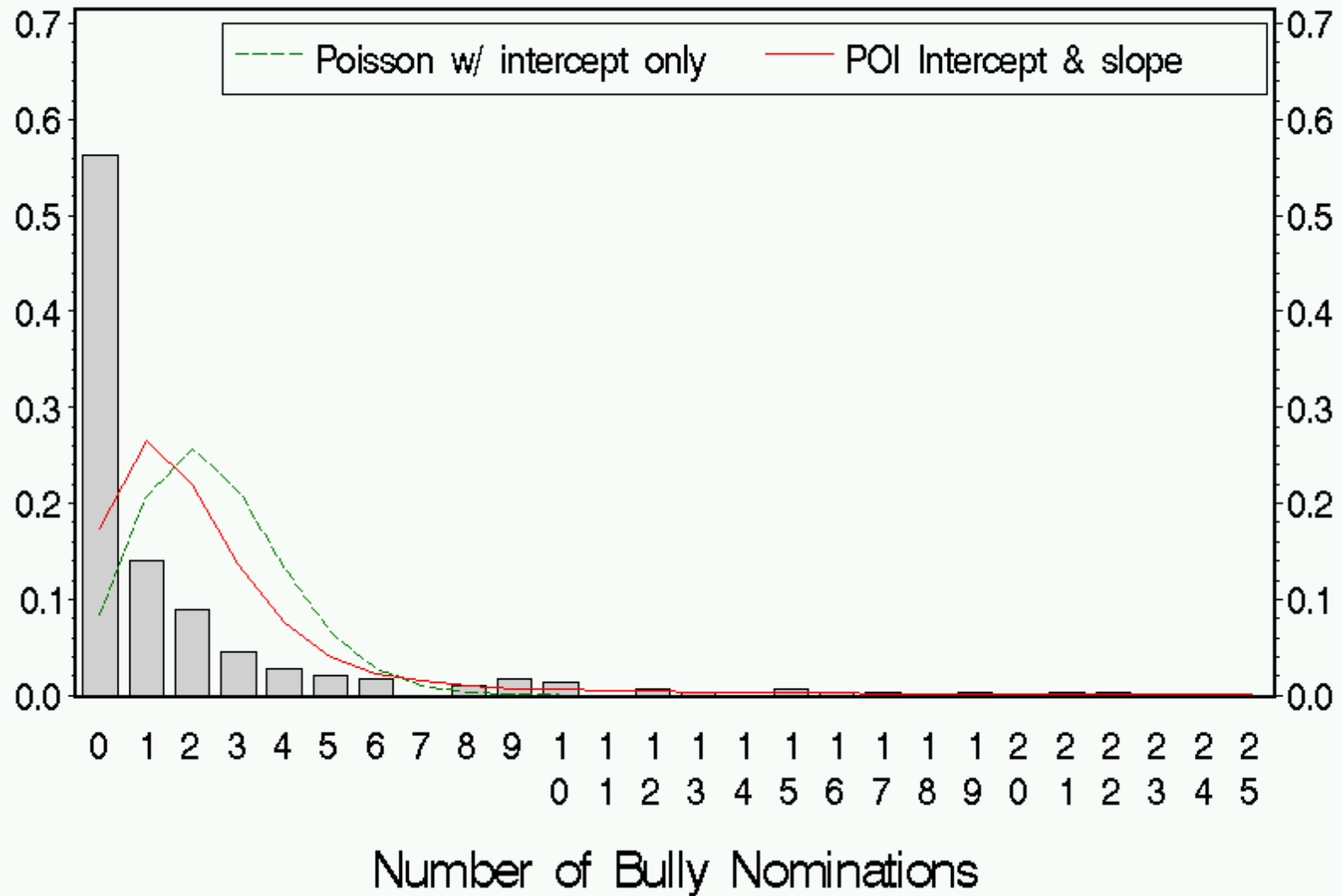
Estimates

- Generalized Linear Models
- Fit of Negative Binomial

Model to Data

Fit of Marginal Distribution

Distribution of Number of Bully Nominations & Fitted Values



Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

- Outline of Bully Example
- The Problem
- Exploratory Analysis
- The Predictor Variable
- The Predictor Variable
- SAS for these
- SAS for Distribution of Peer Nominations
- SAS for Bully Scale
- Relationship Between the Measures
- We Know So Far That...
- Initial Modeling
- Fit of Poisson Regression Model
- Fit of Marginal Distribution

- Revised Model
- Fit Statistics & Parameter

Estimates

- Generalized Linear Models
- Fit of Negative Binomial

Model to Data

Revised Model

To deal with the overdispersion, we'll change the random component to Negative Binomial.

The GLM model has

- Random= **Negative Binominal**
- Linear predictor= $\beta_0 + \beta_1 x_{ij}$.
- log link
- Our next model is

$$Y_{ij} = \underbrace{\mu_{ij}}_{\text{Poisson}} \underbrace{\epsilon_{ij}}_{\text{Gamma}}$$

where

- ◆ $E(\epsilon_{ij}) = 1$
- ◆ $\text{var}(\epsilon_{ij}) = 1/\phi$ (ϕ is the “dispersion” parameter).
- ◆ $E(Y_{ij}|x_{ij}) = \mu_{ij} = \exp[\beta_0 + \beta_1 x_{ij}]$
- ◆ $\text{var}(Y_{ij}|x_{ij}) = \mu_{ij} + \mu_{ij}^2/\phi$
- ◆ and

$$P(Y_{ij} = y) = \frac{\Gamma(y + \phi)}{y! \Gamma(\phi)} \left(\frac{\phi}{\phi + \mu_{ij}} \right)^\phi \left(\frac{\mu_{ij}}{\phi + \mu_{ij}} \right)^y$$

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

- Outline of Bully Example
- The Problem
- Exploratory Analysis
- The Predictor Variable
- The Predictor Variable
- SAS for these
- SAS for Distribution of Peer Nominations
- SAS for Bully Scale
- Relationship Between the Measures
- We Know So Far That...
- Initial Modeling
- Fit of Poisson Regression Model
- Fit of Marginal Distribution

- Revised Model
- Fit Statistics & Parameter

- Estimates
- Generalized Linear Models
- Fit of Negative Binomial

Model to Data

Fit Statistics & Parameter Estimates

$df = 289$ for all of these

Dist	Link	G^2	X^2	X^2/df	AIC	BIC
Poisson	log	1774.60	2977.94	10.30	2160	2168
NegBin	log	244.08	351.13	1.22	998	1010

Parm	Poisson				Negative Binomial			
	est.	se	Wald	p	est.	se	Wald	p
β_0	-0.66	0.09	55.87	< .01	-1.19	0.36	11.30	< .01
β_1	0.81	0.04	540.05	< .01	1.09	0.20	30.66	< .01
$1/\phi$	—				3.50	0.44	—	

For interpretation,

$$\exp(.81) = 2.25 \quad \text{and} \quad \exp(1.09) = 2.98$$

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

- Outline of Bully Example
- The Problem
- Exploratory Analysis
- The Predictor Variable
- The Predictor Variable
- SAS for these
- SAS for Distribution of Peer Nominations
- SAS for Bully Scale
- Relationship Between the Measures
- We Know So Far That...
- Initial Modeling
- Fit of Poisson Regression Model
- Fit of Marginal Distribution

