Effect size measures for nonlinear count regression models

Stefany Coxe, Ph.D.

Florida International University, Department of Psychology

Effect Size

Effect size gives the magnitude of an effect or relationship
Often given in standardized units to aid interpretation

Important adjunct to tests of statistical significance

- How large is it? What does it mean in practical terms?

Standardized mean difference (SMD) is the difference between two group means, divided by standard deviation

- Cohen's d, Hedge's g, Glass' delta

Poisson Regression

Poisson regression is a regression model for count outcomes
Counts violate linear regression assumptions because they are:

- Discrete and strictly positive or 0
- Typically display **non-constant variance** in relationships

Non linear (exponential) relationship

$$\hat{Y} = e^{(b_0 + b_1 X_1 + b_2 X_2 + \cdots + b_p X_p)}$$

Relationships are **heteroscedastic** (non constant variance)
For Poisson regression, conditional variance = conditional mean

Rate ratio: typical measure of effect size for Poisson regression
When X increases 1 unit, $\hat{Y}$ is multiplied by e^{b_1}

Alternative count models **relax** the mean-variance relationship and so can account for overdispersion:

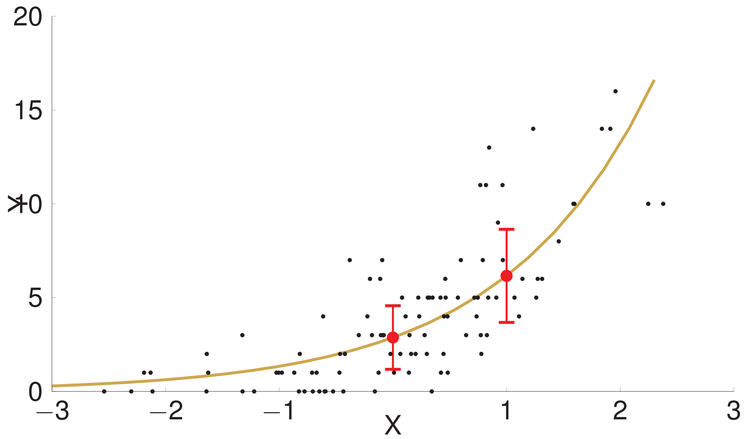
- **Overdispersed Poisson regression** allows a scaling adjustment in the relationship, so $Variance = mean \times scale$
- **Negative binomial regression** models the relationship such that $Variance = mean + alpha \times mean^2$

Purpose

Extend the **standardized mean difference** measure of effect size for use with count regression models:

- Allow for continuous OR categorical predictors
- Take into account heteroscedasticity of count outcomes
- Correctly account for overdispersion in overdispersed Poisson and negative binomial models
- Estimate confidence intervals for effect size

Poisson Regression Model



- Relationship between X and Y for simulated count outcome Y
- Yellow line is the predicted line for a Poisson regression
- Red dots are predicted means when $X = 0$ and $X = 1$
- Error bars are $\pm$ 1 SD
- Note **nonlinear relationship** and **nonconstant variance**

Implementation

Programmed in R using the "shiny" package
`https://stefany.shinyapps.io/RcountD/`

Predicted (model-based) means at $X = 0$ and $X = 1$ are used

- Group predictors (0/1): mean difference between groups
- Continuous predictors: effect as X changes from 0 to 1

Variance at $X = 0$ is used to standardize

- Group predictors (0/1): "control group" variance is used
- Continuous predictors: variance at $X = 0$ is used

Confidence intervals are estimated using **Monte Carlo simulation**

- General method that can be applied to any function

Mean center or standardize predictor to improve interpretation

Inputs:

- Type of regression model: Poisson, overdispersed Poisson, negative binomial
- Intercept coefficient and standard error
- Slope coefficient and standard error (for effect of interest)
- Dispersion parameter (for OD Poisson and NB only)
- Confidence interval percentage (default: 95%)
- Number of Monte Carlo replications (default: 2000)
- Random number seed (to re-create results; default: 12345)

Outputs:

