2016

Regression Models for Count Data Based on the Double Poisson Distribution

Rebecca Wardrop
University of South Carolina

Regression Models for Count Data Based on the Double Poisson Distribution

by

Rebecca Wardrop

Bachelor of Arts
College of Wooster 2014

Submitted in Partial Fulfillment of the Requirements

for the Degree of Master of Science in Public Health in

Biostatistics

The Norman J. Arnold School of Public Health

University of South Carolina

2016

Accepted by:

James W. Hardin, Director of Thesis

James R. Hussey, Reader

Robert R. Moran, Reader

Paul Allen Miller, Vice Provost and Interim Dean of Graduate Studies

Abstract

This paper explores the double Poisson distribution. The probability mass function and the difficulties associated with derivative-based optimization for this distribution are discussed. Stata software developed for estimation of double Poisson regression is detailed. Simulations are used to test the software. Data which are over-, under-, and equidispersed relative to the Poisson are generated and the software is utilized to estimate a regression model, a zero-inflated model, and a marginalized zero-inflated model all based on the double Poisson distribution. The estimated power of the test for $\phi = 1$ for the double Poisson models are compared to the power of the test for $\alpha = 0$ for the negative binomial models. Coefficient estimation is also compared across the models.

Table of Contents

List of Tables

Chapter 1

Introduction

Count data are most commonly modeled using regression based on the Poisson distribution. Despite its popularity, this type of regression has a considerable limitation. The Poisson distribution has a single parameter, which does not allow for the variance to vary independently from the mean. This is a significant drawback because count data are often over- or underdispersed relative to Poisson variance. Multiple alternative approaches to address this issue have been popularized, including negative binomial based regression, zero-inflated regression models, and generalized Poisson regression. However, these methods are predominantly used for overdispersed data, and often fall short in modeling data which are underdispersed.

Regression based on the double Poisson distribution proposed by Efron (1986) is an alternative which allows for more accurate regression models for data that are either over- or underdispersed relative to the Poisson. This paper will explore double Poisson regression and zero-inflated extensions of the regression model. The Stata software developed for this paper will be described and utilized in analyzing simulated data. In addition, double Poisson models will be compared to the traditional Poisson approach as well as to negative binomial models.

Chapter 2

The Double Poisson Regression and Extensions

Double exponential families were proposed by Bradley Efron (1986) as a solution to over- and underdispersion. These families extend the distributions of the exponential family to incorporate dispersion parameters, which allow the variance to differ from the mean while still preserving single parameter exponential family properties. Efron (1986) defined the probability mass function for any double exponential family distribution as

$$f_{\mu,\phi,n}(y) = c(\mu,\phi,n)\phi^{\frac{1}{2}} g_{\mu,n}(y)^{\phi} g_{\mu,n}(y)^{1-\phi} dG_n(y) \tag{2.1}$$

Thus, by the definition of a double exponential family, for a variable Y following the double poisson distribution, the exact probability mass function is given by

$$P(Y=y) = f_{\mu,\phi}(y) = c(\mu,\phi)\left(\phi^{\frac{1}{2}} e^{-\phi\mu}\right)\left(\frac{e^{-y}y^{y}}{y!}\right)\left(\frac{e\mu}{y}\right)^{\phi y}, y = 0,1,2,... \tag{2.2}$$

where $c(\mu,\phi)$ is the normalizing constant which ensures the probability mass function sums to unity. The equation for this constant is given as

$$\frac{1}{c(\mu,\phi)} = \sum_{y=0}^{\infty} f_{\mu,\phi}(y) \approx 1 + \frac{1-\phi}{12\mu\phi}\left(1+\frac{1}{\mu\phi}\right) \tag{2.3}$$

Note that the normalizing constant is an approximation of an infinite series. This approximation does ensure that the probability mass function sums to one in most cases; however, when μ is small, the approximation becomes unreliable, often leading to a sum greater than one (Zou, Geedipally, & Lord, 2013). To provide higher accuracy, a k-th partial sum can be used, where k is equal to ten times the largest observed outcome.

If $Y = Y_1, Y_2, ..., Y_n$ represent an independent, identically distributed sample from a double Poisson distribution with parameters μ and ϕ, the joint probability mass function is:

$$f_{\mu,\phi}(y) = \prod_{i=0}^{n} c(\mu, \phi) \left(\phi^{\frac{1}{2}} e^{-\phi\mu}\right) \left(\frac{e^{-y_i} y_i^{y_i}}{y_i!}\right) \left(\frac{e\mu}{y_i}\right)^{\phi y_i} \tag{2.4}$$

Therefore, the log-likelihood function is given by

$$\begin{aligned} \mathcal{L}(\mu, \phi|Y) \quad = \quad & \sum_{i=0}^{n} \left\{ \frac{1}{2} \ln(\phi) - \phi\mu - y_i + y_i \ln y_i - \ln \Gamma(y_i + 1) \right. \\ & \left. +\phi y_i(\ln \mu - \ln y_i + 1) - \ln(c(\mu, \phi)) \right\} \end{aligned} \tag{2.5}$$

Given the above function, the mean is $E(Y) \approx \mu = g(\mathbf{x}\boldsymbol{\beta})$. The natural link function used is $\mu = \exp(\mathbf{x}\boldsymbol{\beta})$ because μ must be positive. Thus, the coefficients in a double Poisson model are interpreted the same way as in Poisson regression.

Further, the variance is given by $Var(Y) \approx \frac{\mu}{\phi}$, where $\phi > 0$, allowing us to characterize dispersion relative to the Poisson. Thus, when ϕ is greater than one, the variance is less than the mean, implying that the data are underdispersed. Conversely, when ϕ is less than one, the variance is greater than the mean and the data are overdispersed. When ϕ is equal to one, the distribution reduces to the Poisson, signifying that the data are equidispersed.

The constant, $c(\mu, \phi)$ is often suppressed in maximum likelihood estimation due to issues with nonlinearity and difficulties calculating derivatives of an infinite series (Chou and Steenhard, 2009). However, in the approach considered in this paper, numeric derivatives are employed to avoid these issues.

2.1 Zero-Inflated Models

As has been considered for Poisson regression, a mixture model can be considered to address the issue of excess zeroes. In this mixture the outcome count variable, Y_i, can take a zero value with probability π, or it can draw an outcome from the double Poisson distribution with probability $1 - \pi$.

The zero excess is modeled by

$$f(y) = \pi + (1 - \pi) f_z(y) \tag{2.6}$$

where we will consider $f_z(y)$ to be the double Poisson probability mass function for the non-zero-excess group and π is the mixing proportion. Thus,

$$P(Y = y) = \begin{cases} \pi + (1 - \pi) c(\mu, \phi) \left(\phi^{\frac{1}{2}} e^{-\phi\mu}\right) & \text{if } y = 0, \\ 1 - \pi) c(\mu, \phi) \left(\phi^{\frac{1}{2}} e^{-\phi\mu}\right) \left(\frac{e^{-y} y^y}{y!}\right) \left(\frac{e\mu}{y}\right)^{\phi y} & \text{if } y = 1, 2, ... \end{cases} \tag{2.7}$$

There are two link functions used in zero-inflated regression. The logit link with the Bernoulli process leads to logistic regression, for which $\pi = \frac{\exp(\mathbf{x}\boldsymbol{\beta})}{1+\exp(\mathbf{x}\boldsymbol{\beta})}$, models the zero excess, while double Poisson regression using the log link models the remaining outcomes with $\mu = \exp(\mathbf{z}\boldsymbol{\gamma})$.

Using this model yields two set of regression coefficients. One set, $\boldsymbol{\beta}$, describes the association of the variables considered with the probability of being in the excess zero group. The other set, $\boldsymbol{\gamma}$, can be interpreted as standard double Poisson regression coefficients. The drawback of this approach is that the two sets of coefficients apply only to the group to which the observation belongs. No inferences can be made based on these coefficients to describe the effect of the explanatory variables in the overall population.

2.2 Zero-Inflated Marginalized Models

Marginalized zero-inflated regression is a simple adaptation of the zero-inflated model to make the resulting coefficients applicable to the overall population. Zero-inflated models yield parameter estimates which describe the association between covariates and a latent outcome. Marginalized models yield parameter estimates which describe the effects of the variables of interest on the marginal mean $E(Y|X)$ (Martin & Hall, 2016).

A simple change is made to the link function to achieve the marginalized model. According to Long et al (2014), the overall population mean μ is given by

$$\mu = \exp\{\mathbf{x}\boldsymbol{\alpha} - \ln(1-\pi)\} \tag{2.8}$$

where $\boldsymbol{\alpha}$ are the coefficients relating the covariate to outcomes from the non-zero excess group. The mixture probability is still modeled by $\pi = \frac{\exp(\mathbf{x}\boldsymbol{\beta})}{1+\exp(\mathbf{x}\boldsymbol{\beta})}$.

Similar to the zero-inflated double Poisson model, this modeling process yields two sets of regression coefficients. The coefficients can be interpreted in the same manner as in the zero-inflated model. The set of coefficients, $\boldsymbol{\beta}$, still describes the association of the covariates with the mixture probability and $\boldsymbol{\alpha}$ are still interpreted as standard double Poisson regression coefficeints. However, the set of coefficients $\boldsymbol{\alpha}$ can be used to make inference about the overall population due to the marginalization.

Chapter 3

Stata Software

The Stata software developed for double Poisson models mimics the syntax for other types of regression. The syntax is given by:

dpoisson *depvar* [*indepvars*] [*if*] [*in*] [*weight*] [*,options*]

where **dpoisson** is the regression command and the dependent variable is the only necessary specification. The output of the software and input specifications are also similar to other regression commands. It includes a model summary containing the log likelihood of the fitted full model and a likelihood ratio test that at least one of the covariates has a non-zero coefficient. The software also outputs a table of parameter estimates including the estimated coefficients and their standard errors, z-test statistics, p-values for the z-test, and 95% confidence intervals.

