

8-1-2013

Double truncated poisson regression model with random effects

Abdalhalim M. Ali Suaiee

Follow this and additional works at: <http://digscholarship.unco.edu/dissertations>

Recommended Citation

Suaiee, Abdalhalim M. Ali, "Double truncated poisson regression model with random effects" (2013). *Dissertations*. Paper 260.

This Text is brought to you for free and open access by the Student Research at Scholarship & Creative Works @ Digital UNC. It has been accepted for inclusion in Dissertations by an authorized administrator of Scholarship & Creative Works @ Digital UNC. For more information, please contact Jane.Monson@unco.edu.

UNIVERSITY OF NORTHERN COLORADO

Greeley, Colorado

The Graduate School

DOUBLE TRUNCATED POISSON REGRESSION MODEL
WITH RANDOM EFFECTS

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy

Abdalhalim M. Ali Suaiee

College of Education and Behavioral Sciences
School of Educational Research, Leadership, and Technology
Department of Applied Statistics and Research Methods

August 2013

This Dissertation by: Abdalhalim M. Ali Suaiee

Entitled: DOUBLE TRUNCATED POISSON REGRESSION MODEL WITH
RANDOM EFFECTS

has been approved as meeting the requirement for the Degree of Doctor of
Philosophy in the College of Education and Behavioral sciences in the School of
Educational Research, Leadership, and Technology.

Accepted by the Doctoral Committee

Jay Schaffer Ph.D., Co-Chair

Trent Lalonde Ph.D., Co-Chair

Robert Pearson Ph.D., Committee Member

Robert Heiny Ph.D., Faculty Representative

Date of Dissertation Defense _____

Accepted by the Graduate School

Linda L. Black, Ed.D., LPC
Acting Dean of the Graduate School and International Admissions

ABSTRACT

Suaiee, Abdalhalim M. Ali. Double Truncated Poisson Regression Model with Random Effects. Published Doctor of Philosophy dissertation, University of Northern Colorado, 2013.

Count data regression models are used for special cases where the response variable takes count values or only non-negative values. Poisson regression models are commonly used to analyze count data. A frequent problem with the use of these models is that the observed variation is greater than expected and mixed Poisson models are alternative models that provide a means of explaining the extra-Poisson variation. Mixed Poisson regression models have extensive research and literature studies, and have been commonly used in fields such as epidemiology, medicine, genetics, economics, engineering, marketing, and in the physical and social sciences. However, in many cases, the analyst does not observe the entire distribution of counts. In such a case, the count data are truncated as the data are observed only over part of the range of the response variable. In this study, we formulate a class of regression models based on a Double Truncated Poisson regression model with random effects. Two different distributions for the random effects, Normal and Gamma, were studied through simulation. Misspecification of these distributions was addressed. Comparisons with the Left Truncated Mixed Poisson model and the regular Mixed Poisson model were presented. It was concluded that with Normal random effects, double and Left Truncated Mixed Poisson models provide a better fit to clustered double truncated count data compared to the regular mixed Poisson model. For Gamma random effects, the Double Truncated Mixed Poisson model

provides a better fit to clustered double truncated count data. These models were used to analyze a Transitional Housing Facility data set.

ACKNOWLEDGMENTS

First and foremost, all praise be to Allah, the Almighty, the Benevolent for His blessings and guidance for giving me the inspiration to embark on this dissertation and instilling in me the strength to see that this dissertation becomes a reality.

I would like to say a big thank you to my amazing supervisor, Dr. Trent Lalonde for believing in me to finish up this dissertation. His advice, encouragement and help mean a lot to me. So I can never thank him enough. The same thanks go to the rest of my supervisory committee, Dr. Jay Schaffer, Dr. Robert Pearson, and Dr. Robert Heiny for their careful reading of the dissertation and helpful suggestions.

I want also to thank Mr. Alan and Mrs. Cathy Hendrickson for their friendship and helping me in editing my writing of this dissertation.

The most significant people to me in the completion of this dissertation are my family members. I thank my parents for their encouragement and care. The thanks also go to my brothers and sisters.

I would like to thank my two wonderful sons, Mosab and Sohib, and my beautiful daughter, Arwa, for filling my life with joy and happiness.

Last, but not least, I am extremely thankful to my beloved wife, Karema. I would not be the successful man I am without you. Thank you for your unlimited support, sacrifice, and encouragement. I love you and I plan to make it up to you.