- Revised Model
- Fit Statistics & Parameter

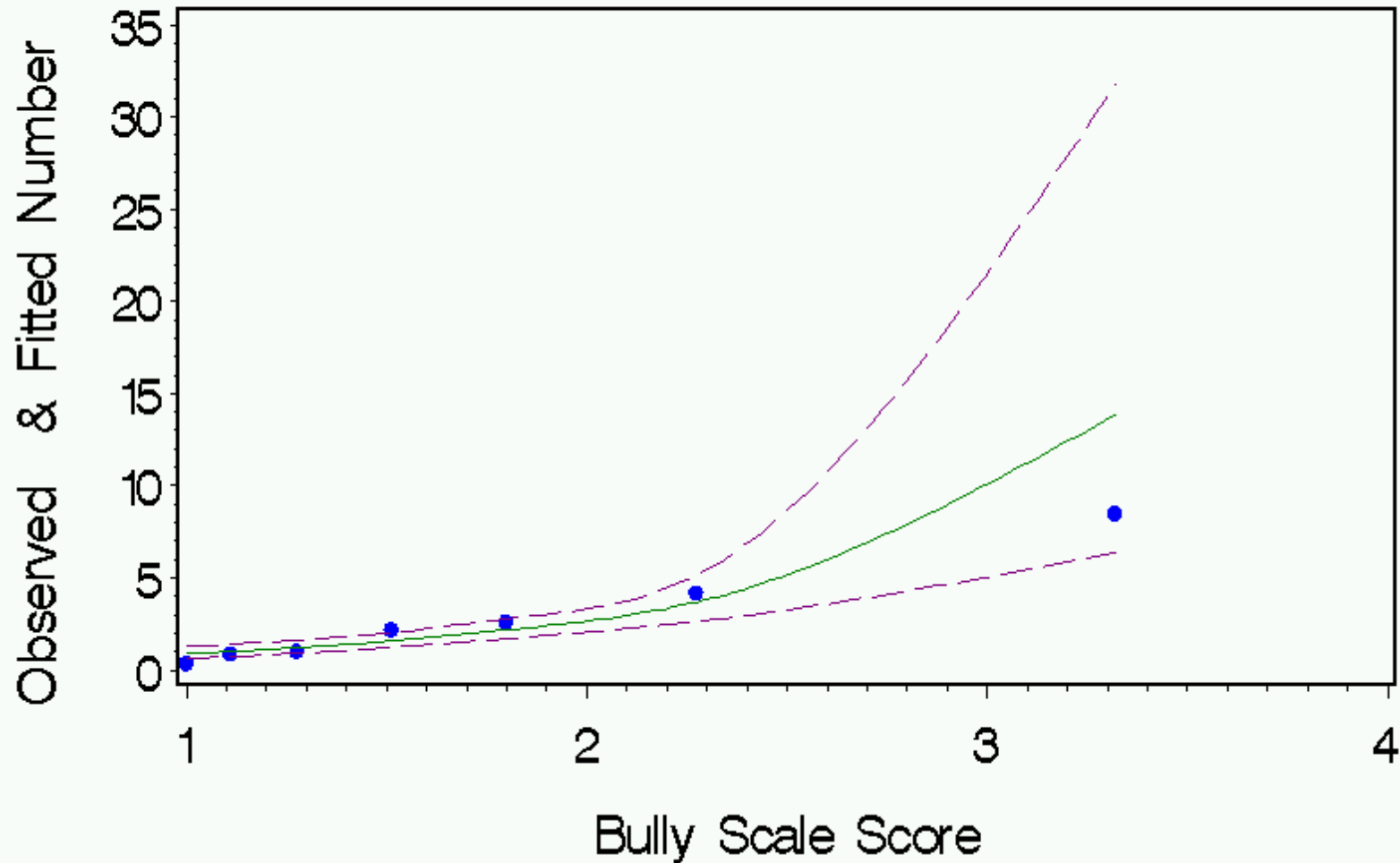
Estimates

- Generalized Linear Models
- Fit of Negative Binomial

Model to Data

Fit of Negative Binomial Model to Data

Observed, Negative Binomial Fitted, & 95% Confidence Bands with a Log Link Function



Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

- Outline of Bully Example
- The Problem
- Exploratory Analysis
- The Predictor Variable
- The Predictor Variable
- SAS for these
- SAS for Distribution of Peer Nominations
- SAS for Bully Scale
- Relationship Between the Measures
- We Know So Far That...
- Initial Modeling
- Fit of Poisson Regression Model
- Fit of Marginal Distribution

- Revised Model
- Fit Statistics & Parameter

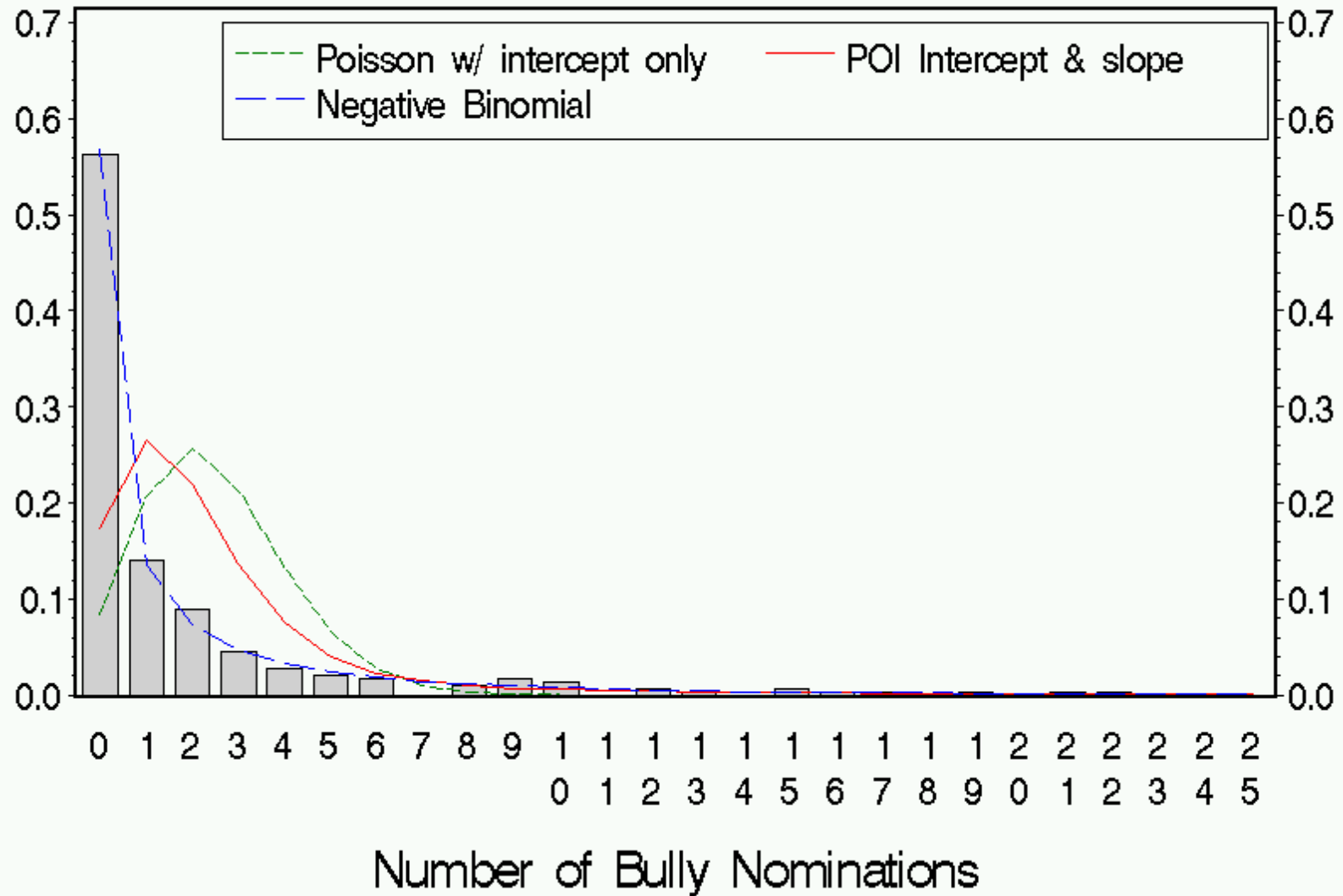
Estimates

- Generalized Linear Models
- Fit of Negative Binomial

Model to Data

Fit of Marginal Distribution

Distribution of Number of Bully Nominations & Fitted Values



- Overview

- Inference for model parameters

- Assessing Model Fit

- Global Fit Statistics

- Residuals

- Confidence intervals for fitted values

- Overdispersion

- Overview of Bully Example

 - Outline of Bully Example
 - The Problem
 - Exploratory Analysis
 - The Predictor Variable
 - The Predictor Variable
 - SAS for these
 - SAS for Distribution of Peer Nominations
 - SAS for Bully Scale
 - Relationship Between the Measures
 - We Know So Far That...
 - Initial Modeling
 - Fit of Poisson Regression Model
 - Fit of Marginal Distribution

- Revised Model
- Fit Statistics & Parameter
- Estimates

 - Generalized Linear Models
 - Fit of Negative Binomial

- Model to Data

Change of the Link Function

The relationship between Y_{ij} and x_{ij} looks like a straight line... The New GLM:

- Negative Binomial

- $\beta_0 + \beta_1 x_{ij}$

- Identity Link function

This model is

$$Y_{ij} = \underbrace{\mu_{ij}}_{\text{Poisson}} \epsilon_{ij} = \underbrace{(\beta_0 + \beta_1 x_{ij})}_{\text{Poisson}} \underbrace{\epsilon_{ij}}_{\text{Gamma}}$$

- $E(\epsilon_{ij}) = 1$

- $\text{var}(\epsilon_{ij}) = 1/\phi$

- $E(Y_{ij} | x_{ij}) = \mu_{ij} = \beta_0 + \beta_1 x_{ij}$

- and

$$P(Y_{ij} = y) = \frac{\Gamma(y + \phi)}{y! \Gamma(\phi)} \left(\frac{\phi}{\phi + \mu_{ij}} \right)^\phi \left(\frac{\mu_{ij}}{\phi + \mu_{ij}} \right)^y$$

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

- Outline of Bully Example
- The Problem
- Exploratory Analysis
- The Predictor Variable
- The Predictor Variable
- SAS for these
- SAS for Distribution of Peer Nominations
- SAS for Bully Scale
- Relationship Between the Measures
- We Know So Far That...
- Initial Modeling
- Fit of Poisson Regression Model
- Fit of Marginal Distribution

- Revised Model
- Fit Statistics & Parameter

- Estimates
- Generalized Linear Models
 - Fit of Negative Binomial

Model to Data

Fit Statistics & Parameter Estimates

$df = 289$ for all of these

Dist	Link	G^2	X^2	X^2/df	AIC	BIC
Poisson	log	1774.60	2977.94	10.30	2160	2168
NegBin	log	244.08	351.13	1.22	998	1010
NegBin	Identity	244.33	337.09	1.17	992	1003

Parm	Log Link				Identity Link			
	est.	se	Wald	p	est.	se	Wald	p
β_0	-1.19	0.36	11.30	< .01	-2.64	0.65	16.44	< .01
β_1	1.09	0.20	30.66	< .01	3.07	0.57	29.50	< .01
$1/\phi$	3.50	0.44	—		3.35	0.43		

For interpretation, a one unit change in bully scale leads to

$\exp(1.09) = 2.98$ times larger or 3.07 more nominations

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

- Outline of Bully Example
- The Problem
- Exploratory Analysis
- The Predictor Variable
- The Predictor Variable
- SAS for these
- SAS for Distribution of Peer Nominations
- SAS for Bully Scale
- Relationship Between the Measures
- We Know So Far That...
- Initial Modeling
- Fit of Poisson Regression Model
- Fit of Marginal Distribution