The first step in estimating regression parameters is to identify valid starting values for model. The software identifies the starting value by analytically obtaining the solution for a constant only model such that $E(y) = \exp(\beta_0)$. Given that β_0 is valid for the constant only model, the vector of coefficients $\widehat{\boldsymbol{\theta}} = (\beta_0, 0, 0, ..., 0)$ must also be valid for the model containing the considered predictors. The log-likelihood for this model is stored and Stata can then optimize and solve for $\widehat{\boldsymbol{\theta}}_{final}$ using the same procedure used for other distributions.

In addition to estimating a double Poisson regression, the software also estimates a traditional Poisson regression using the variables entered. The optimized log-likelihood for the Poisson model is stored and compared to the optimized double Poisson log-likelihood in a likelihood ratio test of whether $\phi = 1$, in which case the model does not differ from a Poisson.

The commands for zero-inflated and zero-inflated marginalized double Poisson regression run similarly to the double Poisson command while incorporating the mixture. These commands also mirror the input and output of the commonly used zero-inflated commands built into Stata.

zidp *depvar* [*indepvars*] [*if*] [*in*] [*weight*] , **inflate**(*varlist*[,**offset**(*varname*)] |**_cons**) [*options*]

zimdp *depvar* [*indepvars*] [*if*] [*in*] [*weight*] , **inflate**(*varlist*[,**offset**(*varname*)] |**_cons**) [*options*]

where the **zidp** and **zimdp** are the regression commands for zero-inflated and zero-inflated marginalized, respectively. The dependent variable and the inflate variable list are the only necessary specifications.

Chapter 4

Simulation Approach

To evaluate the double Poisson model, a simulation study was conducted. First, the size of the likelihood ratio test comparing the model to the Poisson model was checked. Data were simulated using the Poisson distribution such that $E(y) = \exp(0.25 + \ln(1.5)x_1 - 0.25x_2)$, where $x_1 \sim$ Bernoulli(0.5) and $x_2 \sim$ Uniform(0, 1). Then the outcome variable, y, was modeled using double Poisson regression with covariates x_1 and x_2. The simulation was run with 1000 repetitions for sample sizes of 30, 100, and 1000 observations. To evaluate whether the size of the test is appropriate, the number of times the test statistic exceeds $\chi^2_{1,0.95}$ was counted. This count is expected to be approximately 5% of the recorded test statistic values.

Then, data were simulated from a double Poisson distribution for which $E(y) = \exp(0.25 + \ln(1.5)x_1 - 0.25x_2)$, where $x_1 \sim$ Bernoulli(0.5) and $x_2 \sim$ Uniform(0, 1), and $\phi = 0.5$, indicating overdispersion relative to the Poisson distribution. These simulated data were modeled using double Poisson regression, negative binomial regression, and Poisson regression. Again the simulation was run with 1000 repetitions for sample sizes of 30, 100, and 1000 observations. For the double Poisson and negative binomial models, the number of times the likelihood ratio test statistic exceeded $\chi^2_{1,0.95}$ was counted, yielding an estimate of the power. Further, the number of times the 95% confidence intervals for the coefficients contain the true value is counted for all three regression models. This process was repeated using data simulated from double Poisson distribution for which $\phi = 1$ and $\phi = 1.5$, indicating equi- and under-dispersion relative to the Poisson, respectively.

A similar approach was used to compare the zero-inflated models. Data were again simulated from a double Poisson distribution for which $E(y) = \exp(0.25 + \ln(1.5)x_1 - 0.25x_2)$, where $x_1 \sim \text{Bernoulli}(0.5)$ and $x_2 \sim \text{Uniform}(0,1)$, and $\phi = 0.5$, $\phi = 1$, and $\phi = 1.5$. A Bernoulli variable was then generated for which the probability was given by $P(Z = 1) = \frac{1}{1+\exp(0.6-2.25x_1)}$. If the Bernoulli variable was equal to 1, then the outcome variable was replaced with a 0. This yields an outcome variable, y, that follows a zero-inflated double Poisson distribution. This outcome variable was modeled using zero-inflated double Poisson regression, zero-inflated marginalized double Poisson regression, zero-inflated negative binomial regression, and zero-inflated Poisson regression. The simulation was repeated 1000 times for sample sizes of 30, 100, and 1000 observations. Again, the likelihood ratio test statistic was used to estimate power, and 95% confidence intervals for the coefficients were considered. The zero-inflated simulations was limited to 20 iterations.

These simulations allow for comparison of the double Poisson regression models with their more commonly used counterparts. Power of the tests that $\phi = 1$ and $\alpha = 0$ for double Poisson and negative binomial regression, respectively, can be compared to determine how double Poisson regression compares to negative binomial regression. Further, the mean value of the coefficient estimates can be examined and compared to the true values to determine their accuracy.

Chapter 5

Results

The likelihood ratio test incorrectly rejected the null hypothesis for approximately 5% of the simulations when the outcome values were from a Poisson distribution. When each variable had 30 observations, the test incorrectly rejected in 8.1% of the simulations. This value is higher than expected, which may be due to a size distortion in the likelihood ratio test. Because the rejection region produced is an asymptotic result, the size of a likelihood ratio test is an approximation which becomes more accurate as the sample size increases (Cizek, Hardle, & Weron, 2005). This percentage decreased to 6.2% incorrect rejections when there were 1000 observations. This result is reasonably near the 5% significance level and allows for the conclusion that the test is performing as expected.

It is important to note that not all simulations converge, which leads to fewer observations. The least number of successful simulations was 369 for the underdispersed zero-inflated negative binomial regression model with 1000 observations. Full information on the number of successfully converged simulations and detailed results are provided in the appendix.

5.1 Overdispersed Data Simulation Results

To generate data that are overdispersed relative to the Poisson, we use $\phi = 0.5$, and the power of the test for $\phi = 1$ (and $\alpha = 0$) are determined using the simulation results. For both negative binomial and double Poisson regression, the power of these tests is not discernibly different from 1.000 when the sample size is 1000 observations. This indicates that both models recognize and account for the presence of overdispersion relative to the Poisson in the data. However, when the sample size is smaller there is a slight difference in power between the two types of regression. The negative binomial regression model's test for $\alpha = 0$ correctly rejects 47.3% (95% CI: $(0.442, 0.504)$) and 97.1% (95% CI: $(0.961, 0.981)$) of the time for sample sizes of 30 and 100, respectively. The double Poisson model's likelihood ratio test for $\phi = 1$ slightly outperforms with powers of 52.2% (95% CI: $(0.491, 0.553)$)and 98.8% (95% CI: $(0.981, 0.995)$), respectively.

The double Poisson model is also more effective at accurately estimating the coefficients. Table 5.1 shows the mean coefficient estimates for the three regression models considered. Note that the Double Poisson regression model achieves the closest mean value among the models.

The 95% confidence intervals for the coefficients are also considered. It is expected that the intervals contain the true values of the coefficients in 95% of the simulations. This is true for the intervals produced for the double Poisson regression models. The proportion of intervals not containing the true values of the coefficients range from 0.045 to 0.064 across the three sample sizes considered.

When the sample sizes is 30, the confidence intervals produced for the coefficients in negative binomial regression perform reasonably well with the intervals for β_0, β_1, and β_2, not containing the true value in only 6.5%, 5.8%, and 5.2% of simulations.

Table 5.1 Mean Values of the Coefficients for the Overdispersed Double Poisson Data ($n = 1000$)

	True Values	Poisson	Negative Binomial	Double Poisson
$\boldsymbol{\beta}_0$	0.25	0.3182214	0.3235692	0.2470219
$\boldsymbol{\beta}_1$	≈ 0.405	0.3585286	0.3642256	0.4082656
$\boldsymbol{\beta}_2$	0.25	0.2155314	0.2268638	0.2532754

However, as the sample size is increased, the intervals do not contain the true value in a higher proportion of the simulations. For example, when the sample size is 1000, the 95% confidence interval for β_0 does not contain 0.25 in 19.2% of the simulations performed for the negative binomial model. The confidence intervals produced for the Poisson regression models do not include the true value over 10% of the time for all sample size and all coefficients. The largest proportion of intervals produced for a Poisson model not containing the true value occurs with a sample size of 1000, with 33.1% of confidence intervals not containing the true value of β_0.

These results indicate that double Poisson regression can outperform negative binomial regression in some cases. When the data are overdispersed relative to the Poisson and the outcome values follow a double Poisson distribution, the double Poisson model is a better fitting model than both negative binomial and Poisson models.

5.2 Underdispersed Data Simulation Results

The power results for the data which were generated to be underdispersed relative to the Poisson with $\phi = 1.5$ clearly demonstrate the unique modeling capabilities of double Poisson regression models compared to negative binomial models. The power of the test of $\alpha = 0$ for the negative binomial models fell below 0.005 for all sample sizes considered. This is to be expected because the negative binomial model can only account for overdispersion. Double Poisson regression can account for underdispersion and this is reflected in the estimated power. The estimated power of the test for $\phi = 1$ was 0.510 (95% CI: $(0.479, 0.541)$) when the sample size was 30 observations. It increased to 0.916 (95% CI: $(0.899, 0.933)$) when the sample was 100 and increased further to indiscernibly different from 1.000 when the sample size was 1000 observations.