Table of Contents

Chapter I: INTRODUCTION	1
Objectives Of The Study	7
Research Questions	7
Chapter II: LITERATURE REVIEW	10
Count Models	10
Poisson Regression Model	11
Negative Binomial Regression	13
Truncated Count Data	16
Left Truncated Poisson Model at $k = 0$	16
Left Truncated Poisson Model at $k = c_l$	20
Right Truncated Poisson Model at $k = c_r$	23
Double Truncated Poisson Model	27
Linear Mixed Model	30
Count Model With Random Effect	32
Parameter Estimation	34
Maximum Likelihood	35
Penalized Quasilikelihood	36
Hierarchical Likelihood	39
Truncated Mixed Count Models	40
Chapter III: DOUBLE TRUNCATED MIXED POISSON REGRESSION MODEL	44
Introduction	44
The Double Truncated Mixed Poisson Regression Model	45
H-likelihood Estimation for Normal Random Effect	47
H-likelihood Estimation for Gamma Random Effect	50
Chapter IV: SIMULATION STUDY	56
Simulation Study for the Double Truncated Mixed Poisson Regression Data with Normal Random Effects	57
Simulation Study for the Double Truncated Mixed Poisson Regression Data with Gamma Random Effects	63
Simulation Study for the Double Truncated Mixed Poisson Regression Data with Mis-specified Normal Random Effects	69

Simulation Study for the Double Truncated Mixed Poisson Regression Data with Mis-specified Gamma Random Effects	76
Real Data Example	84
Chapter V: CONCLUSION AND FUTURE RESEARCH	89
Conclusion	89
Future Research	92
REFERENCES	94
Appendices	103
A: Histograms of the Estimated Random Effects	103
B: R Programs	116
C: Transitional Housing Facility Data Set	156

List of Tables

1	Generated Double Truncated with Normal Random Effect	54
2	Generated Double Truncated with Gamma Random Effect	54
3	Double Truncated Poisson Regression Model With Normal Random Effects	59
4	Power and Type I Error Rate for Models with Normal Random Effects	60
5	Double Truncated Poisson Regression Model With Gamma Random Effects	65
6	Power and Type I Error Rate for Models with Gamma Random Effects	66
7	Double Truncated Poisson Regression Model With Miss-Specified Normal Random Effects	70
8	Power and Type I Error Rate for Models with Misspecified Normal Random Effects	71
9	Double Truncated Poisson Regression Model With Mis-Specified Gamma Random Effects	78
10	Power and Type I Error Rate for Models with Misspecified Gamma Random Effects	79
11	Significant Predictors of the Response Variable with Normal Random Effects	86
12	Significant Predictors of the Response Variable with Gamma Random Effects	88

List of Figures

1	Power of Double Truncated Poisson Model with Normal Random Effects	61
2	Type I rate of Double Truncated Poisson Model with Normal Random Effects	62
3	Histogram of the estimated Random effect	62
4	Power of Double Truncated Poisson Model with Gamma Random Effects	67
5	Type I rate of Double Truncated Poisson Model with Gamma Random Effects	68
6	Histogram of the estimated Random effect	68
7	Power of Double Truncated Poisson Model with Miss-specified Normal Random Effects	72
8	Type I rate of Double Truncated Poisson Model with Miss-specified Normal Random Effects	73
9	Power of Double Truncated Poisson Model with Mis-specified Normal Random Effects	74
10	Type I rate of Double Truncated Poisson Model with Mis-specified Normal Random Effects	75
11	Histogram of the estimated Mis-specified Random effect	76
12	Power of Double Truncated Poisson Model with Mis-specified Gamma Random Effects	80

13	Type I rate of Double Truncated Poisson Model with Mis-specified Gamma Random Effects	81
14	Power of Double Truncated Poisson Model with Mis-specified Gamma Random Effects	82
15	Type I rate of Double Truncated Poisson Model with Mis-specified Gamma Random Effects	83
16	Histogram of the estimated Miss-specified Random effect	83

CHAPTER I

INTRODUCTION

Count data regression models are used for special cases in which the response variable takes count values or only non-negative values. Count regression models are commonly used to measures of health utilization such as the number of doctor visits, with explanatory variables including various socio-economic variables, health status, and type of health insurance (Cameron and Sohansson, 1997). Important early developments in count models, have been used in actuarial science, biostatistics, and demography. In the last 10 years these models also extensively took a place in economics, political science, and sociology. The special features of data in their respective fields of application have driven developments that have increased the choice of these models.

In many situations of practical interest the response variable in an experiment or observational study is a count that is assumed to follow the Poisson distribution. An important landmark in the development of the count data regression model was the appearance of the generalized linear models, of which the Poisson regression is a special case. In another case, an event may be thought of as the realization of a point process controlled by some specified rate of occurrence of the event. The number of events may be described as the total number of such realizations over

some unit of time. Count data regression is beneficial in studying the occurrence rate per unit of time conditional on some covariates. Cameron and Trivedi (1998) described that the approach taken to the analysis of count data sometimes depends on how the counts are assumed to arise. Count data can arise from two common ways:

- Counts arise from a direct observation of a point process.
- Counts arise from discretization of continuous latent data.