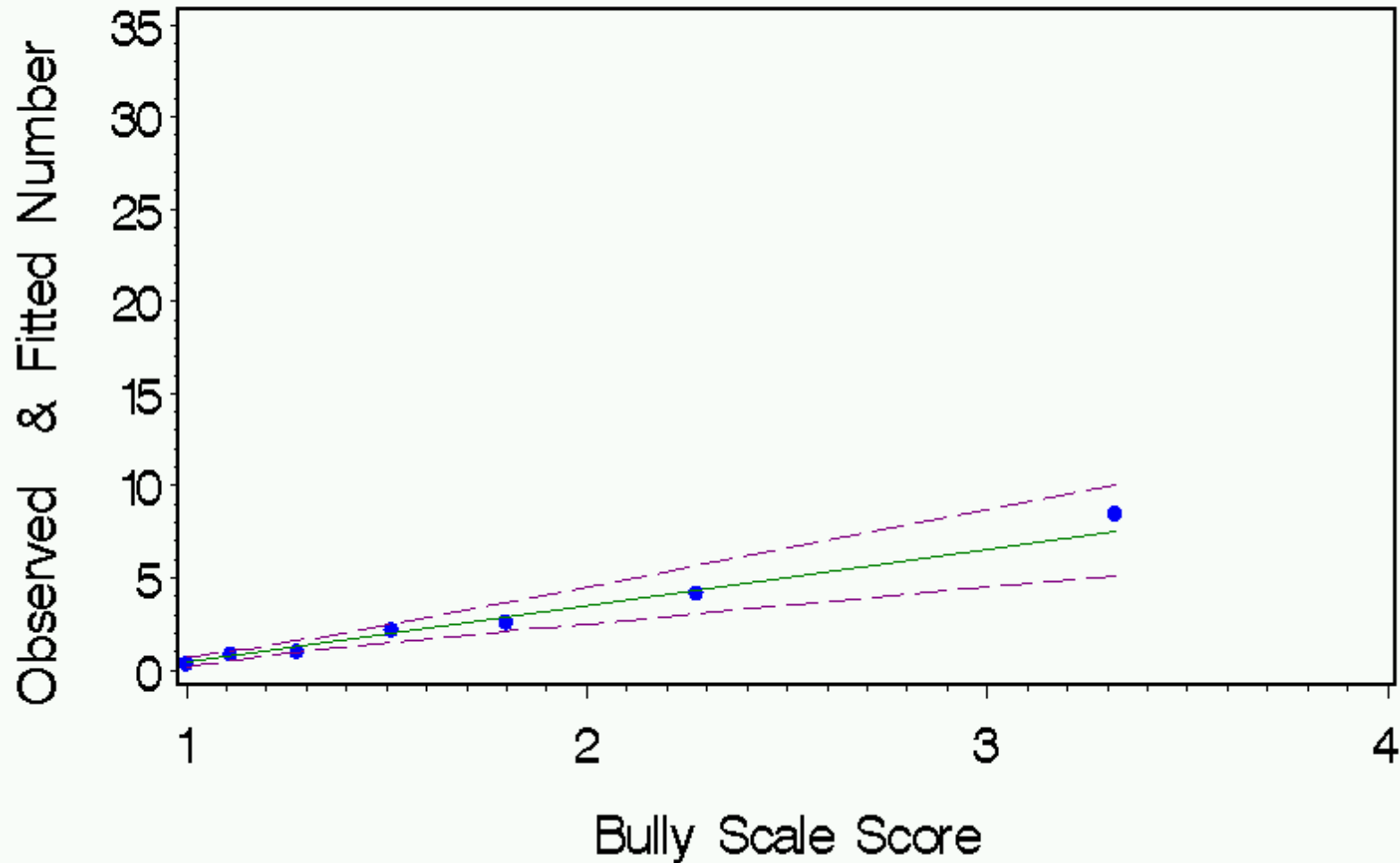
- Revised Model
- Fit Statistics & Parameter

- Estimates
- Generalized Linear Models
- Fit of Negative Binomial

Model to Data

Fit of Negative Binomial Model w/ Identity

Observed, Negative Binomial Fitted, & 95% Confidence Bands with an Identity Link Function



Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

- Outline of Bully Example
- The Problem
- Exploratory Analysis
- The Predictor Variable
- The Predictor Variable
- SAS for these
- SAS for Distribution of Peer Nominations
- SAS for Bully Scale
- Relationship Between the Measures
- We Know So Far That...
- Initial Modeling
- Fit of Poisson Regression Model
- Fit of Marginal Distribution

- Revised Model
- Fit Statistics & Parameter

Estimates

- Generalized Linear Models
- Fit of Negative Binomial

Model to Data

Fit of Marginal Distribution w/ Identity

- Overview

- Inference for model parameters

- Assessing Model Fit

- Global Fit Statistics

- Residuals

- Confidence intervals for fitted values

- Overdispersion

- Overview of Bully Example

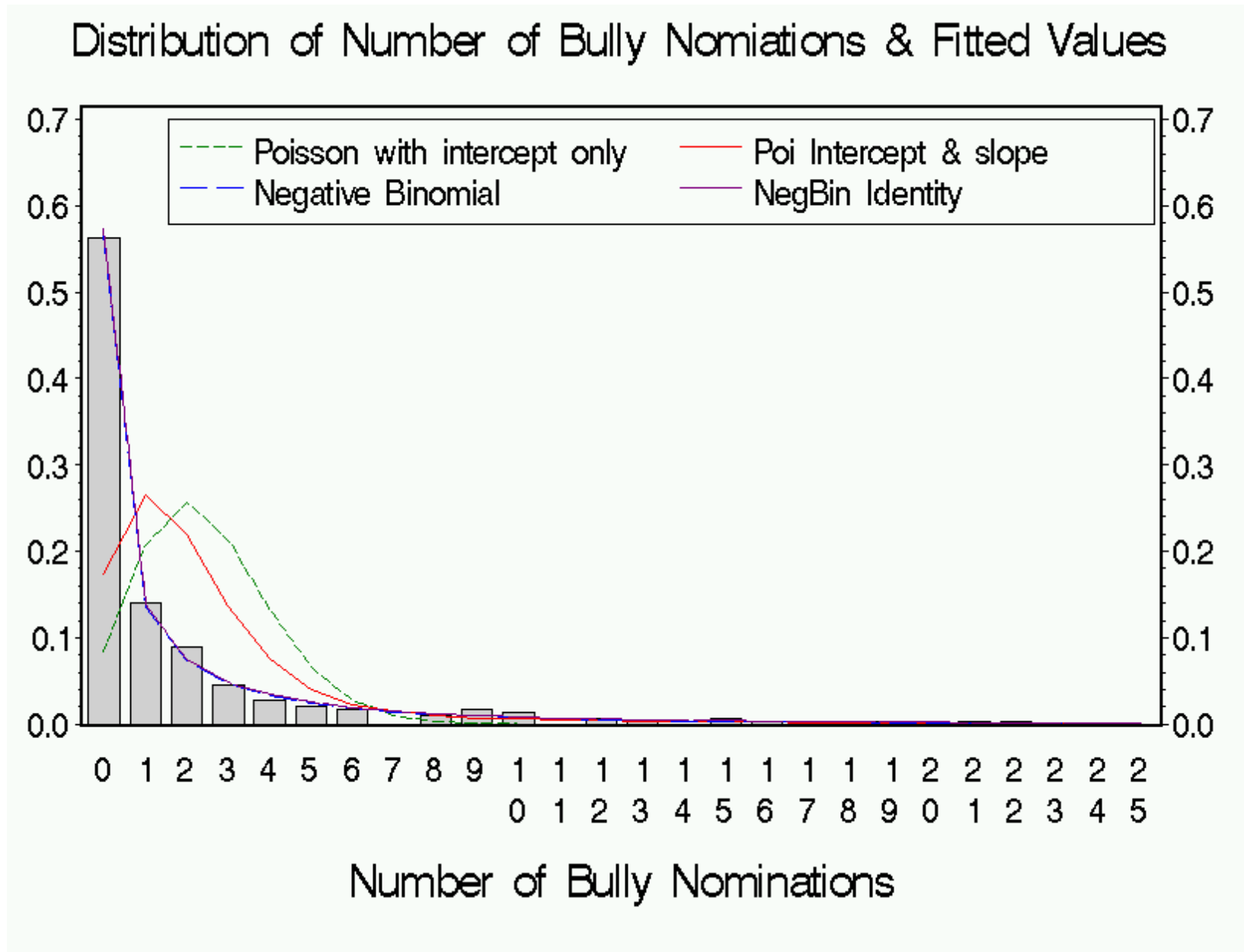
 - Outline of Bully Example
 - The Problem
 - Exploratory Analysis
 - The Predictor Variable
 - The Predictor Variable
 - SAS for these
 - SAS for Distribution of Peer Nominations
 - SAS for Bully Scale
 - Relationship Between the Measures
 - We Know So Far That...
 - Initial Modeling
 - Fit of Poisson Regression Model
 - Fit of Marginal Distribution

- Revised Model
- Fit Statistics & Parameter

Estimates

- Generalized Linear Models
- Fit of Negative Binomial

Model to Data



Conclusion

The bully scale can reasonably be used lieu of the peer nominations.

Support from this comes from

- The similarity of the marginal distributions for the two measures (both positively skewed).
- Goodness of fit of the negative binomial regression with identity link function.