Both negative binomial and double Poisson regression as well as Poisson regression

Table 5.2 Mean Values of the Coefficients for the Underdispersed Double Poisson Data ($n = 1000$)

	True Values	Poisson	Negative Binomial	Double Poisson
$\boldsymbol{\beta}_0$	0.25	0.2658918	0.2658919	0.2486466
$\boldsymbol{\beta}_1$	≈ 0.405	0.4010003	0.4010003	0.4069364
$\boldsymbol{\beta}_2$	0.25	0.2490116	0.2490116	0.2522818

yield fairly accurate estimations of the coefficients, as seen in Table 5.2. Note that the negative binomial coefficients are identical to the Poisson coefficients when the sample size is 1000. This occurs because the test for $\alpha = 0$ did not reject in any of the simulation at this sample size.

For the Double Poisson regression models the proportion of times the 95% confidence intervals for the coefficients did not contain the true values were close to 0.05 for all sample sizes considered. This indicates that the confidence intervals are performing as expected. The proportions for both the negative binomial and Poisson regression models were lower, indicating that a higher proportion of the 95% confidence intervals contained the true values of the coefficients.

5.3 Equidispersed Data Simulation Results

Expectations for the analysis of the likelihood ratio test for the data which are equidispersed relative to the Poisson are different from those for both the under- and overdispersed data. Because the double Poisson distribution simplifies to the Poisson when $\phi = 1$, the percentage of rejections of the test of $\phi = 1$ and $\alpha = 0$ should approximately equal the significance level. This occurs for the double Poisson regression models with a proportion of rejections of 0.084 (95% CI: $(0.067, 0.101)$), 0.058 (95% CI: $(0.044, 0.072)$), and 0.051 (95% CI: $(0.037, 0.065)$) when the sample sizes are 30, 100, and 1000, respectively. The proportion of rejections for the negative binomial

Table 5.3 Mean Values of the Coefficients for the Equidispersed Double Poisson Data ($n = 1000$)

	True Values	Poisson	Negative Binomial	Double Poisson
$\boldsymbol{\beta}_0$	0.25	0.2478945	0.248452	0.2476479
$\boldsymbol{\beta}_1$	≈ 0.405	0.4074398	0.4069786	0.4076067
$\boldsymbol{\beta}_2$	0.25	0.2531785	0.2528882	0.2532765

models are all below the significance level of 0.05, with the highest estimate at 0.023 when n=1000.

Following expectations, all three models produce accurate estimates of the coefficients, as seen in Table 5.3. In addition, the 95% confidence intervals produced by the models all contain the true value of the coefficient in approximately 95% of the cases.

5.4 Overdispersed Zero-Inflated Data Simulation Results

Data are generated to be overdispersed relative to the Poisson using $\phi = 0.5$ in these simulations. Thus, it is expected that the tests for $\phi = 1$ and $\alpha = 0$ would be rejected and the proportion of times the simulations correctly reject yields an estimate of the power. For all three tests considered, the power increases with sample size, as expected. The power estimates for the test of $\alpha = 0$ in the zero-inflated negative binomial model are similar to the estimates for the test of $\phi = 1$ in the zero-inflated double Poisson model, with estimates of 0.942 (95% CI: $(0.928, 0.956)$) and 0.944 (95% CI: $(0.930, 0.958)$) when n=1000, respectively. The power estimates for the zero-inflated marginalized double Poisson test of $\phi = 1$ were only slightly higher with a power of 0.949 (95% CI: $(0.935, 0.963)$) when n=1000. Thus, all three models correctly identified overdispersion relative to the Poisson with approximately the same power.

Table 5.4 Mean Values of the Coefficients for the Overdispersed Zero-Inflated Double Poisson Data ($n = 1000$)

	True Values	ZIP	ZINB	ZIDP	ZIMDP
$\boldsymbol{\beta}_0$	0.25	0.6192273	0.4892453	0.2312524	-0.2141435
$\boldsymbol{\beta}_1$	$\approx$ 0.405	0.2850461	0.309542	0.4057106	-0.9804658
$\boldsymbol{\beta}_2$	0.25	0.1686814	0.1855295	0.243487	0.2461409

Only the zero-inflated double Poisson model estimates the non-zero excess coefficients well. As seen in Table 5.4, both the zero-inflated Poisson and the zero-inflated negative binomial overestimate β_0 and underestimate both β_1 and β_2, on average. Recall, the zero-inflated marginalized double Poisson is not expected to estimate the coefficients used to generate the data because these coefficients apply to the overall data and not just the non-zero excess group of data. However, note that the mean of the β_2 estimates is close to its true value. This occurs because x_2 does not have a role in the zero inflation, thus the non-marginalized coefficient is expected to be the same as the marginalized coefficient.

The 95% confidence intervals for the coefficients reflect these results as well. Approximately 95% of the confidence intervals for the coefficients produced by the zero-inflated double Poisson model contain the true values of the coefficients. The confidence intervals produced by the other three models considered do not due to their inaccurate estimates of the coefficients themselves. However, all four models considered accurately identify that x_2 does not have a role in the zero inflation. For all models, approximately 5% of the 95% confidence intervals for the inflation coefficient for x_2 do not contain the true value of 0, which is expected given the significance level of 0.05.

5.5 Underdispersed Zero-Inflated Data Simulation Results

To generate data which are underdispersed relative to the Poisson, the dispersion parameter $\phi = 1.5$. Thus, similarly to the uninflated data simulations considered previously, the test for $\alpha = 0$ in zero-inflated negative binomial regression should not reject. The test performs as expected with an estimated power of 0.003 (95% CI: $(0, 0.007)$) when n = 30 and estimated powers of 0 for the other sample sizes considered. For the test of $\phi = 1$, the zero-inflated marginalized double Poisson model yields a power slightly higher than the non-marginalized model. When n = 1000, the estimated power for the zero-inflated double Poisson model is 0.680 (95% CI: $(0.651, 0.709)$) and the estimated power of the marginalized model is 0.821 (95% CI: $(0.795, 0.847)$).

Again, zero-inflated double Poisson regression is the only model that produces estimates near the real values of coefficients for the non-zero excess group. As seen in Table 5.5, the zero-inflated Poisson and zero-inflated negative binomial models overestimate β_1 and β_2 but underestimate β_0. The mean value of the coefficient estimates for β_0 and β_1 for the zero-inflated marginalized double Poisson are not expected to be near the values used to generate the distribution. As expected, the estimate for β_2 is accurate.

The results concerning the 95% confidence intervals are similar to the results from the zero-inflated overdispersed data. The confidence intervals produced in the zero-inflated double Poisson simulations do not contain the true value approximately 5%

Table 5.5 Mean Values of the Coefficients for the Underdispersed Zero-Inflated Double Poisson Data ($n = 1000$)

	True Values	ZIP	ZINB	ZIDP	ZIMDP
$\boldsymbol{\beta}_0$	0.25	0.0476455	0.0403438	0.2476349	-0.1911103
$\boldsymbol{\beta}_1$	≈ 0.405	0.5017133	0.5083450	0.4082345	-0.9722466
$\boldsymbol{\beta}_2$	0.25	0.2871264	0.3019847	0.2457806	0.2487429

of the time. Intervals produced by the other models do not contain the true values more often because their estimates of the coefficients were inaccurate. All four models produced confidence intervals that contained the true value for the inflation coefficient for x_2 approximately 95% of the time, as expected.

5.6 Equidispersed Zero-Inflated Data Simulation Results

Given that the data are generated using $\phi = 1$ leading the double Poisson to reduce to a Poisson distribution, it is expected that the test for $\phi = 1$ would fail to reject in approximately 95% of cases (using a significance level of 0.05). This is the case for both the zero-inflated double Poisson and the zero-inflated marginalized double Poisson models. When n = 1000, the test for $\phi = 1$ incorrectly rejects in 5.1% of the simulations (95% CI: $(0.037, 0.065)$) for the zero-inflated double Poisson model and 5.8% (95% CI: $(0.044, 0.072)$) for the marginalized model. Similar to the uninflated cases, the proportion of rejections for the test of $\alpha = 0$ are all below the significance level of 0.05.

The zero-inflated Poisson, negative binomial, and double Poisson models all accurately estimate the coefficients for the non-zero excess group. The mean values of the coefficients produced by the simulations are given in Table 5.6. The zero-inflated marginalized double Poisson model again does not accurately estimate β_0 or β_1 as the coefficients have been modified to apply to the overall population. Again, this model's estimate of β_2 is accurate on average because x_2 is not associated with the zero inflation.

Table 5.6 Mean Values of the Coefficients for the Equidispersed Zero-Inflated Double Poisson Data ($n = 1000$)

	ZIP	ZINB	ZIDP	ZIMDP	
$\boldsymbol{\beta}_0$	0.25	0.2492326	0.2435707	0.2421238	-0.1941309
$\boldsymbol{\beta}_1$	≈ 0.405	0.4057610	0.4072516	0.4077789	-0.9818405
$\boldsymbol{\beta}_2$	0.25	0.2454314	0.2461351	0.2458520	0.2456174

The confidence intervals produced by the regression models reinforce these results. The 95% confidence intervals for the coefficients contained the true values in approximately 95% of simulations in the three models which produced accurate estimates. Further, all models identified that x_2 is not involved in the zero-inflation. Thus, the confidence intervals for the inflated coefficient for x_2 contained the true value of 0 in approximately 95% of simulations as well.

Chapter 6

Discussion

The majority of the simulations produced expected results. The double Poisson models performs as well as or better than the negative binomial models when the data are generated to be overdispersed relative to the Poisson. This indicates that double Poisson regression is a viable alternative to a negative binomial model. The results also indicate that double Poisson regression incorporates and correctly identifies the possibility of underdispersion relative to the Poisson. This is an advantage over negative binomial regression, which can only model overdispersion. When data are generated to be equidispersed relative to the Poisson, both the double Poisson and negative binomial models can correctly identify the equidispersion and model the data appropriately.