In the first case, examples are the number of telephone calls arriving at a central telephone exchange, the number of monthly absences at a workplace, the number of airline accidents, the number of hospital admissions, and so forth. The data may also consist of inter-arrival times for events. In the second case, consider the following example. Credit rating of agencies may be stated as AAA, AAB, AA, A, BBB, B, and so forth, where AAA indicates the greatest credit. Suppose one codes these as $y = 0, 1, \dots, m$. These are pseudocounts that can be analyzed using a count regression. But one may also regard this as an ordinal ranking that can be modeled using a suitable latent variable model such as ordinal multinomial.

Typically, a characteristic of count data is that the counts occur over some fixed area or observation period. Count data, even though numeric, can create some problems if analyzed using the regular linear regression because only nonnegative integer values can occur. Thus, count data can potentially result in a highly skewed distribution, one that is cut off at zero. Therefore, it is often unreasonable to

assume that the response variable and the resulting errors have a normal distribution, making linear regression a less appropriate option for analysis. A suitable way to deal with count data is to use the Poisson distribution and log link function in the analysis. The regression model that uses these kinds of options is called the Poisson regression or the Poisson log-linear regression model.

The natural stochastic model for counts is a Poisson distribution for the number of occurrences of the event, with density

$$P(Y_i = y_i) = \frac{e^{-\lambda_i} \lambda_i^{y_i}}{y_i!}, \quad y_i = 0, 1, 2, \dots$$

where λ_i is the intensity or rate parameter. The mean and variance are

$$E[Y_i] = \lambda_i,$$

$$V(Y_i) = \lambda_i.$$

This shows the equality of mean and variance property of the Poisson distribution.

The Poisson regression model is derived from the Poisson distribution by parameterizing the relationship between the mean parameter λ_i and covariates \mathbf{x}_i .

The standard assumption is to use the exponential mean parameterization,

$$\lambda_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta}), \quad i = 1, \dots, n,$$

where by assumption there are k linearly independent covariates, usually including a constant for the intercept.

In some studies, inclusion in the sample requires that sampled individuals have been engaged in the activity of interest. In many cases, “the analyst does not observe the entire distribution of counts. In particular the zeros often are not observed” (Grogger and Carson, 1991). In such a case, the count data are truncated where the data will be observed only over part of the range of the dependent variable. Moreover, the prevalence of this form of truncation may be attributed to the tendency of an observational apparatus to become active only when at least one event occurs (Johnson and Kotz, 1969). For example, the size of groups observed in public places (Coleman and James, 1961), the number of accidents per worker in a factory (David and Johnson, 1952), or the number of visits to a hospital, could be described by a zero-truncated or a positive Poisson model. More examples of truncated counts include the number of bus trips made per week in surveys taken on buses, the number of shopping trips made by individuals sampled at a mall, and the number of unemployment spells among a pool of unemployed. In all these cases, the data are generally said to be left-truncated. Right truncation happens from loss of observations greater than some specified value. For example, if we want to model the number of times (per week) an in-vehicle navigation system is used on the morning commute to work during weekdays, the data are right truncated at 5, which is the maximum number of uses in any given week.

Yet another possible distinction from the regular Poisson model is when the data generating process excludes the observation of count outcomes outside of a specific interval. For example, observations on the number of items purchased when

there is a limit per customer would exclude the zero class, since customers who do not buy the item will not be identified, and would exclude observations above the limit. Similarly, the number of serious criminal offences committed over a certain time period by felons under California's three strikes law, is characteristic of a double truncation. Data of this type might be considered to follow a doubly truncated Poisson distribution. Problems of truncation in a Poisson distribution, of one form or another, had early consideration by Bliss (1948), David and Johnson (1952), Moore (1952, 1954), Rider (1953), and Plackett (1953). However, it was not until Cohen (1954), that various forms of truncation were concisely generalized into the doubly truncated Poisson model.

The doubly truncated Poisson model encompasses singly-left or right truncation, as well as the regular Poisson model. However, the most frequent truncation phenomenon to arise is where the Poisson distribution otherwise appropriately describes the data generating process except for the omission of the zero-class, which is a special case of the doubly truncated Poisson distribution. Unless the likelihood function is suitably modified, truncation leads to inconsistent parameter estimates and the bias of estimators will be severe (Cameron and Trivedi, 2001).

Count data has been extensively used for the analysis of discrete data and can show overdispersion compared to the Poisson distribution. For example, for counts of the number of epileptic seizures, there is a very large individual variation in the seizures rate (Hougaard, Lee, and Whitmore, 1997). A frequent problem with the

use of these models is that the observed variation is greater than expected which can lead to an overdispersion. Overdispersion can be associated with parameter heterogeneity, and one way to take care of heterogeneity is by way of mixed models (Mullahy, 1997).

Mixed Poisson regression models have attracted a great deal of attention in the last 10 years. Recently, Karlis and Xekalaki (2005) have addressed the concern of the Poisson Mixed Model. Mixed Poisson models are alternative models that provide a means of explaining the extra-Poisson variation. The term mixed model refers to the use of both fixed and random effects in the same analysis. Mixed Poisson distributions have been used widely in scientific fields for modeling non-homogeneous populations. Mixed Poisson models can be used to model rare events and have been used for modeling practical applications amongst others in the field of market research, and accidents and sickness. Mixed Poisson model can be used in hospital admissions modeling when the process of events occurring over time is assumed to be a random process for each individual where each individual has his own intensity of event occurrence Woehl (2008).