Qualifications (i.e., more to be done):

- Add in other variables known to be related to bullying (e.g., gender) to try to account for extra variability (i.e., systematic vs random).
- More modeling that takes into account peer groupings (i.e., see whether there are “errors” or systematic differences between peer groups).

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

- Outline of Bully Example
- The Problem
- Exploratory Analysis
- The Predictor Variable
- The Predictor Variable
- SAS for these
- SAS for Distribution of Peer Nominations
- SAS for Bully Scale
- Relationship Between the Measures
- We Know So Far That...
- Initial Modeling
- Fit of Poisson Regression Model
- Fit of Marginal Distribution

- Revised Model
- Fit Statistics & Parameter

- Estimates
- Generalized Linear Models
- Fit of Negative Binomial

Model to Data

Zero Inflated Models

Models for situations where there might be two underlying types or groups: one group that follows the regression model and the other that just gives 0's.

Recommended supplemental reading:

Long, J.S. (1997). *Regression Models for Categorical and Limited Dependent Variables*.

Donald Erdman, Laura Jackson, Arthur Sinko (2008). *Zero-Inflated Poisson and Zero-Inflated Negative Binomial Models Using the COUNTREG Procedure* (Paper 322-2008). SAS Institute Inc., Cary, NC.
→ [PROC COUNTREG is in SAS v9.2](#)

Skrondal, A., & Rabe-Hesketh, S. (2004). *Generalized Latent Variable Modeling: Multilevel, Longitudinal and Structural Equation Models*. NY: Chapman & Hall

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

- Zero Inflated Models
- Basic Zero Inflated Model (e.g., "ZIP")
- ZIP Model (continued)
- ZIP and Bully Nominations
- Extending the ZIP
- ZIP Model Parameter Estimates
- ZIP w/ logit model and Bully Nominations
- Comparing all Fitted
- Comparing all Fitted
- Mix and Match

SAS for these Models

Fitting GLMS
Generalized Linear Models

Statistic inference & the

Basic Zero Inflated Model (e.g., “ZIP”)

- The basic model is essentially a latent class type model of the form

$$P(Y_{ij} = y|x_{ij}) = \begin{cases} \pi + (1 - \pi)P(0|x_{ij}) & \text{for } y = 0 \\ (1 - \pi)P(y|x_{ij}) & \text{for } y > 0 \end{cases}$$

where

- ◆ π = the probability of being in the “zero only” type or class.
- ◆ $P(0|x_{ij})$ and $P(y|x_{ij})$ are based on some model, such as Poisson or Negative Binomial regression.

- ZIP model is a zero inflated Poisson usually with a log link:

$$P(Y_{ij} = y|x_{ij}) = \begin{cases} \pi + (1 - \pi) \exp(-\mu_{ij}) & \text{for } y = 0 \\ (1 - \pi) \frac{\exp(-\mu_{ij}) \mu_{ij}^y}{y!} & \text{for } y > 0 \end{cases}$$

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

- Zero Inflated Models
- Basic Zero Inflated Model (e.g., “ZIP”)

- ZIP Model (continued)
- ZIP and Bully Nominations
- Extending the ZIP
- ZIP Model Parameter

Estimates

- ZIP w/ logit model and Bully Nominations
- Comparing all Fitted
- Comparing all Fitted
- Mix and Match

SAS for these Models

Fitting GLMS
Generalized Linear Models

Statistic inference & the

ZIP Model (continued)

$$P(Y_{ij} = y | x_{ij}) = \begin{cases} \pi + (1 - \pi) \exp(-\mu_{ij}) & \text{for } y = 0 \\ (1 - \pi) \frac{\exp(-\mu_{ij}) \mu_{ij}^y}{y!} & \text{for } y > 0 \end{cases}$$

■ Mean

$$E(Y_{ij} | x_{ij}) = (0 \times \pi) + \mu_{ij} \times (1 - \pi) = \mu_{ij} - \mu_{ij}\pi$$

■ Variance:

$$\text{var}(Y_{ij} | x_{ij}) = \mu_{ij}(1 - \pi)(1 + \mu_{ij}\pi)$$

- Note that if $\pi = 0$, we simply have a standard Poisson regression with log link.
- Extending the ZIP model by noting that class membership is dichotomous, so we can do a logistic regression (or other model for binary data) on the probability of class membership, e.g., a logit model,

$$\log\left(\frac{\pi_{ij}}{1 - \pi_{ij}}\right) = \gamma_0 + \gamma_1 z_{1ij} + \dots + \gamma_q z_{qij}$$

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

- Zero Inflated Models
- Basic Zero Inflated Model (e.g., "ZIP")
- ZIP Model (continued)
- ZIP and Bully Nominations
- Extending the ZIP
- ZIP Model Parameter Estimates
- ZIP w/ logit model and Bully Nominations
- Comparing all Fitted
- Comparing all Fitted
- Mix and Match

SAS for these Models

Fitting GLMS
Generalized Linear Models

Statistic inference & the

ZIP and Bully Nominations

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

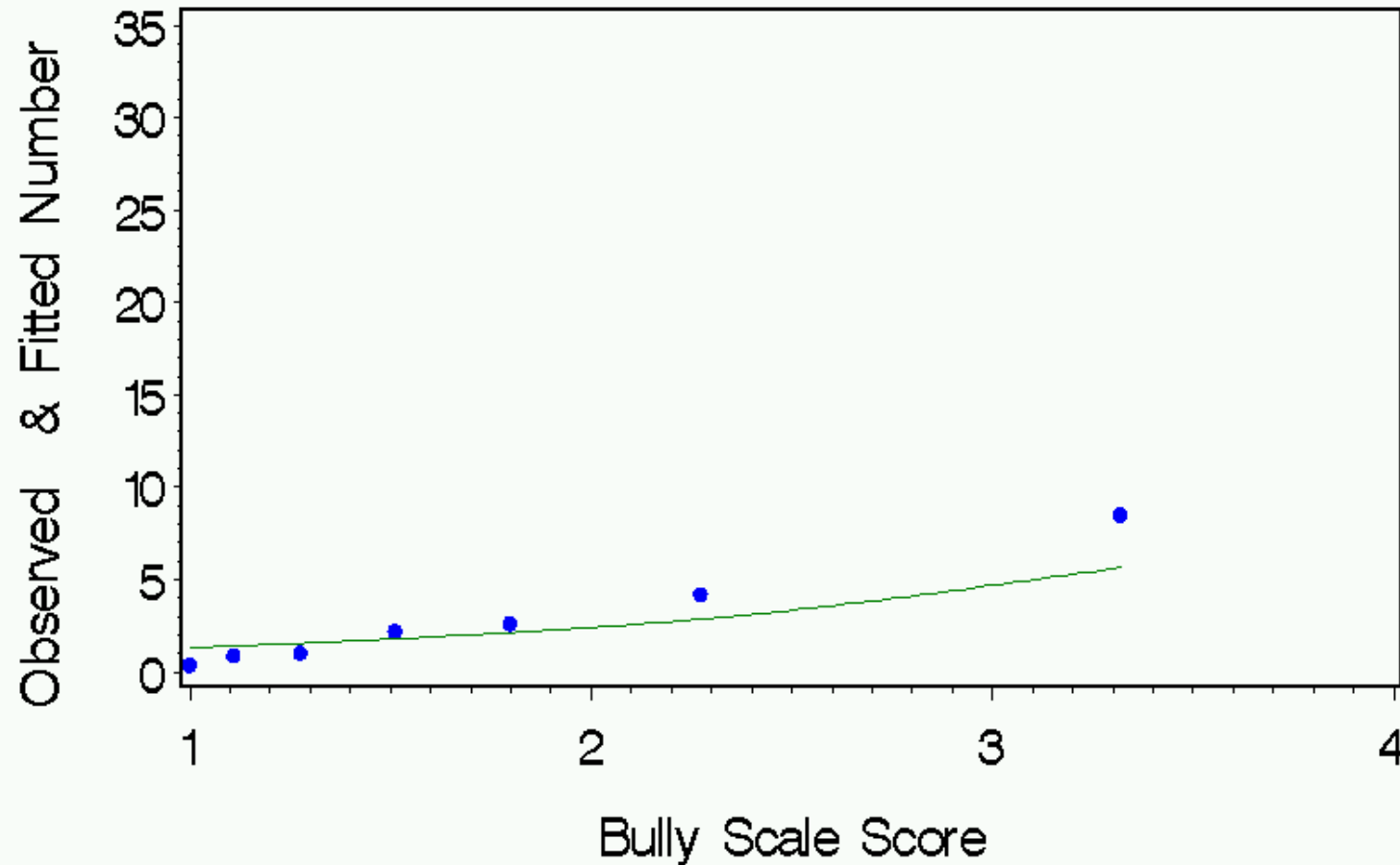
- Zero Inflated Models
- Basic Zero Inflated Model (e.g., "ZIP")
- ZIP Model (continued)
- ZIP and Bully Nominations
- Extending the ZIP
- ZIP Model Parameter Estimates
- ZIP w/ logit model and Bully Nominations
- Comparing all Fitted
- Comparing all Fitted
- Mix and Match

SAS for these Models

Fitting GLMS
Generalized Linear Models

Statistic inference & the

Observed & ZIP1 Fitted
for Number of Nominations as a Bully



Extending the ZIP

- Since class membership is dichotomous, so we can do a logistic regression (or other model for binary data) on the probability of class membership

- For example,

$$\log \left(\frac{\pi_i}{1 - \pi_i} \right) = \gamma_0 + \gamma_1 z_{1i} + \dots + \gamma_q z_{qi}$$