- Predicted mean and SD when X = 0 and X = 1
- Rate ratio and confidence interval
- Standardized mean difference and confidence interval
- Histogram of Monte Carlo distribution for each effect size measure, with mean and confidence limits indicated

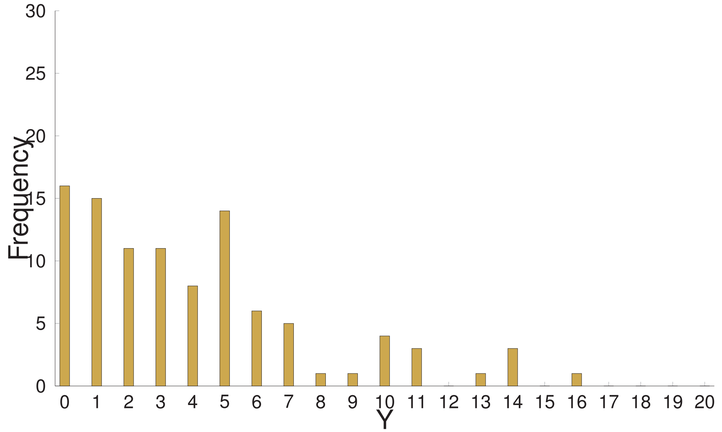
Effect size measures for nonlinear count regression models

Stefany Coxe, Ph.D.

Florida International University, Department of Psychology

Example 1: Poisson regression

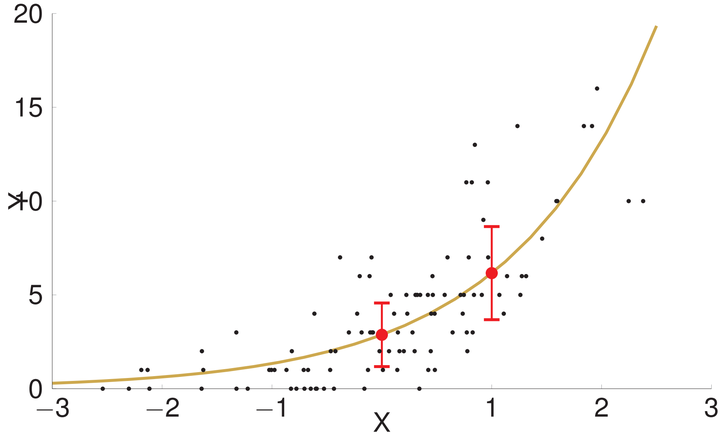
Distribution of the count outcome: **Variance ≈ mean**



Relationship between X and Y: **Exponential**

$\hat{Y} = e^{1.05558 + 0.76274X}$

Standard errors are 0.06577 and 0.05546, respectively



Example 1 output

- When predictor = 0, the predicted outcome mean is 2.874 with a standard deviation of 1.695
- When predictor = 1, the predicted outcome mean is 6.161 with a standard deviation of 2.482

Exponential effect size

- The mean when predictor = 1 is **2.144 times larger** than the mean when predictor = 0.
- The 95 % CI for this estimate is [1.917 , 2.385]

Standardized mean difference effect size (Cohen's d)

- The mean when predictor = 1 is **1.94 standard deviations higher** than the mean when predictor = 0.
- The 95 % CI for this estimate is [1.455 , 2.502].

Example 2 output

- When predictor = 0, the predicted outcome mean is 2.8 with a standard deviation of 2.482
- When predictor = 1, the predicted outcome mean is 7.077 with a standard deviation of 5.343

Exponential effect size

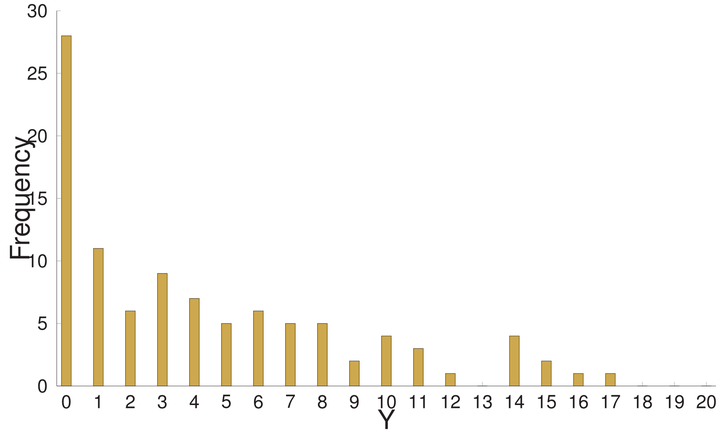
- The mean when predictor = 1 is **2.528 times larger** than the mean when predictor = 0.
- The 95 % CI for this estimate is [2.063 , 3.068].