Very few studies have discussed truncated Poisson regression with mixed effects. Saei and Chambers (2005) described a random components truncated Poisson model that can be applied to clustered and zero-truncated count data. Nasiri(2011) reviewed the application of mixed Poisson models to health care events where the problem of individuals falling into the zero class is discussed and estimates are compared. It is very important to notice that ignoring truncation will

cause biased estimates of the Poisson regression parameters. Also, misspecification of the random effect distribution can lead to severe bias in the random effect component prediction (Litieri, Alonso, and Molenbergs, 2008).

Objectives Of The Study

The objectives of the study are

1. To construct a doubly-truncated mixed effects Poisson regression model.
2. To estimate the parameters of this model by Hierarchical Generalized Linear Models (HGLM) estimation.
3. To carry out a simulation study to
 - study two different distributions for the random effect in the context of the model.
 - Address the misspecification of random effect distribution in the context of the model.
4. To Compare this model to left truncated Poisson regression model with mixed effects and ordinary Poisson regression model with mixed effects.
5. To apply the model to real data.

Research Questions

The following questions are to be addressed in this study:

- Q1 What are the assumptions and systematic model formula for the Double Truncated Poisson regression model with random effects?
- Q2 How will the parameters be estimated for the Double Truncated Poisson regression model with random effects?
- Q3 Is there a difference between using normal random effect and gamma random effect in the context of the model estimation?
- Q4 How will the misspecification of the random effect distribution affect the parameter estimate bias, parameter estimate standard errors, power, and Type I error rate?
- Q5 How is the proposed model going to perform compared to left truncated Poisson mixed model and regular Poisson mixed model?

This study will help researchers realize the use of doubly-truncated Poisson regression with mixed effects in analyzing count data. Besides focusing on parameter estimation, this study will also help to highlight the interpretation of coefficients. This study will help overcome the truncation and overdispersion problems that occur in Poisson regression which, if ignored, may cause underestimation of standard errors (Cox, 1983) and which consequently gives misleading inference about the regression parameters.

This dissertation consists of 5 chapters. Chapter I gives a general outline of the study. It begins with the explanation of count data. This includes the characteristics of count data which is very important throughout the study. Chapter I also gives a brief description of the Poisson regression model, left truncated Poisson regression, right truncated Poisson regression, the double truncated Poisson regression, and mixed Poisson regression model.

Chapter II, provides a review of count data models, the Poisson regression model, and the Negative Binomial regression model. This chapter also discusses the basic idea of truncated count data, left truncated Poisson Regression model, right truncated Poisson Regression model, and the double truncated Poisson Regression model. The main properties include mean, variance and maximum likelihood, described for each case. Also, chapter II provides a brief discussion of the linear mixed model and the count regression model with mixed effects. Mixed effects Poisson regression model with corresponding parameter estimation methods, which include Maximum Likelihood, Penalized Quasilikelihood, and Hierarchical Likelihood are discussed. Existing work on mixed truncated Poisson regression models is discussed

Chapter III gives clear descriptions on the formulation of the doubly-truncated Poisson regression model with mixed effects. The process of mean

model parameter estimation is fully presented as well as the estimation of dispersion components. Further, we present the properties of estimators and the advantages of using this model.

In Chapter IV, a simulation study is conducted to illustrate the model and estimation method. A real data set is used to demonstrate the benefits of using doubly-truncated Poisson regression model with mixed effects.

In chapter V, the conclusions of the study are discussed, including some potential future work topics.

CHAPTER II

LITERATURE REVIEW

Count Models

Count models are appropriate when the dependent variable is a count bound between zero and infinity. It represents the number of events that occur in a given time frame. A count variable counts the number of discrete events of interest that occurs during a time interval. Examples are the occurrence of cancer cases in a hospital during a given period of time, the number of cars that pass through a toll station per day, and phone calls received at a call center. Winkelmann (1995) studied the number of live births over a specified age interval of the mother, where the interest was to analyze the variation in terms of the mother's schooling, age, and household income. Another example of count modeling is studied by Cameron, Trivedi, Milne and Piggott (1988), where they studied the number of times that individuals utilize a health service, such as visits to a doctor or days in the hospital in the past year. The most popular methods to model count data are Poisson and negative binomial regression (Saffari and Adnan, 2011). Poisson regression is the more popular of the two and is applied to various fields. The Poisson model assumes that the mean equals the variance. However, count data often show overdispersion compared to the Poisson distribution. Overdispersion occurs when the variance exceeds the mean. Also, count data can be truncated where some values in a specific range cannot be observed. Therefore, the inclusion of random effects, by accounting for individual differences, may decrease the possibility of overdispersion.