The zero-inflated simulations lead to similar conclusions. When the data are overdispersed, the zero-inflated double Poisson models perform as well as the zero-inflated negative binomial models in terms of power. However, the coefficient estimates produced by the zero-inflated double Poisson model are much better than the estimates for both the zero-inflated Poisson and zero-inflated negative binomial models. The zero-inflated double Poisson model can still recognize underdispersion in the presence of zero-inflation, while the zero-inflated negative binomial model cannot. Finally, when there is equidispersion and zero-inflation in the data, the zero-inflated double Poisson model accurately estimates the coefficients, on average, and incorrectly rejects the test at the significance level.

Although these results are promising, there is further research to be done. For example, the double Poisson models considered here could be compared to other alternative models such as generalized Poisson or quasi-likelihood models. Further, double Poisson regression could be used to model data generated using the negative binomial distribution to evaluate the model's performance. Finally, the models considered here should be applied to real data to determine the applicability in a real world setting.

References

Chou, N. T. & Steenhard, D. (2009). A Flexible Count Data Regression Model Using SAS PROC NLMIXED. *SAS Language Reference Dictionary, 4.* SAS Institute Inc., Cary, N.C.

Cizek, P., Hardle, W., & Weron, R. (2005). Likelihood Ratio Test. *Statistical Tools for Finance and Insurance.*

Efron, B. (1986). Double Exponential Families and Their Use in Generalized Linear Regression. *Journal of American Statistical Association, 81*(295), 709-721.

Long, D. L., Preisser, J. S., Herring, A. H., & Golin, C. E. (2014). A Marginalized Zero-Inflated Poisson Regression Model with Overall Exposure Effects. *Statistics in Medicine, 33*(29), 5151-5165.

Martin, J. & Hall, D. B. (2016). Marginalized Zero-Inflated Regression Models for Count Data. Manuscript submitted for publication.

Zao, Y., Geedipally, S. R., & Lord, D. (2013). Evaluating the Double Poisson Generalized Linear Model. *Accident Analysis & Prevention, 59*, 497-505.

Appendix A

Full Results

A.1 Simulations Testing Size of the Likelihood Ratio Test

Table A.1 Summary Table of the Test Statistics

	Mean	Std. Dev.	Minimum	Maximum
$n = 30$	1.236398	1.752519	0.0000135	13.68773
$n = 100$	1.120399	1.607786	0.0000018	12.74409
$n = 1000$	1.092899	1.533460	0.0000011	12.86238

A.2 Overdispersed Data Simulation Tables

Table A.2 Summary Table for Double Poisson Regression on Overdispersed Data $n = (30)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	1000	5.252571	4.832934	0.0006935	32.17255
β_0	1000	0.1794764	0.5389272	-2.742061	1.552799
β_1	1000	0.3984457	0.4414885	-1.076519	2.289521
β_2	1000	0.256771	0.7450893	-1.931567	3.851027
ϕ	1000	0.6022566	0.2000388	0.1713118	1.35781

Table A.3 Summary Table for Double Poisson Regression on Overdispersed Data $n = (100)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	1000	20.57958	10.50707	0.8151261	72.67791
β_0	1000	0.2358402	0.2725943	-0.8794453	1.105812
β_1	1000	0.4020919	0.2343711	-0.4087247	1.313856
β_2	1000	0.2409035	0.3919653	-1.279709	1.773609
ϕ	1000	0.5247047	0.994492	0.2238183	0.8825715

Table A.4 Summary Table for Double Poisson Regression on Overdispersed Data $n = (1000)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	1000	214.7114	33.31131	123.7116	328.0908
β_0	1000	0.2470219	0.0867090	-0.1012812	0.4872373
β_1	1000	0.4082656	0.0724030	0.1484432	0.6705977
β_2	1000	0.2532754	0.1233902	-0.1896524	0.6627502
ϕ	1000	0.5014819	0.0300975	0.4128276	0.5983468

Table A.5 Estimated Power of Test for $\phi = 1$ for Overdispersed Data ($\phi = 0.5$) by Sample Size

	$n = 30$	$n = 100$	$n = 1000$
power	0.522	0.988	1.000

Table A.6 Proportion of 95% Confidence Intervals Not Containing True Coefficient Value for Double Poisson Regression on Overdispersed Data

	$n = 30$	$n = 100$	$n = 1000$
β_0	0.048	0.051	0.054
β_1	0.045	0.051	0.055
β_2	0.047	0.064	0.057

Table A.7 Summary Table for Negative Binomial Regression on Overdispersed Data $n = (30)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	999	4.782994	4.687690	0	32.05104
β_0	999	0.2703667	0.4507859	-2.009444	1.572041
β_1	999	0.3574001	0.3818474	-0.8521115	2.015971
β_2	999	0.2331022	0.6650585	-1.8109394	2.923036
α	999	0.3795166	0.2627993	0.000000	1.468267

Table A.8 Summary Table for Negative Binomial Regression on Overdispersed Data $n = (100)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	1000	19.2873	10.4616	0.5829353	71.26928
β_0	1000	0.3177267	0.2329328	-0.5926527	1.113593
β_1	1000	0.3586104	0.2055743	-0.3705224	1.028521
β_2	1000	0.2164119	0.3507152	-1.025018	1.542849
α	1000	0.4460906	0.1589968	0.0674152	1.183223

Table A.9 Summary Table for Negative Binomial Regression on Overdispersed Data $n = (1000)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	1000	203.0857	32.95358	114.7579	306.4024
β_0	1000	0.3235692	0.0750832	0.0238671	0.5410829
β_1	1000	0.3642256	0.0639928	0.1309904	0.5917957
β_2	1000	0.2268638	0.1109632	-0.1782868	0.5693384
α	1000	0.4666695	0.0493975	0.3279201	0.6434028

Table A.10 Estimated Power of Test for $\alpha = 0$ for Overdispersed Data by Sample Size

	$n = 30$	$n = 100$	$n = 1000$
power	0.473	0.971	1.000

Table A.11 Proportion of 95% Confidence Intervals Not Containing True Coefficient Value for Negative Binomial Regression on Overdispersed Data

	$n = 30$	$n = 100$	$n = 1000$
β_0	0.065	0.062	0.192
β_1	0.058	0.060	0.100
β_2	0.052	0.062	0.070

Table A.12 Summary Table for Poisson Regression on Overdispersed Data $n = (30)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
β_0	1000	0.2730359	0.4482954	-1.569089	1.505727
β_1	1000	0.3559254	0.3798942	-0.799080	2.031450
β_2	1000	0.2296032	0.6530374	-1.617517	3.103088

Table A.13 Summary Table for Poisson Regression on Overdispersed Data $n = (100)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
β_0	1000	0.3182214	0.2334504	-0.6200951	1.105406
β_1	1000	0.3585286	0.2056042	-0.3682128	1.031536
β_2	1000	0.2155314	0.3501057	-1.091168	1.580662

Table A.14 Summary Table for Poisson Regression on Overdispersed Data
$n = (1000)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
β_0	1000	0.3234253	0.0747642	0.0201285	0.5373703
β_1	1000	0.3642342	0.0639324	0.1338744	0.5931463
β_2	1000	0.2271444	0.1104250	-0.1693781	0.5882490

Table A.15 Proportion of 95% Confidence Intervals Not Containing True Coefficient Value for Poisson Regression on Overdispersed Data

	$n = 30$	$n = 100$	$n = 1000$
β_0	0.159	0.166	0.331
β_1	0.123	0.150	0.229
β_2	0.128	0.147	0.161

A.3 Underdispersed Data Simulation Tables

Table A.16 Summary Table for Double Poisson Regression on Underdispersed Data $n = (30)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	1000	4.753740	3.925275	0.0000779	23.85881
β_0	1000	0.2228024	0.2716971	-0.7643115	1.043774
β_1	1000	0.4058389	0.2320543	-0.2733662	1.301085
β_2	1000	0.2619937	0.3951523	-1.154958	1.556912
ϕ	1000	1.710077	0.4578059	0.6394044	4.543094

Table A.17 Summary Table for Double Poisson Regression on Underdispersed Data $n = (100)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	1000	10.97120	5.718872	0.0002565	35.33717
β_0	1000	0.2465704	0.1489448	-0.4223371	0.7562209
β_1	1000	0.4051084	0.1275869	0.0022292	0.8138250
β_2	1000	0.2482277	0.2180415	-0.5844415	1.177946
ϕ	1000	1.560927	0.1965056	0.9014463	2.333706

Table A.18 Summary Table for Double Poisson Regression on Underdispersed Data $n = (1000)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	1000	89.94805	17.18504	42.94072	145.0057
β_0	1000	0.2486466	0.0467694	0.0829530	0.3830697
β_1	1000	0.4069364	0.0402929	0.2690029	0.5525697
β_2	1000	0.2522818	0.0674886	0.0051806	0.4811692
ϕ	1000	1.503652	0.0595233	1.324215	1.680926

Table A.19 Estimated Power of Test for $\phi = 1$ for Underdispersed Data ($\phi = 1.5$) by Sample Size

	$n = 30$	$n = 100$	$n = 1000$
power	0.510	0.916	1.000

Table A.20 Proportion of 95% Confidence Intervals Not Containing True Coefficient Value for Double Poisson Regression on Underdispersed Data

	$n = 30$	$n = 100$	$n = 1000$
β_0	0.056	0.061	0.057
β_1	0.044	0.050	0.054
β_2	0.055	0.065	0.057

Table A.21 Summary Table for Negative Binomial Regression on Underdispersed Data $n = (30)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	996	0.0054689	0.0967834	0	2.84234
β_0	996	0.2456628	0.2652086	-0.792607	1.050039
β_1	996	0.3957214	0.2268664	-0.2692037	1.281893
β_2	996	0.2547358	0.3861745	-1.146024	1.463288
α	996	0.001018	0.0117172	0	0.2843519