Standardized mean difference effect size (Cohen's d)

- The mean when predictor = 1 is **1.724 standard deviations higher** than the mean when predictor = 0.
- The 95 % CI for this estimate is [1.144 , 2.43].

Example 2: Negative binomial regression

Distribution of the count outcome: **Variance > mean**

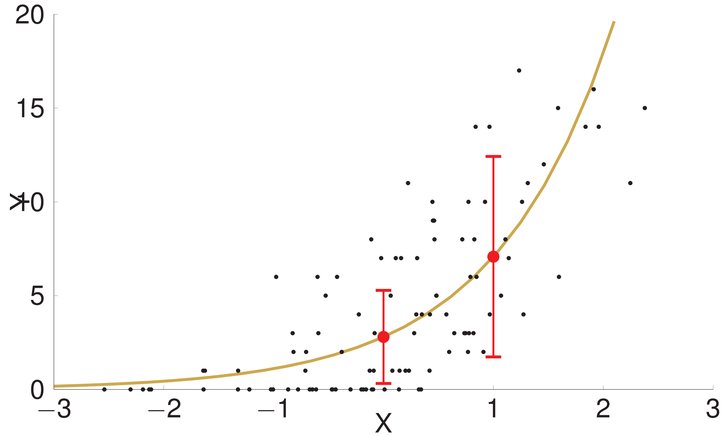


Relationship between X and Y: **Exponential**

$\hat{Y} = e^{1.02947 + 0.92741X}$

Standard errors are 0.09815 and 0.10074, respectively

Alpha = 0.42863 ($Variance = mean + alpha \times mean^2$)



Effect size measures for nonlinear count regression models

Stefany Coxe, Ph.D.

Florida International University, Department of Psychology

Effect size measures for Poisson regression models

Submit

Regression model

Poisson regression

Intercept/constant regression coefficient

1.05558

Standard error of intercept

0.06577

Regression coefficient for effect of interest

0.76274

Standard error of effect of interest

0.05546

Dispersion

0

For overdispersed Poisson regression, dispersion is square root of phi: (Scale) parameter from SPSS or square of Scale parameter from SAS

For negative binomial regression, dispersion is alpha: (Negative binomial) parameter from SPSS or Dispersion from SAS

Confidence interval %

95% CI

Replications

2000

Seed

12345

Monte Carlo methods use random number generation. To obtain the exact same results each time you use this app, use the same seed value each time

Submit

Means and standard deviations

When predictor = 0, the predicted outcome mean is 2.874 with a standard deviation of 1.695 .

When predictor = 1, the predicted outcome mean is 6.161 with a standard deviation of 2.482 .

Exponential effect size

Effect size = 1 indicates no effect. Values greater than 1 indicate that the outcome mean when predictor = 1 is HIGHER than the outcome mean when predictor = 0 while values less than 1 indicate that the outcome mean when predictor = 1 is LOWER than the outcome mean when predictor = 0.

The mean when predictor = 1 is 2.144 times larger than the mean when predictor = 0.

The 95 % CI for this estimate is [1.917 , 2.385].

Standardized mean difference effect size (Cohen's d)

Effect size = 0 indicates no effect. Positive values indicate that the outcome mean when predictor = 1 is higher than the outcome mean when predictor = 0, while negative values indicate that the outcome mean when predictor = 1 is lower than the outcome mean when predictor = 0.

The mean when predictor = 1 is 1.94 standard deviations higher than the mean when predictor = 0.

The 95 % CI for this estimate is [1.455 , 2.502].