In this chapter, we first review the standard Poisson Regression count model and the negative regression model. Then, we take a closer look at truncated count data dealing with left, right and doubly truncated Poisson count models. We also present the generalized linear mixed model, moving to the count model with random effects and the mixed effects Poisson regression model. Estimation methods are detailed for each data situation. Finally, we present current work on truncated count models with random effect.

Poisson Regression Model

The Poisson regression model is the standard approach to count data analysis. It has three basic assumptions. First the conditional distribution of the dependent variable, second the specification of the mean parameter, and third the independence of the distribution for all observations. When a Poisson Model is appropriate for an outcome Y , the probabilities of observing a specific, y , are given by

$$P(Y = y) = \frac{e^{-\lambda} \lambda^y}{y!},$$

where λ is the population rate parameter. The Poisson random variable Y has the mean $E[Y] = \lambda$, and variance $Var(Y) = \lambda$. The equality of the mean and variance provides a quick check on whether a Poisson model might be appropriate for a sample of observations. For example, if the sample mean is very different from the sample variance, then one probably does not have Poisson data. The Poisson regression model is derived from the Poisson distribution by parameterizing the relation between the mean parameter λ and covariates \mathbf{x} (Cameron and Trivedi, 2001). Suppose that we have a sample of n observations Y_1, \dots, Y_n which can be treated as realizations of independent Poisson random variables, with $Y_i \sim Poi(\lambda_i)$,

and suppose that we want to let the mean λ_i depend on a vector of explanatory variables \mathbf{x}_i . The standard assumption is to use the exponential mean parametrization,

$$\lambda_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta}), \quad i = 1, 2, \dots, n.$$

The distribution of Y will be conditional on the regressors \mathbf{x} , so the conditional distribution is $Y_i | \mathbf{x}_i \sim^{iid} Poi(\lambda_i)$, and the probability mass function is

$$f(Y_i | \lambda_i) = \frac{e^{-\lambda_i} \lambda_i^{y_i}}{y_i!}; \quad y_i = 0, 1, 2, \dots, \quad (1)$$

where $\lambda_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta})$. In this expression, \mathbf{x}_i is a vector of covariates and $\boldsymbol{\beta}$ is a vector of parameters. The conditional mean and variance of the distribution are given by

$$\lambda(\mathbf{X}_i) = E[Y_i | \mathbf{X}_i] = Var(Y_i | \mathbf{X}_i) = \exp(\mathbf{x}_i^T \boldsymbol{\beta}).$$

The coefficients $\boldsymbol{\beta}$ can be interpreted as average proportionate change in $E[Y_i | \mathbf{x}_i]$ for a unit change in \mathbf{x}_i , (Grogger and Carson, 1991). Using (1) and the assumption that the responses $(Y_i | \mathbf{x}_i)$ are independent, the likelihood function is

$$L(\boldsymbol{\beta}; y_i) = \prod_{i=1}^n \frac{e^{-\lambda_i} \lambda_i^{y_i}}{y_i!}.$$

The principle of maximum likelihood states that for the given data, we need to maximize L as a function of the regression parameters. To aid the numerical optimization routine, we use the log-likelihood so that we can work with a sum

instead of a product. Thus, the log-likelihood is

$$\begin{aligned} l(\boldsymbol{\beta}; \mathbf{y}_i) &= \sum_{i=1}^n (y_i \ln \lambda_i - \lambda_i - \ln(y_i!)) \\ &= \sum_{i=1}^n (y_i \mathbf{x}_i^T \boldsymbol{\beta} - \exp(\mathbf{x}_i^T \boldsymbol{\beta}) - \ln(y_i!)) . \end{aligned} \quad (2)$$

To estimate the parameters using the maximum likelihood estimation (MLE) method, we take derivatives of (2) with respect to $\boldsymbol{\beta}$. Setting the derivatives to zero, the first order condition maximum likelihood is

$$\sum_{i=1}^n (y_i - \exp(\mathbf{x}_i^T \boldsymbol{\beta})) \mathbf{x}_i = 0.$$

The advantages of using MLE are that maximum likelihood provides consistent estimators, are asymptotically normally distributed, and are asymptotically minimum variance unbiased estimators as the sample size increases. The most common hypothesis testing follows by using Wald tests.

Negative Binomial Regression

In case of modeling count data, it is usually our first tendency is to use Poisson regression. In practice, we often find Poisson regression model is not appropriate to model count data, though Poisson regression models are often presented as the regular approach for such data. On the other hand, the negative binomial regression model is much more flexible and is therefore likely to fit the count data better, if the data are not Poisson. In addition, Poisson and negative binomial regression models differ in terms of their assumptions of the conditional mean and variance of the response variable.