Table A.22 Summary Table for Negative Binomial Regression on Underdispersed Data $n = (100)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	999	0.0007803	0.0210217	0	0.6522511
β_0	999	0.265392	0.1472285	-0.3740571	0.7667754
β_1	999	0.3976704	0.1252107	0.0021633	0.7904198
β_2	999	0.2446858	0.214767	-0.5820748	1.143437
α	999	0.0000968	0.0023587	0.0000000	0.069384

Table A.23 Summary Table for Negative Binomial Regression on Underdispersed Data $n = (1000)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	1000	0	0	0	0
β_0	1000	0.2658919	0.0464366	0.1020487	0.3991504
β_1	1000	0.4010003	0.0398293	0.2640791	0.5448037
β_2	1000	0.2490116	0.0666635	0.0051014	0.4745244
α	1000	0.0000000	0.0000000	0	0.000000

Table A.24 Estimated Power of Test for $\alpha = 0$ for Underdispersed Data by Sample Size

	$n = 30$	$n = 100$	$n = 1000$
power	0.004	0.001	0.000

Table A.25 Proportion of 95% Confidence Intervals Not Containing True Coefficient Value for Negative Binomial Regression on Underdispersed Data

	$n = 30$	$n = 100$	$n = 1000$
β_0	0.013	0.021	0.025
β_1	0.014	0.020	0.024
β_2	0.014	0.023	0.015

Table A.26 Summary Table for Poisson Regression on Underdispersed Data $n = (30)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
β_0	1000	0.2439336	0.2671132	-0.7926069	1.050039
β_1	1000	0.3969512	0.2273209	-0.2692037	1.281851
β_2	1000	0.2570239	0.3884445	-1.146145	1.463273

Table A.27 Summary Table for Poisson Regression on Underdispersed Data $n = (100)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
β_0	1000	0.2652744	0.1472012	-0.3740571	0.7667754
β_1	1000	0.3980615	0.1257572	0.0021633	0.7904197
β_2	1000	0.2444606	0.2147778	-0.5820748	1.143437

Table A.28 Summary Table for Poisson Regression on Underdispersed Data $n = (1000)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
β_0	1000	0.2658918	0.0464366	0.1020487	0.3991504
β_1	1000	0.4010003	0.0398283	0.2640791	0.5448037
β_2	1000	0.2490116	0.0666635	0.0051014	0.4745244

Table A.29 Proportion of 95% Confidence Intervals Not Containing True Coefficient Value for Poisson Regression on Underdispersed Data

	$n = 30$	$n = 100$	$n = 1000$
β_0	0.011	0.020	0.025
β_1	0.010	0.020	0.024
β_2	0.011	0.022	0.015

A.4 Equidispersed Data Simulation Tables

Table A.30 Summary Table for Double Poisson Regression on Equidispersed Data $n = (30)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	1000	1.215848	1.708021	0.0000005	15.40638
β_0	1000	0.2047934	0.3334695	-0.8818812	1.185652
β_1	1000	0.4084277	0.2881226	-0.4457929	1.418979
β_2	1000	0.26526	0.4924153	-1.581586	2.097764
ϕ	1000	1.150437	0.309281	0.4568165	2.870911

Table A.31 Summary Table for Double Poisson Regression on Equidispersed Data $n = (100)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	1000	1.098242	1.535847	0.0000009	14.79622
β_0	1000	0.2444993	0.1809511	-0.4853991	0.870132
β_1	1000	0.4040129	0.1568595	-0.0717125	0.8560722
β_2	1000	0.2445719	0.1674831	-0.747699	1.270228
ϕ	1000	1.037485	0.1480723	0.6399453	1.69829

Table A.32 Summary Table for Double Poisson Regression on Equidispersed Data $n = (1000)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	1000	1.023906	1.423115	0.0000001	10.85135
β_0	1000	0.2476479	0.056632	0.0378391	0.4195695
β_1	1000	0.4076067	0.0489802	0.2230007	0.5835963
β_2	1000	0.2532765	0.0829342	-0.0551421	0.4977086
ϕ	1000	0.9999967	0.0444457	0.8645381	1.126228

Table A.33 Estimated Power of Test for $\phi = 1$ for Equidispersed Data ($\phi = 1$) by Sample Size

	$n = 30$	$n = 100$	$n = 1000$
power	0.084	0.058	0.051

Table A.34 Proportion of 95% Confidence Intervals Not Containing True Coefficient Value for Double Poisson Regression on Equidispersed Data

	$n = 30$	$n = 100$	$n = 1000$
β_0	0.046	0.054	0.044
β_1	0.052	0.052	0.048
β_2	0.053	0.061	0.046

Table A.35 Summary Table for Negative Binomial Regression on Equidispersed Data $n = (30)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	993	0.2127494	0.6368952	0	7.607454
β_0	993	0.2126885	0.3327328	-0.8924742	1.193916
β_1	993	0.4054783	0.2862568	-0.4434526	1.425267
β_2	993	0.2640799	0.4889527	-1.401134	2.112465
α	993	0.0324342	0.0705431	0	0.5380058

Table A.36 Summary Table for Negative Binomial Regression on Equidispersed Data $n = (100)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	993	0.3789205	0.9962619	0	10.74076
β_0	993	0.2481477	0.1789776	-0.4908126	0.8670005
β_1	993	0.4013752	0.1553192	-0.0718296	0.8542461
β_2	993	0.2440914	0.2653531	-0.7399939	1.281
α	993	0.0254452	0.0459111	0	0.3084958

Table A.37 Summary Table for Negative Binomial Regression on Equidispersed Data $n = (1000)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	911	0.5173525	1.072668	0	8.442913
β_0	911	0.248452	0.0575425	0.0377225	0.4197744
β_1	911	0.4069786	0.0493231	0.2229566	0.5838591
β_2	911	0.2528882	0.0840476	-0.0551652	0.4959723
α	911	0.0106571	0.014305	0	0.0768921

Table A.38 Estimated Power of Test for $\alpha = 0$ for Equidispersed Data by Sample Size

	$n = 30$	$n = 100$	$n = 1000$
power	0.005	0.016	0.023

Table A.39 Proportion of 95% Confidence Intervals Not Containing True Coefficient Value for Negative Binomial Regression on Equidispersed Data

	$n = 30$	$n = 100$	$n = 1000$
β_0	0.049	0.043	0.040
β_1	0.043	0.049	0.046
β_2	0.041	0.049	0.055

Table A.40 Summary Table for Poisson Regression on Equidispersed Data $n = (30)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
β_0	1000	0.2124704	0.3320219	-0.892474	1.172475
β_1	1000	0.405326	0.2861443	-0.4264888	1.425267
β_2	1000	0.2637161	0.487943	-1.413682	2.112465

Table A.41 Summary Table for Poisson Regression on Equidispersed Data $n = (100)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
β_0	1000	0.24716	0.1805091	-0.4908126	0.866805
β_1	1000	0.4028542	0.1524052	-0.0718296	0.8542756
β_2	1000	0.2439122	0.2667755	-0.7430642	1.281

Table A.42 Summary Table for Poisson Regression on Equidispersed Data $n = (1000)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
β_0	1000	0.2478945	0.0565986	0.0377129	0.4192198
β_1	1000	0.4074398	0.0490007	0.2229717	0.5838558
β_2	1000	0.2531785	0.0828944	-0.0551652	0.4965525

Table A.43 Proportion of 95% Confidence Intervals Not Containing True Coefficient Value for Poisson Regression on Equidispersed Data

	$n = 30$	$n = 100$	$n = 1000$
β_0	0.041	0.045	0.044
β_1	0.040	0.048	0.049
β_2	0.040	0.051	0.046

A.5 Zero-Inflated Overdispersed Data Simulation Tables

Table A.44 Summary Table for Zero-Inflated Double Poisson Regression on Zero-Inflated Overdispersed Data $n = (30)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	832	2.133595	2.701721	0	14.53686
β_0	832	0.1781066	4.11781	-111.1633	4.106563
β_1	832	0.1360411	2.14738	-9.20358	45.78965
β_2	832	0.30117	3.484169	-12.06849	77.7935
γ_0	832	-22.61041	138.7684	-881.5131	456.9386
γ_1	832	27.79032	111.7806	-455.5499	788.0507
γ_2	832	1.425697	236.5352	-1197.677	1792.086
$\ln\phi$	832	0.4030363	1.464342	-4.96128	8.088055

Table A.45 Summary Table for Zero-Inflated Double Poisson Regression on Zero-Inflated Overdispersed Data $n = (100)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	1000	2.122629	2.89807	0.0000155	16.81099
β_0	1000	0.0040202	2.358131	-52.54269	1.906128
β_1	1000	0.2310982	1.51532	-23.11324	18.13352
β_2	1000	0.2134907	1.753881	-16.49625	31.42447
γ_0	1000	-6.083835	51.14625	-767.0179	139.4602
γ_1	1000	6.060447	33.39295	-248.985	743.0478
γ_2	1000	3.100089	66.42996	-851.1842	816.7293
$\ln\phi$	1000	-0.4643417	0.7783319	-4.537404	3.421355

Table A.46 Summary Table for Zero-Inflated Double Poisson Regression on Zero-Inflated Overdispersed Data $n = (1000)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	1000	15.92631	8.818674	0.0800209	53.8463
β_0	1000	0.2312524	0.2725079	-1.535472	0.8578113
β_1	1000	0.4057106	0.1829238	-0.222983	1.327217
β_2	1000	0.243487	0.2668995	-0.894491	1.140538
γ_0	1000	-0.6370899	0.566746	-12.00446	0.4305602
γ_1	1000	2.289708	0.4816239	1.54492	13.15071
γ_2	1000	0.0021899	0.4262011	-1.527482	1.68484
$\ln \phi$	1000	-0.6784985	0.2398018	-1.90263	-0.0444952