- For our Bully nominations, we could try

$$\log \left(\frac{\pi_{ij}}{1 - \pi_{ij}} \right) = \gamma_0 + \gamma_1 (\text{bully scale})_{ij}$$

- A comparison of how well various ZIP models fit the data:

Model								
Dist.	link	for π	df	G^2	X^2	AIC	BIC	
Poi	log	none	288	1579.04	814.82	1585	1596	
Poi	log	logit	287	1561.26	824.00	1569	1584	
Poi	Ident	logit	287	1553.76	825.90	1562	1576	

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

- Zero Inflated Models
- Basic Zero Inflated Model (e.g., "ZIP")
- ZIP Model (continued)
- ZIP and Bully Nominations
- Extending the ZIP
- ZIP Model Parameter Estimates
- ZIP w/ logit model and Bully Nominations
- Comparing all Fitted
- Comparing all Fitted
- Mix and Match

GAS for these Models

Fitting GLMS
Generalized Linear Models

Statistic inference & the

ZIP Model Parameter Estimates

and How to interpret them:

parm	ZIP w/o model for π				ZIP With Logit model for π			
	est	se	Wald	p	est	se	Wald	p
β_0	0.48	0.10	0.28	< .01	0.50	0.10	0.30	< .01
β_1	0.60	0.04	0.52	< .01	0.60	0.04	0.52	< .01
γ_0	0.21	0.12	-0.03	.09	1.50	0.35	0.82	< .01
γ_1					-0.77	0.20	-1.16	< .01

ZIP w/o model for π :

$$\exp(0.60) = 1.82 \quad \text{and} \quad \hat{\pi} = \frac{\exp(0.21)}{1 + \exp(0.21)} = .55$$

ZIP With Logit model for π :

$$\exp(0.60) = 1.82 \quad \text{and} \quad \hat{\pi} = \frac{\exp(1.50 - 0.77(\text{bullysc})_{ij})}{1 + \exp(1.50 - 0.77(\text{bullysc})_{ij})}$$

Note that $\exp(-.77) = 0.46$.

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

- Zero Inflated Models
- Basic Zero Inflated Model (e.g., "ZIP")

- ZIP Model (continued)
- ZIP and Bully Nominations
- Extending the ZIP
- ZIP Model Parameter

Estimates

- ZIP w/ logit model and Bully Nominations
- Comparing all Fitted
- Comparing all Fitted
- Mix and Match

SAS for these Models

Fitting GLMS
Generalized Linear Models

Statistic inference & the

ZIP w/ logit model and Bully Nominations

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

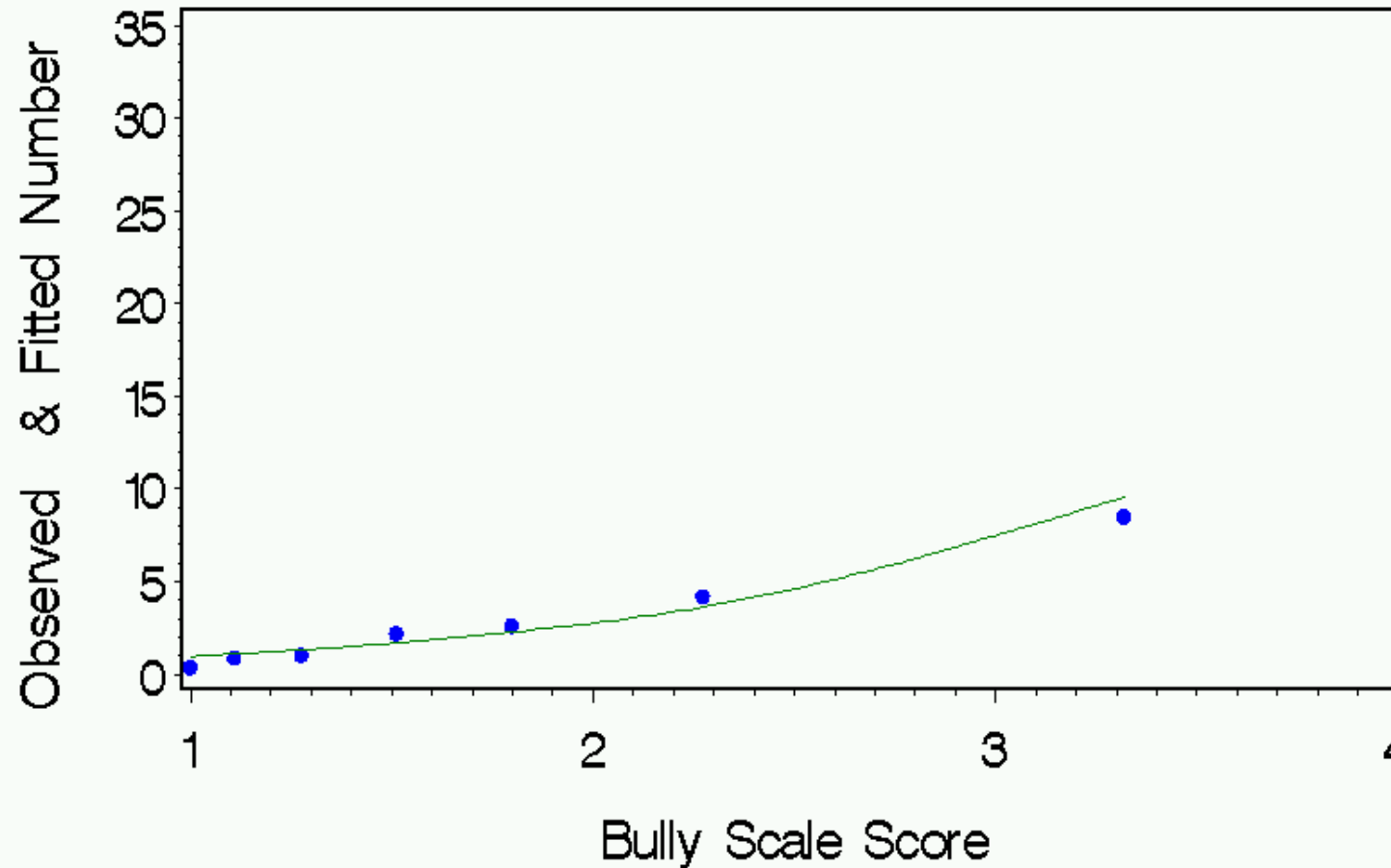
- Zero Inflated Models
- Basic Zero Inflated Model (e.g., "ZIP")
- ZIP Model (continued)
- ZIP and Bully Nominations
- Extending the ZIP
- ZIP Model Parameter Estimates
- ZIP w/ logit model and Bully Nominations
- Comparing all Fitted
- Comparing all Fitted
- Mix and Match