Poisson models assume that the conditional mean and variance of the distribution are equal. Negative binomial regression models do not assume an equal mean and variance and particularly correct for overdispersion in the data, which is when the variance is greater than the conditional mean (Osgood, 2000). The properties of the negative binomial models can be described as the Poisson-Gamma Mixed Model. The Poisson-Gamma model has properties that are very similar to the Poisson model, in which the response variable Y following a Poisson distribution with a mean λ where an additional model error is assumed to follow a Gamma distribution. The Poisson-Gamma is a mixture of two distributions and was first derived by Greenwood and Yule (1920). This mixture distribution was developed to account for over-dispersion that is commonly observed in discrete or count data (Lord, Washington, and Ivan, 2005). It became very popular because the conjugate distribution has a closed form and leads to the negative binomial distribution.

Suppose that given λ , Y has a Poisson distribution with mean λ , and λ has a gamma distribution, $\Gamma(k, \mu)$. The gamma probability density function for λ is

$$f(\lambda; k, \mu) = \frac{(k/\mu)^k}{\Gamma(k)} \exp(-k\lambda/\mu) \lambda^{k-1}, \quad \lambda \geq 0.$$

This gamma distribution has

$$E[\lambda] = \mu, \quad Var(\lambda) = \frac{\mu^2}{k}.$$

The parameter $k > 0$ describes the shape. The density is skewed to the right, but the degree of skewness decreases as k increases.

Marginally, the gamma mixture of the Poisson distributions yields the negative binomial distribution for Y . Its probability mass function is

$$p(y; k, \mu) = \frac{\Gamma(y+k)}{\Gamma(k)\Gamma(y+1)} \left(\frac{k}{\mu+k} \right)^k \left(1 - \frac{k}{\mu+k} \right)^y, \quad y = 0, 1, 2, \dots$$

This negative binomial distribution has

$$E[Y] = \mu, \quad \text{Var}(Y) = \mu + \mu^2/k.$$

The index k^{-1} is called the dispersion parameter. As $k^{-1} \rightarrow 0$, the gamma distribution has $\text{var}(\lambda) \rightarrow 0$ and it converges to a degenerate distribution at μ . Similarly, the negative binomial distribution then has $\text{var}(\lambda) \rightarrow \mu$ and it converges to the Poisson distribution with mean μ .

Negative binomial models for counts permit μ to depend on explanatory variables (Lawless 1987). Such that

$$\mu_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta}).$$

For k fixed, a negative binomial model is a GLM. Thus, the likelihood equations for the regression parameters $\boldsymbol{\beta}$ are special cases of the likelihood equations for an ordinary GLM with variance function $\text{Var}(\mu) = \mu + \mu^2/k$. The usual iterative reweighted least squares algorithm applies for ML model fitting (Agresti, 2002). When k is unknown, ML fitting can use a Newton-Raphson routine on all the parameters simultaneously.

Truncated Count Data

Count data frequently arise as outcomes of an underlying count process in continuous time (Winkelmann, 2008). When count data are observed only over part of the range of the response variable, then the data are called truncated count data (Cameron and Trivedi, 2001). Examples of truncated data arise in many contexts. The number of accidents per worker in a factory and the number of visits to a hospital can be examples of zero truncated data or generally called left truncated data. On the other hand, if we want to model the number of times (per week) an in-vehicle navigation system is used on the morning commute to work during weekdays, the data are right truncated at 5, which is the maximum number of uses in any given week. Examples of right truncated data may arise when observations greater than some particular value are missed (Cameron and Trivedi, 2001). Left truncation at zero has been known as the most common form of truncation (Winkelmann, 2008). Truncated Poisson and negative binomial models have been discussed by Creel and Loomis (1990), and Grogger and Carson (1991). However, left truncation at value $k = c_l$ and right truncation at value $k = c_r$ can also be the case.

Left Truncated Poisson Model at $k = 0$

Count data in which zero counts cannot be observed are called zero truncated count data, or positive count data (Gurmu, 1991). Examples of zero truncated count data include the number of bus trips made per week in surveys taken on buses, the number of shopping trips made by individuals sampled at a mall, and the number of unemployment spells among a pool of unemployed (Cameron and Trivedi, 2001). Grogger and Carson (1991) apply the truncated model to the

number of recreational fishing trips taken by a sample of Alaskan Fishermen. Such data can be modeled using a truncated version of the Poisson distribution. For the Poisson probability function, a model for count truncated on the left at value $k = 0$ is showed by Grogger and Carson (1991) as follows.