Table A.47 Estimated Power of Test for $\phi = 1$ (ZIDP) for Zero-Inflated Overdispersed Data ($\phi = 0.5$) by Sample Size

	$n = 30$	$n = 100$	$n = 1000$
power	0.191	0.185	0.944

Table A.48 Proportion of 95% Confidence Intervals Not Containing True Coefficient Value for Zero-Inflated Double Poisson Regression on Zero-Inflated Overerdispersed Data

	$n = 30$	$n = 100$	$n = 1000$
β_0	0.279	0.168	0.060
β_1	0.179	0.055	0.045
β_2	0.156	0.068	0.048

Table A.49 Summary Table for Zero-Inflated Marginalized Double Poisson Regression on Zero-Inflated Overdispersed Data $n = (30)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	584	1.987453	2.876549	0	17.56076
β_0	584	-0.379276	2.63248	-48.0155	3.325235
β_1	584	-0.8532335	1.473254	-7.596896	19.69532
β_2	584	0.2588249	2.117019	-8.139478	29.51854
γ_0	584	-1.25416	4.676127	-26.2365	13.17938
γ_1	584	1.475935	11.87727	-216.4305	29.16136
γ_2	584	-0.2984907	5.497769	-67.5534	16.79074
$\ln \phi$	584	0.6394292	1.562738	-4.442347	9.032578

Table A.50 Summary Table for Zero-Inflated Marginalized Double Poisson Regression on Zero-Inflated Overdispersed Data $n = (100)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	972	1.979347	2.730136	0	16.58419
β_0	972	-0.5993433	3.534369	-74.88737	0.9771867
β_1	972	-1.138786	3.010834	-80.04905	35.1757
β_2	972	0.2279579	2.525642	-44.65187	43.48825
γ_0	972	-1.882026	4.854911	-42.55289	3.707883
γ_1	972	3.502794	4.641642	-18.02847	24.69835
γ_2	972	-0.0500096	3.376262	-34.33688	42.37562
$\ln \phi$	972	-0.4371546	0.7781494	-5.109459	2.372452

Table A.51 Summary Table for Zero-Inflated Marginalized Double Poisson Regression on Zero-Inflated Overdispersed Data $n = (1000)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	1000	15.90248	8.794727	0.1172237	53.73215
β_0	1000	-0.2141435	0.1927202	-1.496524	0.2342533
β_1	1000	-0.9804658	0.1814214	-1.543492	-0.1770775
β_2	1000	0.2461409	0.2353194	-0.579769	1.013883
γ_0	1000	-0.6346799	0.6093537	-13.80679	0.4833203
γ_1	1000	2.288801	0.5312893	1.528534	14.90846
γ_2	1000	-0.0046805	0.3990355	-1.589375	1.455736
$\ln \phi$	1000	-0.6772788	0.2400818	-1.901218	-0.0425737

Table A.52 Estimated Power of Test for $\phi = 1$ (ZIMDP) for Zero-Inflated Overdispersed Data ($\phi = 0.5$) by Sample Size

	$n = 30$	$n = 100$	$n = 1000$
power	0.175	0.172	0.949

Table A.53 Proportion of 95% Confidence Intervals Not Containing True Coefficient Value for Zero-Inflated Double Poisson Regression on Zero-Inflated Underdispersed Data

	$n = 30$	$n = 100$	$n = 1000$
β_0	0.101	0.087	0.789
β_1	0.293	0.719	0.999
β_2	0.092	0.056	0.037

Table A.54 Summary Table for Zero-Inflated Negative Binomial Regression on Zero-Inflated Overdispersed Data $n = (30)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	997	0.7541072	1.699569	0	12.35512
β_0	997	0.3841768	1.144328	-10.66285	6.569091
β_1	997	-1016.61	24553.79	-745789.4	7.469898
β_2	997	0.1554803	2.410565	-11.59645	35.00412
γ_0	997	-34.34995	179.1517	-1176.265	741.9562
γ_1	997	52.98017	156.9622	-631.7735	888.6647
γ_2	997	-1.178229	301.3836	-1189.27	1371.395
$\ln \alpha$	997	-59.23692	652.2322	-20220.47	2.176748

Table A.55 Summary Table for Zero-Inflated Negative Binomial Regression on Zero-Inflated Overdispersed Data $n = (100)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	1000	1.723751	2.837179	0	17.0316
β_0	1000	0.4464112	0.4894456	-1.82785	1.941923
β_1	1000	0.2196791	0.5934819	-4.687164	1.866187
β_2	1000	0.1756218	0.7502654	-2.734546	3.278298
γ_0	1000	-5.359116	50.45087	-759.2	201.8522
γ_1	1000	6.787451	41.96632	-157.9382	706.8228
γ_2	1000	1.961829	67.81144	-876.5267	790.0993
$\ln \alpha$	1000	-11.86491	35.27753	-635.861	1.121496

Table A.56 Summary Table for Zero-Inflated Negative Binomial Regression on Zero-Inflated Overdispersed Data $n = (1000)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	1000	15.59047	8.765664	0.0299717	52.79344
β_0	1000	0.4892453	0.1357596	0.0255157	0.88585
β_1	1000	0.309542	0.1305862	-0.1173529	0.7485971
β_2	1000	0.1855295	0.2008234	-0.4614204	0.7620929
γ_0	1000	-0.2149397	0.2571214	-1.171198	0.4858699
γ_1	1000	2.005918	0.2029011	1.427884	2.745974
γ_2	1000	-0.047537	0.3450838	-1.152159	1.181995
$\ln \alpha$	1000	-1.425232	0.4779846	-4.73018	-0.1171445

Table A.57 Estimated Power of Test for $\alpha = 0$ for Zero-Inflated Overdispersed Data by Sample Size

	$n = 30$	$n = 100$	$n = 1000$
power	0.061	0.152	0.942

Table A.58 Proportion of 95% Confidence Intervals Not Containing True Coefficient Value for Zero-Inflated Negative Binomial Regression on Zero-Inflated Overdispersed Data

	$n = 30$	$n = 100$	$n = 1000$
β_0	0.163	0.172	0.450
β_1	0.243	0.062	0.103
β_2	0.099	0.083	0.063

Table A.59 Summary Table for Zero-Inflated Poisson Regression on Zero-Inflated Overdispersed Data $n = (30)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
β_0	799	0.4983656	1.172824	-8.30338	7.195813
β_1	799	-36.66717	578.4792	-13275.25	3.997535
β_2	799	0.0595742	2.206714	-32.49429	9.625678
γ_0	799	-7.4149	79.04917	-852.5837	680.128
γ_1	799	6.890093	56.42241	-458.6639	799.5233
γ_2	799	3.389028	105.9378	-935.5638	997.2407

Table A.60 Summary Table for Zero-Inflated Poisson Regression on Zero-Inflated Overdispersed Data $n = (100)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
β_0	998	0.5857607	0.4138963	-1.224743	1.874372
β_1	998	0.2042048	0.5778218	-4.35569	2.236064
β_2	998	0.1718985	0.6667866	-2.685396	2.371947
γ_0	998	-0.0233368	0.7505241	-6.310448	2.592301
γ_1	998	1.87766	1.396938	-15.04947	7.501427
γ_2	998	-0.0530051	1.184981	-8.443797	7.163661

Table A.61 Summary Table for Zero-Inflated Poisson Regression on Zero-Inflated Overdispersed Data $n = (1000)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
β_0	1000	0.6192273	0.1195583	0.2575465	0.921964
β_1	1000	0.2850461	0.1186082	-0.0982816	0.6647238
β_2	1000	0.1686814	0.1875641	-0.4554318	0.7471405
γ_0	1000	0.0545556	0.1935774	-0.5514506	0.5925385
γ_1	1000	1.861867	0.1846711	1.301673	2.398921
γ_2	1000	-0.070469	0.3052671	-1.009377	0.9626796

Table A.62 Proportion of 95% Confidence Intervals Not Containing True Coefficient Value for Zero-Inflated Poisson Regression on Zero-Inflated Overdispersed Data

	$n = 30$	$n = 100$	$n = 1000$
β_0	0.183	0.293	0.915
β_1	0.141	0.081	0.242
β_2	0.101	0.095	0.126

A.6 Zero-Inflated Underdispersed Data Simulation Tables

Table A.63 Summary Table for Zero-Inflated Double Poisson Regression on Zero-Inflated Underdispersed Data $n = (30)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	814	2.309645	2.992807	0	22.31187
β_0	814	0.2887323	0.7163162	-8.261283	2.549531
β_1	814	0.2392639	0.7385209	-3.404855	3.291436
β_2	814	0.1621271	0.982643	-4.100286	9.687471
γ_0	814	-13.10087	86.89228	-979.1992	243.7704
γ_1	814	19.72557	85.37293	-209.833	918.7481
γ_2	814	-1.819478	132.7113	-860.77331	988.1268
ϕ	814	15.83195	103.6352	0.1046322	2666.534

Table A.64 Summary Table for Zero-Inflated Double Poisson Regression on Zero-Inflated Underdispersed Data $n = (100)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	971	2.315337	2.836186	0	22.57828
β_0	971	0.2431836	0.3323262	-1.198493	1.306455
β_1	971	0.3659784	0.3219345	-2.048198	1.312403
β_2	971	0.22655	0.4505751	-1.865561	1.705456
γ_0	971	-2.033651	4.470884	-24.1934	1.648847
γ_1	971	3.791049	4.416231	-10.63267	22.0744
γ_2	971	-0.081105	1.573195	-15.26265	7.830997
ϕ	971	1.993873	1.338398	0.4565765	30.52935