SAS for these Models

Fitting GLMS
Generalized Linear Models

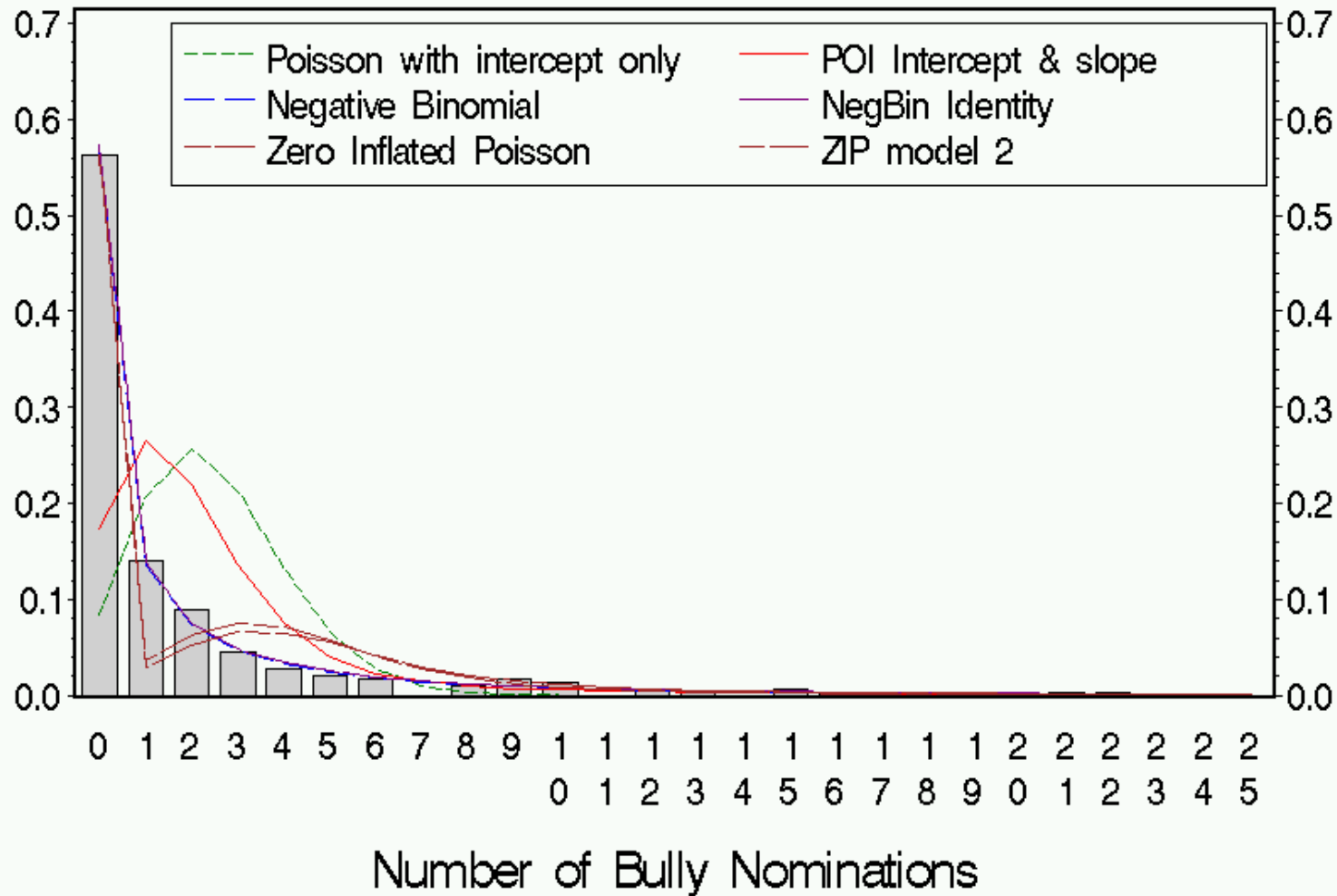
Statistic inference & the

Observed & ZIP2 Fitted
for Number of Nominations as a Bully



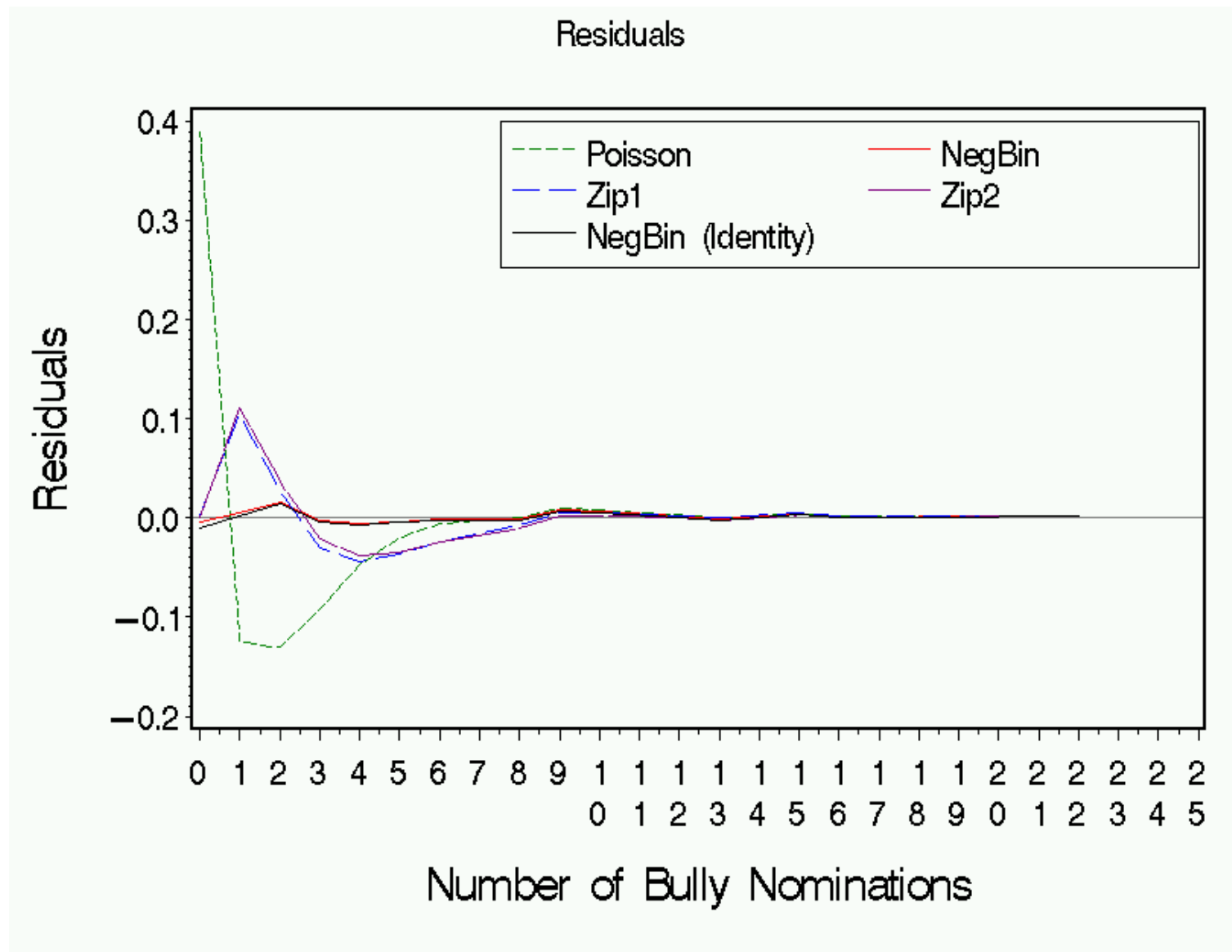
Comparing all Fitted

Distribution of Number of Bully Nominations & Fitted Values



- Overview
- Inference for model parameters
- Assessing Model Fit
- Global Fit Statistics
- Residuals
- Confidence intervals for fitted values
- Overdispersion
- Overview of Bully Example
- Zero Inflated Models
 - Zero Inflated Models
 - Basic Zero Inflated Model (e.g., "ZIP")
 - ZIP Model (continued)
 - ZIP and Bully Nominations
 - Extending the ZIP
 - ZIP Model Parameter Estimates
 - ZIP w/ logit model and Bully Nominations
 - Comparing all Fitted
 - Comparing all Fitted
 - Mix and Match
- GAS for these Models

Comparing all Fitted



Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

- Zero Inflated Models
- Basic Zero Inflated Model (e.g., "ZIP")
- ZIP Model (continued)
- ZIP and Bully Nominations
- Extending the ZIP
- ZIP Model Parameter Estimates
- ZIP w/ logit model and Bully Nominations
- Comparing all Fitted
- Comparing all Fitted
- Mix and Match

GAS for these Models

Fitting GLMS
Generalized Linear Models

Statistic inference & the

Mix and Match

- You can also have a zero inflated Negative Binomial model.
- You can specify a model other than logit for the mixing probability.
- You can do all this as a multi-level (random effects) model.
- In SAS:
 - ◆ If you use v 9.1, to fit a ZIP you have to use PROC NLMIXED or PROC GENMOD with programming statements.
 - ◆ If you use v 9.2, a ZIP can be fit easily using PROC GENMOD.
 - ◆ For a zero inflated Negative Binomial you can use PROC NLMIXED.
 - ◆ v 9.2, PROC COUNTREG in ETS; however, it doesn't appear to be in the version that I have. Documentation on it can be found at <http://support.sas.com/documentation/cdl/en/etsug/60372/HTML/default/countreg.toc.htm>

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

- Zero Inflated Models
- Basic Zero Inflated Model (e.g., "ZIP")
- ZIP Model (continued)
- ZIP and Bully Nominations
- Extending the ZIP
- ZIP Model Parameter Estimates
- ZIP w/ logit model and Bully Nominations
- Comparing all Fitted
- Comparing all Fitted
- Mix and Match

SAS for these Models

Fitting GLMS
Generalized Linear Models

Statistic inference & the

SAS for these Models

These work for v 9.1 and beyond:

```
/* Poisson Regression */  
proc genmod data=bullynom;  
  model bullynm = bullysc / link=log dist=poi type3;  
  output out=genmodpoi pred=fitpoi upper=uppoi  
         lower=lopoi stdreschi=res_poi;  
  title1 'Poisson Regression';
```

```
/* Negative Binomial Regression */  
proc genmod data=bullynom;  
  model bullynm = bullysc / link=log dist=negbin type3 ;  
  output out=genmodnb pred=nbfit  
         upper=nbup lower=nblo stdreschi=res_negbin;  
  title1 'Negative Binomial';
```

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted
values

Overdispersion

Overview of Bully Example

Zero Inflated Models

SAS for these Models

- SAS for these Models
- SAS for these Models
- SAS v 9.1 using NLMIXED

Fitting GLMS

Statistic inference & the
Likelihood function

“Deviance”

Summary

SAS for these Models

These work for v 9.2 and beyond:

```
/* Zero Inflated Poission Regression w/o model for inflation
   probability*/
proc genmod data=bullynom;
  model bullynm = bullysc / link=log dist=zip type3 obstats;
  zeromodel / link=logit;
  output out=zip1 pred=zipfit1 ;
  title1 'Zero Inflated Poisson';
```

```
/* Zero Inflated Poission Regression */
proc genmod data=bullynom;
  model bullynm = bullysc / link=log dist=zip type3;
  zeromodel bullysc / link=logit;
  output out=zip2 pred=zipfit2 ;
  title1 'ZIP w/ logit model for inflation probability';
```

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

SAS for these Models

- SAS for these Models
- SAS for these Models
- SAS v 9.1 using NLMIXED

Fitting GLMS

Statistic inference & the Likelihood function

“Deviance”

Summary

SAS v 9.1 using NLMIXED

```
proc nlmixed data=bullynom;
/* Some starting values */
parm beta0= -2.4076 beta1= 1.0168 a0=1;
/* linear predictor for the inflation probability */
    linpinf = a0 + a1*bullysc;
/* infprob = inflation probability for zeros * /
/* = logistic transform of the linear predictor*/
    infprob = 1/(1+exp(-linpinf));
/* Poisson mean */
    mu = exp( beta0 + beta1*bullysc);
/* Build the ZIP log likelihood */
if bullynm=0 then
    ll = log(infprob + (1-infprob)*exp(-mu));
else ll = log((1-infprob)) - mu + bullynm*log(mu)
    - lgamma(bullynm + 1);
model bullynm ~ general(ll);
title 'Zero Inflated Poisson regression';

SAS demo...
```

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

SAS for these Models

- SAS for these Models
- SAS for these Models
- SAS v 9.1 using NLMIXED

Fitting GLMS

Statistic inference & the Likelihood function

“Deviance”

Summary

Fitting GLMS

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

SAS for these Models

Fitting GLMS

- Fitting GLMS
- Newton-Raphson

Statistic inference & the Likelihood function

“Deviance”

Summary

Important for understanding inferential procedures.