Assume $f(y|\lambda)$ is the density function of Poisson distribution and $F(y|\lambda) = Pr(Y \leq y)$ is the cumulative distribution function. Thus the zero truncated Poisson density, where the value of $y = 0$ is omitted, is given by

$$f(Y = y_i|\lambda, Y > 0) = Pr(Y_i = y_i|Y_i > 0) = \frac{e^{-\lambda_i} \lambda_i^{y_i}}{(1 - e^{-\lambda_i}) y_i!}, \quad (3)$$

where y_i takes only positive integer values larger than 0. The first two moments of the distribution are readily shown to be

$$\begin{aligned} E[Y_i|x_i, Y_i > 0] &= \sum_{y=1}^{\infty} y_i \frac{\lambda_i^{y_i}}{(e^{\lambda_i} - 1) y_i!} \\ &= \frac{\lambda_i}{e^{\lambda_i} - 1} \sum_{y_i} \frac{\lambda_i^{y_i-1}}{y_i - 1}, \end{aligned} \quad (4)$$

and

$$E[Y_i^2|X_i, Y_i > 0] = \frac{\lambda_i^2 e^{\lambda_i}}{e^{\lambda_i} - 1} + \frac{\lambda_i e^{\lambda_i}}{e^{\lambda_i} - 1},$$

giving the variance as

$$Var(Y_i|X_i, Y_i > 0) = \frac{\lambda_i e^{\lambda_i}}{e^{\lambda_i} - 1} - \frac{\lambda_i^2 e^{\lambda_i}}{(e^{\lambda_i} - 1)^2}. \quad (5)$$

It can be seen from (4) and (5) that the positive Poisson model represents underdispersion since the conditional variance is less than the conditional mean. The second part of the right side of (5) may be considered as an adjustment factor, and may be used for tests of overdispersion (Gurmu,1991).

Maximum Likelihood Estimation

To estimate the regression parameters, we need to specify a likelihood criterion. Using the conditional distribution function (3), a joint likelihood of independent responses is

$$L(\boldsymbol{\beta}, y_i) = \prod_{i=1}^m \frac{\lambda_i^{y_i}}{(e^{\lambda_i} - 1)y_i!}.$$

Under the assumption of independent observations of (3), and in the context of the left truncated Poisson model, the joint log-likelihood function is

$$l(\boldsymbol{\beta}; y_i) = \sum_{i=1}^m [y_i \mathbf{x}_i^T \boldsymbol{\beta} - \ln(e^{\lambda_i} - 1) - \ln(y_i!)] .$$

The algorithm needs the first and second order derivation, then it is purely a matter of numerical optimization, and those have been presented by Grogger and Carson (1991).

The first partial derivatives with respect to $\boldsymbol{\beta}$ are

$$\begin{aligned} \frac{\partial l}{\partial \beta_h} &= \sum_{i=1}^m \left[y_i x_i - \frac{1}{e^{e^{\mathbf{x}_i^T \boldsymbol{\beta}}} - 1} e^{e^{\mathbf{x}_i^T \boldsymbol{\beta}}} e^{\mathbf{x}_i^T \boldsymbol{\beta}} x_i \right] \\ &= \sum_{i=1}^m x_{ih} \left[y_i - \frac{e^{\mathbf{x}_i^T \boldsymbol{\beta}}}{1 - \exp(e^{\mathbf{x}_i^T \boldsymbol{\beta}})} \right]. \end{aligned}$$

Setting the derivatives equal to zero, the first order condition for maximum likelihood is

$$\sum_{i=1}^m x_{ih} \left[y_i - \frac{e^{x_i^T \beta}}{1 - \exp(e^{x_i^T \beta})} \right] = 0.$$

To ensure the log-likelihood has been maximized, it would be sufficient to determine that the second partial derivatives are negative. The second partial derivatives with respect to β are

$$\begin{aligned} \frac{\partial^2 l}{d\beta_h d\beta_{h'}} &= - \sum_{i=1}^m x_{ih} \left[\frac{(1 - e^{-x_i^T \beta}) x_{ih'} e^{x_i^T \beta} - e^{x_i^T \beta} (-e^{-x_i^T \beta} (-e^{x_i^T \beta}) x_{ih'})}{(1 - e^{-\lambda_i})^2} \right] \\ &= - \sum_i x_{ih} x_{ih'} e^{x_i^T \beta} \left[\frac{1 - (1 + e^{x_i^T \beta}) \exp(e^{x_i^T \beta})}{(1 - \exp(e^{x_i^T \beta}))^2} \right]. \end{aligned}$$

The MLE $\hat{\beta}$ is asymptotically normal with mean β and variance matrix $I(\beta)^{-1} = -E \left[\frac{\partial^2 l}{d\beta_h d\beta_{h'}} \right]$. In maximum likelihood estimation of truncated models, a misspecification of the underlying distribution leads to inconsistency due to the presence of the adjustment factor. If the distribution is misspecified as the truncated Poisson, then the conditional mean is misspecified and the MLE will be inconsistent (Cameron and Trivedi, 1998).

It can be seen that the zero truncated Poisson distribution does not have equal first and second moments. Furthermore, misspecification of the distribution implies that the first conditional truncated moment will also be misspecified. This can result in inconsistent estimates of β 's if the parent (untruncated) distribution is incorrectly specified (Cameron and Trivedi, 2001).