Table A.65 Summary Table for Zero-Inflated Double Poisson Regression on Zero-Inflated Underdispersed Data $n = (1000)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	1000	6.811821	4.885725	0.0006023	27.0165
β_0	1000	0.2476349	0.1016913	-0.0960578	0.5384977
β_1	1000	0.4082345	0.0884915	0.1445086	0.7663254
β_2	1000	0.2457806	0.1342592	-0.2076208	0.6763009
γ_0	1000	-0.6426249	0.4529513	-11.4965	0.139475
γ_1	1000	2.295038	0.4415634	1.725331	13.83176
γ_2	1000	0.0032312	0.3290151	-1.77265	0.9792663
ϕ	1000	1.527222	0.2297083	0.7670832	2.259559

Table A.66 Estimated Power of Test for $\phi = 1$ (ZIDP) for Zero-Inflated Underdispersed Data ($\phi = 1.5$) by Sample Size

	$n = 30$	$n = 100$	$n = 1000$
power	0.345	0.477	1.000

Table A.67 Proportion of 95% Confidence Intervals Not Containing True Coefficient Value for Zero-Inflated Double Poisson Regression on Zero-Inflated Undererdispersed Data

	$n = 30$	$n = 100$	$n = 1000$
β_0	0.154	0.134	0.052
β_1	0.136	0.061	0.055
β_2	0.098	0.065	0.049

Table A.68 Summary Table for Zero-Inflated Marginalized Double Poisson Regression on Zero-Inflated Underdispersed Data $n = (30)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	569	2.664401	2.986046	0	21.98856
β_0	569	-0.1782046	0.629007	-4.660252	1.165942
β_1	569	-0.9098232	0.698114	-3.027589	1.207218
β_2	569	0.2006247	1.002848	-2.519929	6.038893
γ_0	569	-2.272942	5.482098	-38.4613	8.158757
γ_1	569	3.66158	4.919977	-14.41459	22.59969
γ_2	569	-0.0008668	3.385123	-25.522252	24.10665
ϕ	569	5.393133	10.88438	0.1274091	137.8333

Table A.69 Summary Table for Zero-Inflated Marginalized Double Poisson Regression on Zero-Inflated Underdispersed Data $n = (100)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	861	2.458371	3.159909	0	22.19787
β_0	861	-0.2082199	0.3138351	-1.66119	0.5435619
β_1	861	-1.013116	0.4203697	-2.644108	-0.00593
β_2	861	0.2463699	0.5260561	-1.299841	2.009291
γ_0	861	-2.26909	4.67438	-24.36633	1.372494
γ_1	861	3.929263	4.608943	0.5700393	26.48319
γ_2	861	-0.0001736	1.542478	-10.66527	9.298804
ϕ	861	2.032437	1.014404	0.3695808	7.871971

Table A.70 Summary Table for Zero-Inflated Marginalized Double Poisson Regression on Zero-Inflated Underdispersed Data $n = (1000)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	811	7.97494926	4.73217	0	26.60237
β_0	811	-0.1911103	0.0950854	-0.4462402	0.0673363
β_1	811	-0.9722466	0.1273112	-1.389485	-0.5624285
β_2	811	0.2487429	0.1580862	-0.2685802	0.6641619
γ_0	811	-0.5606924	0.2207503	-1.369603	0.0849186
γ_1	811	2.215329	0.193657	1.725976	2.907735
γ_2	811	-0.0033563	0.2620449	-0.7548783	0.7981545
ϕ	811	1.595005	0.1966327	1.056722	2.258638

Table A.71 Estimated Power of Test for $\phi = 1$ (ZIMDP) for Zero-Inflated Underdispersed Data ($\phi = 1.5$) by Sample Size

	$n = 30$	$n = 100$	$n = 1000$
power	0.381	0.827	1.000

Table A.72 Proportion of 95% Confidence Intervals Not Containing True Coefficient Value for Zero-Inflated Double Poisson Regression on Zero-Inflated Underdispersed Data

	$n = 30$	$n = 100$	$n = 1000$
β_0	0.098	0.272	1.000
β_1	0.395	0.965	1.000
β_2	0.077	0.064	0.043

Table A.73 Summary Table for Zero-Inflated Negative Binomial Regression on Zero-Inflated Underdispersed Data $n = (30)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	700	0.0555644	0.3643673	0	4.523717
β_0	700	0.0305443	0.7542935	-4.038751	2.479473
β_1	700	-232.5248	5287.229	-139426.6	2.97741
β_2	700	0.2470394	1.242072	-3.868982	5.103167
γ_0	700	-32.90418	128.945	-883.1462	347.7047
γ_1	700	34.01907	102.2508	-240.78	741.1502
γ_2	700	10.08089	195.7164	-1177.117	1094.23
α	700	0.0204393	0.1177949	0	2.100634

Table A.74 Summary Table for Zero-Inflated Negative Binomial Regression on Zero-Inflated Underdispersed Data $n = (100)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	912	0.040989	0.2309539	0	3.002298
β_0	912	0.0184679	0.3505224	-1.161912	1.322203
β_1	912	0.4646196	0.3859305	-2.848041	1.371381
β_2	912	0.2816364	0.5518676	-1.89898	2.408487
γ_0	912	-3.971568	22.06296	-652.6793	1.669625
γ_1	912	5.44318	20.96691	-13.06187	620.6774
γ_2	912	0.8568851	21.12934	-15.32044	634.7809
α	912	0.0151542	0.0630906	0	0.6060215

Table A.75 Summary Table for Zero-Inflated Negative Binomial Regression on Zero-Inflated Underdispersed Data $n = (1000)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	369	0.0013854	0.0212171	0	0.3992475
β_0	369	0.0403438	0.1011167	-0.2169793	0.332651
β_1	369	0.508245	0.0977532	0.2003301	0.7758361
β_2	369	0.3019847	0.1545108	-0.1707408	0.6972074
γ_0	369	-1.421592	0.3238438	-2.353234	-0.4283856
γ_1	369	2.875895	0.2819762	2.144025	3.799212
γ_2	369	0.1598667	0.4117543	-1.818847	1.296522
α	369	0.0002305	0.0027613	0.0000000	0.0490183

Table A.76 Estimated Power of Test for $\alpha = 0$ for Zero-Inflated Underdispersed Data by Sample Size

	$n = 30$	$n = 100$	$n = 1000$
power	0.193	0.380	0.997

Table A.77 Proportion of 95% Confidence Intervals Not Containing True Coefficient Value for Zero-Inflated Negative Binomial Regression on Zero-Inflated Underdispersed Data

	$n = 30$	$n = 100$	$n = 1000$
β_0	0.051	0.078	0.488
β_1	0.093	0.027	0.117
β_2	0.050	0.049	0.051

Table A.78 Summary Table for Zero-Inflated Poisson Regression on Zero-Inflated Underdispersed Data $n = (30)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
β_0	762	0.0045865	0.7696562	-4.038752	2.479473
β_1	762	-1277.135	34168.05	-943076.3	2.977411
β_2	762	0.2882669	1.283963	-3.868047	5.103214
γ_0	762	-30.86574	128.485	-1129.057	243.7741
γ_1	762	31.83491	102.9453	-283.0776	855.2153
γ_2	762	11.16204	189.5885	-1324.598	1315.663

Table A.79 Summary Table for Zero-Inflated Poisson Regression on Zero-Inflated Underdispersed Data $n = (100)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
β_0	971	0.0265464	0.3519426	-1.161911	1.310109
β_1	971	0.4614552	0.3971816	-2.848219	1.371381
β_2	971	0.284251	0.5530689	-1.856019	2.466667
γ_0	971	-3.456722	21.02066	-644.212	1.669625
γ_1	971	4.944292	20.00103	-14.61213	612.6532
γ_2	971	0.7767993	20.1281	-19.45194	623.5672

Table A.80 Summary Table for Zero-Inflated Poisson Regression on Zero-Inflated Underdispersed Data $n = (1000)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
β_0	1000	0.0476455	0.1002426	-0.2496546	0.332648
β_1	1000	0.5017133	0.1001709	0.1881921	0.7860225
β_2	1000	0.2871264	0.156743	-0.2729485	0.6970015
γ_0	1000	-1.407472	0.3391572	-2.508697	-0.4283855
γ_1	1000	2.88108	0.295877	2.134593	4.088398
γ_2	1000	0.133703	0.4151311	-1.81862	1.337481

Table A.81 Proportion of 95% Confidence Intervals Not Containing True Coefficient Value for Zero-Inflated Poisson Regression on Zero-Inflated Underdispersed Data

	$n = 30$	$n = 100$	$n = 1000$
β_0	0.058	0.080	0.453
β_1	0.112	0.033	0.107
β_2	0.060	0.049	0.046

A.7 Zero-Inflated Equidispersed Data Simulation Tables

Table A.82 Summary Table for Zero-Inflated Double Poisson Regression on Zero-Inflated Equidispersed Data $n = (30)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	865	2.198325	2.806806	0	17.68703
β_0	865	0.1637019	1.418418	-27.06942	3.778333
β_1	865	0.194739	1.105758	-5.494853	9.776934
β_2	865	0.2672161	1.664783	-10.03173	17.08929
γ_0	865	-23.89318	122.0419	-990.4853	407.9695
γ_1	865	31.43477	109.001	-200.7866	928.6336
γ_2	865	-2.140277	211.7555	-1197.792	1224.58
$\ln \phi$	865	0.7559685	1.312486	-4.203314	8.408788