Unlike normal (ordinary) regression, there is no “closed” form equation from which we can obtain estimates of model parameters. We much use some sort of iterative algorithm.

Two commonly used algorithms are

1. Fisher scoring
2. Newton-Raphson

For Binomial logistic regression and Poisson log-linear models, Fisher scoring simplifies to Newton-Raphson, so we’ll talk a little about Newton-Raphson.

Newton-Raphson

Newton-Raphson is an iterative algorithm

- It requires initial estimates (educated guesses) for the parameter estimates.
- On each cycle the estimates are “up-dated” by approximating the log-likelihood function by a simpler polynomial function that has the shape of a concave parabola.
- The cycles are repeated until the fitted values (parameter estimates) change less than some specified criterion (a very small number).

Newton-Raphson is sometimes referred to as “**iteratively reweighted least squares**”.

Because it is a type of weighted least squares. The weights change from cycle to cycle and depend on variability (which is not constant with means in Binomial & Poisson distributions).

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

SAS for these Models

Fitting GLMS

- Fitting GLMS
- Newton-Raphson

Statistic inference & the Likelihood function

“Deviance”

Summary

Statistic inference & the Likelihood function

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

SAS for these Models

Fitting GLMS

Statistic inference & the Likelihood function

● Statistic inference & the Likelihood function

● Wald Test

● Efficient Score Test

● Efficient Score Test

(continued)

● Likelihood Ratio

We can now consider each of the procedures we discussed for testing hypotheses about model parameters and see what information each of them uses (and therefore how they differ):

1. Wald tests.
2. Likelihood ratio tests.
3. Efficient score tests.

For purpose of illustration, consider the following log-likelihood function for the parameter β in a Poisson regression.

"Deviance"

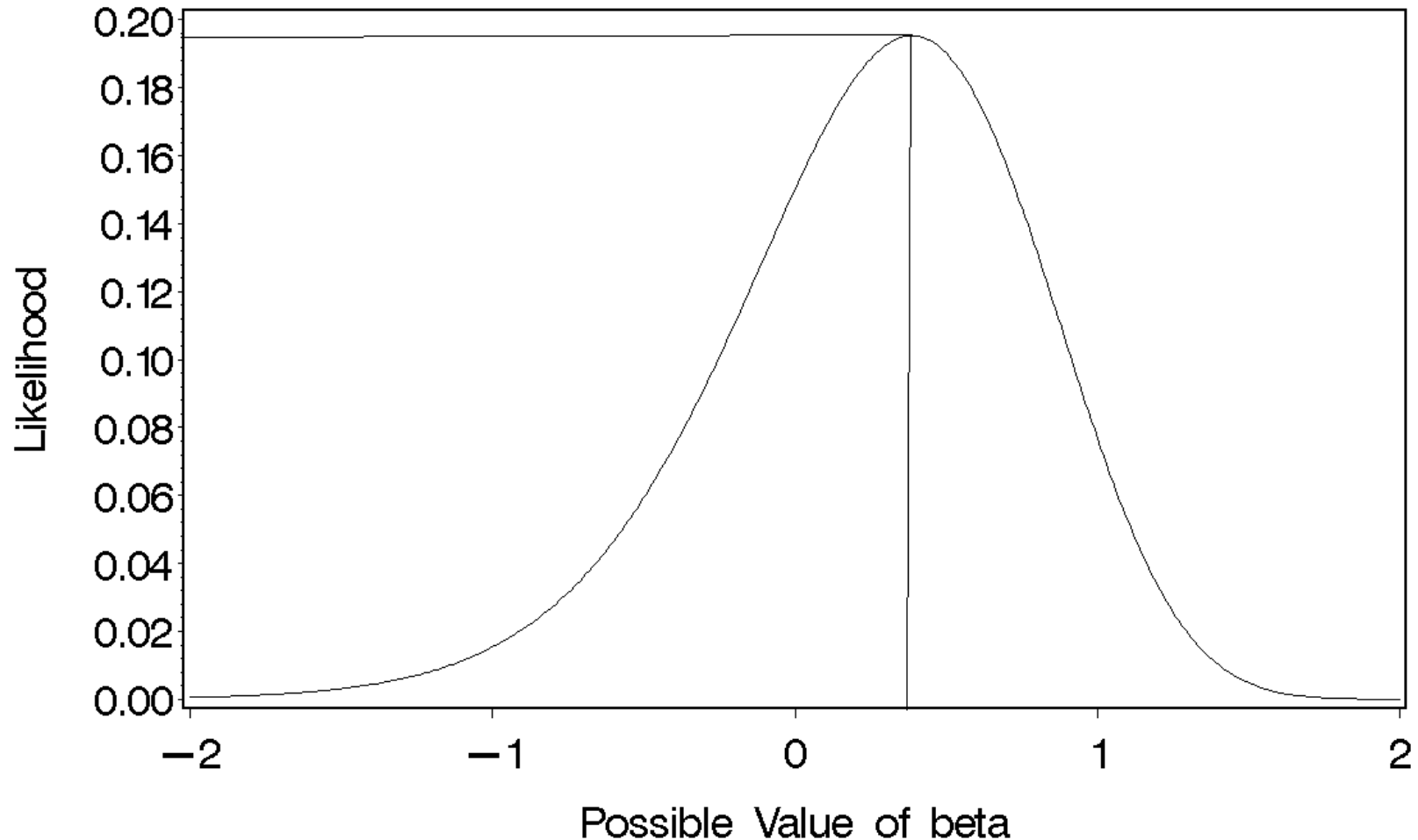
Generalized Linear Models

Summary

Wald Test

Only uses information about the log-likelihood function at the maximum likelihood estimated of $\hat{\beta}$.

Likelihood Function for Very Simple Case



Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

SAS for these Models

Fitting GLMS

Statistic inference & the Likelihood function

● Statistic inference & the Likelihood function

● Wald Test

● Efficient Score Test

● Efficient Score Test

(continued)

● Likelihood Ratio

"Deviance"

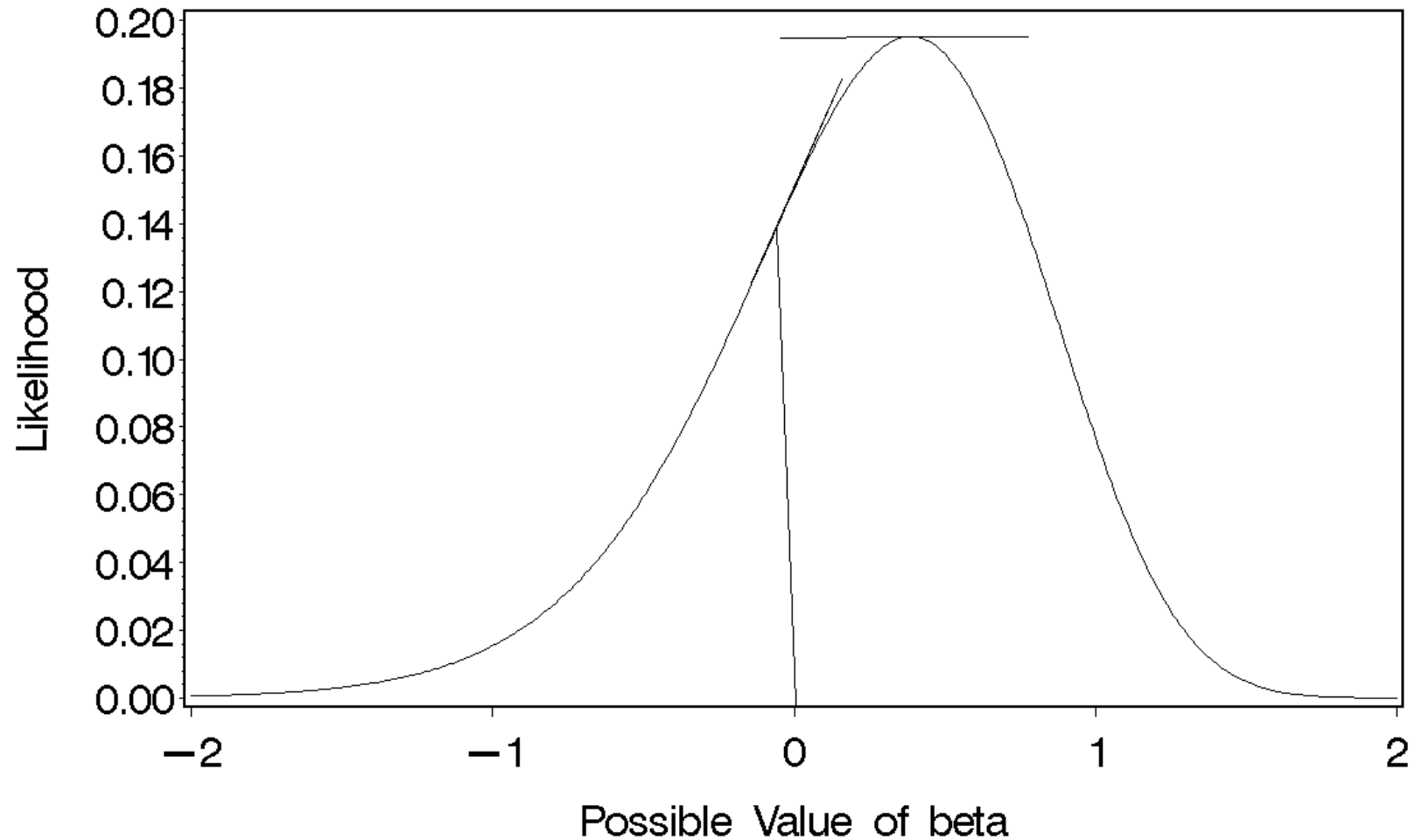
Generalized Linear Models

Summary

Efficient Score Test

uses information about the slope of the function at the null hypothesis value of $\beta = 0$.

Likelihood Function for Very Simple Case



Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

SAS for these Models

Fitting GLMS

Statistic inference & the Likelihood function

● Statistic inference & the Likelihood function

● Wald Test

● Efficient Score Test

● Efficient Score Test

(continued)

● Likelihood Ratio

"Deviance"

Generalized Linear Models

Summary

Efficient Score Test (continued)

- It compares the slope at $\beta = 0$ to the slope at the maximum likelihood estimate, which is 0 (i.e., the derivative of the function equals 0 at the MLE for β).
- The further the MLE of β is from zero, the greater the slope at $\beta = 0$ tends to be.
- The score statistic equals the square of the ratio of the slope (derivative) at $\beta = 0$ to its ASE
- The score statistic has an approximate chi-squared distribution with $df = 1$.
- You (usually) don't have to estimate the model to perform this test.
- Examples include CMH for conditional independence.
- This test can be performed even when $\hat{\beta}$ is infinite (whereas the Wald test cannot).

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

SAS for these Models

Fitting GLMS

Statistic inference & the Likelihood function

● Statistic inference & the Likelihood function

● Wald Test

● Efficient Score Test

● Efficient Score Test

(continued)

● Likelihood Ratio

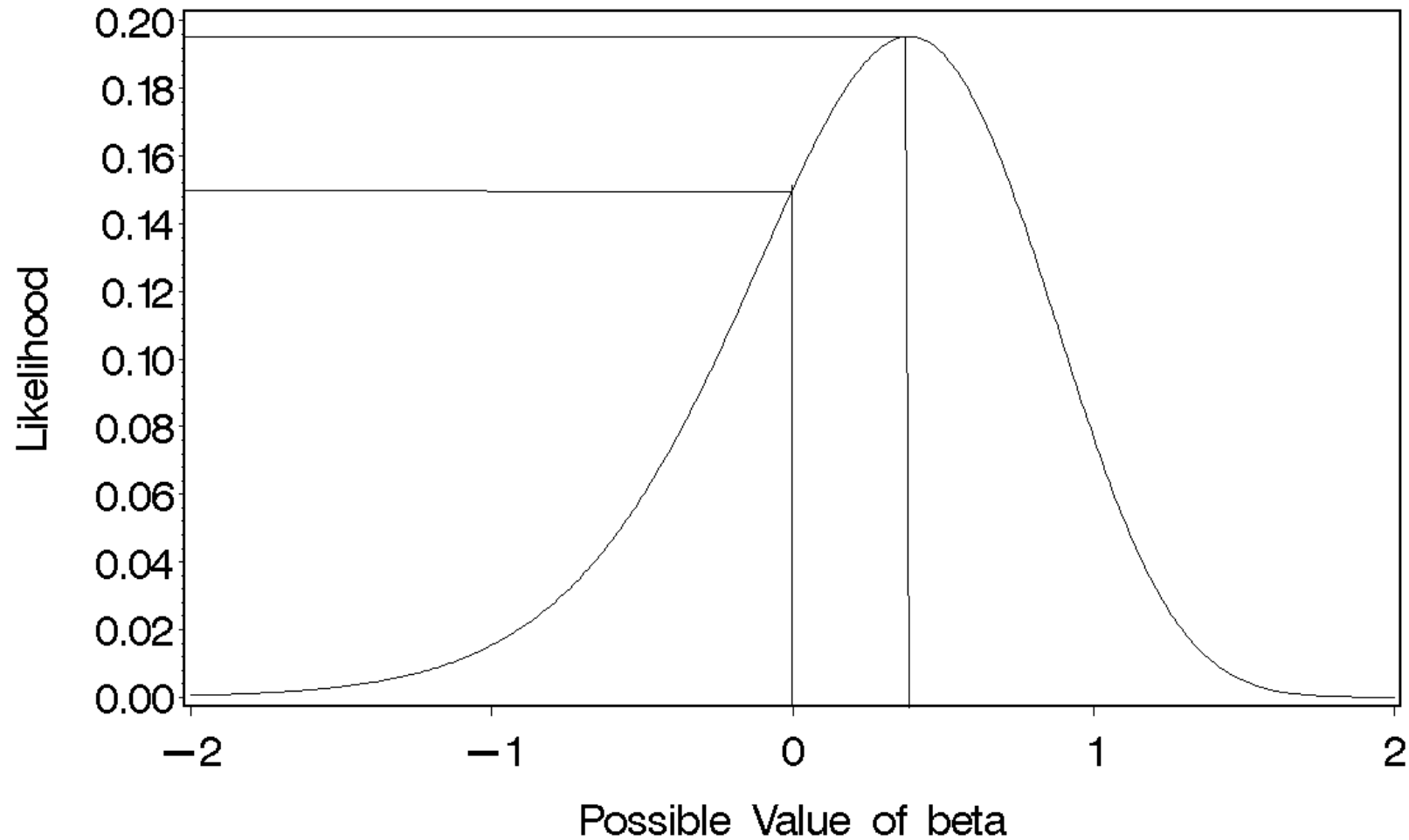
"Deviance"

Generalized Linear Models

Summary

Likelihood Ratio

Likelihood Function for Very Simple Case



Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

SAS for these Models

Fitting GLMS

Statistic inference & the Likelihood function

● Statistic inference & the Likelihood function

● Wald Test

● Efficient Score Test

● Efficient Score Test

(continued)

● Likelihood Ratio

"Deviance"

Generalized Linear Models

Summary

“Deviance”

(This is an analog to regression models decomposition of sum of squares.)

Let

L_S equal the maximum of the log-likelihood function of the most complex model; that is, the model with as many parameters as there are observations — the “**Saturated Model**”.

L_M equal the maximum of the log-likelihood function of a simpler model (of interest).

The deviance compares the log-likelihood value for the saturated and some simpler model.

$$\text{Deviance} = -2(L_M - L_S)$$

For Poisson loglinear (regression) models and Binomial logit models,

$$\text{Deviance} = G^2$$

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

SAS for these Models

Fitting GLMS

Statistic inference & the Likelihood function

“Deviance”

- “Deviance”
- Deviance
- Deviance & Model Comparison

Summary

Deviance

So, for lots of GLMs, Deviance has an approximate chi-squared distribution (. . . and for some GLMs, it doesn't — which is another good reason why programs that fit GLMs don't automatically print out p -values).

When Deviance has an approximate chi-squared distribution, the “residual” degrees of freedom equal

$$df = \# \text{ responses} - \# \text{ nonredundant model parameters}$$

Deviance Residuals components of Deviance (and for binomial logit and Poisson regression G^2). They are alternatives to Pearson residuals. These also can be adjusted so that they are approximately $\mathcal{N}(0, 1)$.

Deviance & Model Comparison — a very useful property.

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

SAS for these Models

Fitting GLMS

Statistic inference & the Likelihood function

“Deviance”

- “Deviance”
- Deviance
- Deviance & Model Comparison

Summary

Deviance & Model Comparison

- Suppose that we have two models, M_0 and M_1 , where M_0 is a special case of M_1 (i.e., they are “nested”).
- Assuming that the more complex model M_1 fits the data, the likelihood ratio test that the simpler model fits (i.e., that you don't need the terms that are in M_1 but are not in M_0) is



$$\begin{aligned} -2(L_0 - L_1) &= -2(L_O - L_S) - \{-2(L_1 - L_S)\} \\ &= \text{Deviance}_0 - \text{Deviance}_1 \\ &= G_O^2 = G_1^2 \end{aligned}$$

- with degrees of freedom

$$df = df_0 - df_1$$

where df_0 and df_1 are the residual df 's for models M_0 and M_1 , respectively. Equivalently, degrees of freedom equal the number of non-redundant parameters in M_1 that are not in M_0 .

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

SAS for these Models

Fitting GLMS

Statistic inference & the Likelihood function

“Deviance”

- “Deviance”
- Deviance
- Deviance & Model Comparison

Summary

Summary

GLM theory unifies important models for continuous and discrete responses.

Some common models as GLMs

Random Component	Link Function	Systematic Component	Model
Normal	Identity	Continuous	Regression
Normal	Identity	Categorical	ANOVA
Normal	Identity	Mixed	ANCOVA
Binomial	Logit	Mixed	Logistic regression
Poisson	Log	Mixed	Loglinear
Multinomial	Generalized Logit	Mixed	Multinomial response

The next section of the course will cover logistic regression, loglinear models, and multinomial response models in more detail.

Overview

Inference for model parameters

Assessing Model Fit

Global Fit Statistics

Residuals

Confidence intervals for fitted values

Overdispersion

Overview of Bully Example

Zero Inflated Models

SAS for these Models

Fitting GLMS

Statistic inference & the Likelihood function

"Deviance"

Summary

● Summary