Table A.83 Summary Table for Zero-Inflated Double Poisson Regression on Zero-Inflated Equidispersed Data $n = (100)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	1000	1.38113	1.86622	0.0000000	16.54779
β_0	1000	0.215969	0.5014021	-2.815358	1.344093
β_1	1000	0.3475096	0.4625878	-2.887919	2.437807
β_2	1000	0.2216713	0.647476	-2.609328	0.9166827
γ_0	1000	-5.496268	47.77197	-755.3585	52.92838
γ_1	1000	6.425895	28.31859	-12.26355	397.6176
γ_2	1000	1.342896	65.31023	-739.2885	874.5712
$\ln \phi$	1000	0.1808797	0.5570817	-2.037059	3.378365

Table A.84 Summary Table for Zero-Inflated Double Poisson Regression on Zero-Inflated Equidispersed Data $n = (1000)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	1000	1.042161	1.435652	0.0000000	12.15031
β_0	1000	0.2435707	0.1471744	-0.4610452	0.6105229
β_1	1000	0.4072516	0.1159599	0.0549916	0.8752079
β_2	1000	0.2461351	0.1764856	-0.3840854	0.8193016
γ_0	1000	-0.6386518	0.3861874	-2.717016	0.1749864
γ_1	1000	2.285794	0.3159541	1.60098	4.94845
γ_2	1000	0.0085614	0.3770795	-1.70056	1.277595
$\ln \phi$	1000	0.0133284	0.180894	-0.5745605	0.5640699

Table A.85 Estimated Power of Test for $\phi = 1$ (ZIDP) for Zero-Inflated Equidispersed Data ($\phi = 1$) by Sample Size

	$n = 30$	$n = 100$	$n = 1000$
power	0.190	0.090	0.051

Table A.86 Proportion of 95% Confidence Intervals Not Containing True Coefficient Value for Zero-Inflated Double Poisson Regression on Zero-Inflated Equierdispersed Data

	$n = 30$	$n = 100$	$n = 1000$
β_0	0.194	0.148	0.055
β_1	0.175	0.074	0.054
β_2	0.147	0.080	0.052

Table A.87 Summary Table for Zero-Inflated Marginalized Double Poisson Regression on Zero-Inflated Equidispersed Data $n = (30)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	596	2.674445	3.454039	0	19.92291
β_0	596	-0.359515	1.479807	-22.789	2.106761
β_1	596	-0.8451682	0.931639	-4.589203	8.256331
β_2	596	0.2269025	1.284803	-8.710286	7.332118
γ_0	596	-3.049758	8.877381	-94.13109	16.85328
γ_1	596	2.298788	25.47982	-516.8484	91.16903
γ_2	596	0.4261391	6.015723	-49.29602	69.77798
$\ln \phi$	596	1.112853	1.615508	-4.010528	9.035553

Table A.88 Summary Table for Zero-Inflated Marginalized Double Poisson Regression on Zero-Inflated Equidispersed Data $n = (100)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	979	1.323259	2.042823	0	16.75977
β_0	979	-0.2549354	0.3930702	-2.262858	0.702777
β_1	979	-1.044422	0.5039631	-3.291687	0.5436941
β_2	979	0.2494317	0.6457823	-3.70103	2.934878
γ_0	979	-2.0423	4.858545	-52.33106	1.588178
γ_1	979	3.591073	4.466294	-27.04757	37.94584
γ_2	979	0.0847905	2.60205	-16.58895	54.64605
$\ln \phi$	979	0.2071104	0.5406726	-1.937616	2.278199

Table A.89 Summary Table for Zero-Inflated Marginalized Double Poisson Regression on Zero-Inflated Equidispersed Data $n = (1000)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	1000	1.005071	1.644315	-5.994286	12.30609
β_0	1000	-0.1941309	0.1100534	-0.5451792	0.1258163
β_1	1000	-0.9818405	0.1428507	-1.414079	-0.5897589
β_2	1000	0.2456174	0.1770326	-0.3734551	0.7349361
γ_0	1000	-0.6300764	0.3679389	-2.462905	0.1649693
γ_1	1000	2.27883	0.3099801	1.572749	4.228553
γ_2	1000	0.0023828	0.3216856	-1.057091	1.10575
$\ln \phi$	1000	0.0155452	0.1797121	-0.703139	0.5651565

Table A.90 Estimated Power of Test for $\phi = 1$ (ZIMDP) for Zero-Inflated Equidispersed Data ($\phi = 1$) by Sample Size

	$n = 30$	$n = 100$	$n = 1000$
power	0.237	0.103	0.058

Table A.91 Proportion of 95% Confidence Intervals Not Containing True Coefficient Value for Zero-Inflated Double Poisson Regression on Zero-Inflated Underdispersed Data

	$n = 30$	$n = 100$	$n = 1000$
β_0	0.128	0.206	0.987
β_1	0.337	0.897	1.000
β_2	0.097	0.062	0.044

Table A.92 Summary Table for Zero-Inflated Negative Binomial Regression on Zero-Inflated Equidispersed Data $n = (30)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	999	0.3430899	1.005143	0	8.190478
β_0	999	0.1102255	0.90277	-4.7523	6.878152
β_1	999	-818.6724	21039.56	-650362.3	3.730715
β_2	999	0.2182893	1.515883	-10.49579	5.868969
γ_0	999	-41.8345	181.8069	-1033.212	903.9393
γ_1	999	61.575	162.2408	-634.6985	900.0986
γ_2	999	-1.148071	306.6287	-1367.465	1358.924
$\ln \alpha$	999	-70.07806	463.8605	-12769.59	1.522146

Table A.93 Summary Table for Zero-Inflated Negative Binomial Regression on Zero-Inflated Equidispersed Data $n = (100)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	1000	0.3194701	0.9896302	0	10.70889
β_0	1000	0.167725	0.4071276	-1.558262	1.279037
β_1	1000	0.3712495	0.454315	-2.829159	1.780212
β_2	1000	0.237648	0.6358636	-1.770516	3.069723
γ_0	1000	-4.260324	35.77535	-716.1304	15.1207
γ_1	1000	5.386305	22.13186	-12.01218	372.4251
γ_2	1000	1.056847	45.4132	-866.0796	746.572
$\ln \alpha$	1000	-59.02715	709.2648	-22102.26	0.6472901

Table A.94 Summary Table for Zero-Inflated Negative Binomial Regression on Zero-Inflated Equidispersed Data $n = (1000)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
χ_0^2	1000	0.3852798	0.9238408	0	6.954323
β_0	1000	0.2326946	0.1139971	-0.145539	0.5736641
β_1	1000	0.4102844	0.1100898	0.0661271	0.7596129
β_2	1000	0.2486717	0.1745261	-0.3126534	0.7640797
γ_0	1000	-0.6642127	0.2609197	-1.557955	0.0750959
γ_1	1000	2.297204	0.2192825	1.664365	3.0747
γ_2	1000	0.0175866	0.3713317	-1.608321	1.255453
$\ln \alpha$	1000	-20.05531	53.36124	-1249.06	-1.448466

Table A.95 Estimated Power of Test for $\alpha = 0$ for Zero-Inflated Equidispersed Data by Sample Size

	$n = 30$	$n = 100$	$n = 1000$
power	0.022	0.021	0.016

Table A.96 Proportion of 95% Confidence Intervals Not Containing True Coefficient Value for Zero-Inflated Negative Binomial Regression on Zero-Inflated Equidispersed Data

	$n = 30$	$n = 100$	$n = 1000$
β_0	0.088	0.082	0.035
β_1	0.209	0.045	0.051
β_2	0.79	0.064	0.050

Table A.97 Summary Table for Zero-Inflated Poisson Regression on Zero-Inflated Equidispersed Data $n = (30)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
β_0	809	0.1336022	0.99104426	-8.303386	6.878153
β_1	809	-1495.383	34356.33	-943076.3	3.730804
β_2	809	0.2548731	1.634194	-10.49579	9.625678
γ_0	809	-17.963332	101.9972	-953.6006	410.7667
γ_1	809	21.42698	87.20069	-283.0776	810.6756
γ_2	809	3.048146	163.8066	-1324.598	1174.367

Table A.98 Summary Table for Zero-Inflated Poisson Regression on Zero-Inflated Equidispersed Data $n = (100)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
β_0	998	0.2186108	0.3973711	-1.639569	1.821469
β_1	998	0.3527267	0.5034682	-3.998435	0.7413688
β_2	998	0.2379138	0.6188039	-2.960205	2.689855
γ_0	998	-0.9862194	2.138616	-37.92474	3.336826
γ_1	998	2.634404	2.371497	-13.23346	36.21062
γ_2	998	-0.0017245	2.134505	-20.29166	39.70543

Table A.99 Summary Table for Zero-Inflated Poisson Regression on Zero-Inflated Equidispersed Data $n = (1000)$

Variable	Successful Simulations	Mean	Standard Deviation	Minimum	Maximum
β_0	1000	0.2492326	0.1126308	-0.121473	0.5440814
β_1	1000	0.405761	0.107491	0.0770238	0.6991494
β_2	1000	0.2454314	0.1731714	-0.2589368	0.7610502
γ_0	1000	-0.61936964	0.2490819	-1.378026	0.1287695
γ_1	1000	2.271756	0.2080695	1.1717783	2.894234
γ_2	1000	0.0073199	0.3619541	-1.571762	1.129371

Table A.100 Proportion of 95% Confidence Intervals Not Containing True Coefficient Value for Zero-Inflated Poisson Regression on Zero-Inflated Equidispersed Data

	$n = 30$	$n = 100$	$n = 1000$
β_0	0.110	0.071	0.056
β_1	0.141	0.045	0.051
β_2	0.094	0.062	0.050