Left Truncated Poisson Model at $k = c_l$

Count data in which $y < c_l$ counts cannot be observed is called left truncated count data at $k = c_l$. Left truncation can happen in medical studies when someone wants to study length of survival after the start of the disease. Let Y be the time elapsed between the onset of the disease and death, and if the follow up period starts t units of the time after the onset of the disease, then Y is clearly left truncated by t . Left truncated data are also common in actuarial research, astronomic, demographic, epidemiologic, reliability testing, and other studies. As an example, imagine you wish to study how long people who have been hospitalized for a heart attack survive taking some treatment at home. The start time is taken to be the time of the heart attack. Only those individuals who survive their stay in hospital are able to be included in the study. For the Poisson probability function, a model for counts truncated on the left at the value $k = c_l$ can be posited as

$$\begin{aligned}
 Pr(Y_i = y_i | Y_i > c_l) &= \frac{Pr(Y_i = y_i, Y_i > c_l)}{Pr(Y_i > c_l)} = \frac{Pr(Y_i = y_i)}{1 - Pr(Y_i \leq c_l)} \\
 &= \frac{Pr(Y_i = y_i)}{1 - F_p(c_l)} = Pr(Y_i = y_i)[1 - F_p(c_l)]^{-1} \\
 &= \frac{e^{-\lambda_i} \lambda_i^{y_i}}{y_i!} \left[1 - \sum_{k=0}^{c_l} \frac{e^{-\lambda_i} \lambda_i^k}{k!} \right]^{-1} \\
 &= \frac{\lambda_i^{y_i} e^{-\lambda_i}}{y_i!} \frac{1}{1 - e^{-\lambda_i} \sum_{k=0}^{c_l} \frac{\lambda_i^k}{k!}} \\
 &= \frac{\lambda_i^{y_i}}{y_i!} \left[\frac{1}{e^{\lambda_i} - \sum_{k=0}^{c_l} \frac{\lambda_i^k}{k!}} \right] \\
 &= \frac{\lambda_i^{y_i}}{\left[e^{\lambda_i} - \sum_{k=0}^{c_l} \frac{\lambda_i^k}{k!} \right] y_i!}.
 \end{aligned}$$

In the case of the left truncated Poisson model, the first two moments of the distribution are readily shown to be

$$\begin{aligned} E[Y_i|x_i, Y_i > c_l] &= \sum_{Y_i=c_l}^{\infty} y_i \frac{\lambda_i^{y_i}}{y_i!} \left[\frac{1}{e^{\lambda_i} - \sum_{k=0}^{c_l} \frac{\lambda_i^k}{k!}} \right] \\ &= \frac{\lambda_i e^{-\lambda_i}}{1 - \sum_{k=0}^{c_l} e^{-\lambda_i} \frac{\lambda_i^k}{k!}} \sum_{y_i=c_l+1}^{\infty} \frac{\lambda_i^{y_i-1}}{(y_i-1)!}, \end{aligned} \quad (6)$$

and

$$\begin{aligned} E[Y_i^2|X_i, Y_i > c_l] &= \sum_{y_i=c_l}^{\infty} y_i^2 \frac{\lambda_i^{y_i}}{y_i!} \frac{1}{[e^{\lambda_i} - \sum_{k=0}^{c_l} \frac{\lambda_i^k}{k!}]} \\ &= \frac{\lambda_i e^{-\lambda_i}}{1 - \sum_{k=0}^{c_l} \frac{e^{-\lambda_i} \lambda_i^k}{k!}} \left[\lambda_i \sum_{y_i=c_l+2}^{\infty} \frac{\lambda_i^{y_i-2}}{(y_i-2)!} + \sum_{y_i=c_l+1}^{\infty} \frac{\lambda_i^{y_i-1}}{(y_i-1)!} \right], \end{aligned}$$

giving the variance as

$$Var(Y_i|X_i, Y_i > c_l) = \frac{\lambda_i e^{-\lambda_i}}{1 - F_p(c_l)} \left[\lambda_i \sum_{y_i=c_l+2}^{\infty} \frac{\lambda_i^{y_i-2}}{(y_i-2)!} + \sum_{y_i=c_l+1}^{\infty} \frac{\lambda_i^{y_i-1}}{(y_i-1)!} \right] - \left[\frac{\lambda_i e^{-\lambda_i}}{[1 - F_p(c_l)]} \sum_{y_i=c_l+1}^{\infty} \frac{\lambda_i^{y_i-1}}{(y_i-1)!} \right]^2. \quad (7)$$

Equations (6) and (7) show that the mean of the left-truncated random variable exceeds the corresponding mean of the untruncated distribution model, whereas the variance of left-truncated random variable is smaller than the corresponding variance of the untruncated distribution model. Cameron and Trivedi (1998) expressed the relationship between the truncated mean and untruncated distribution mean as

$$E[y_i|y_i \geq c_l] = E[y_i] + \delta_i, \quad \delta_i > 0,$$

where δ_i is the difference between the truncated and untruncated means.

Maximum Likelihood Estimation

The left truncated Poisson model whose parameters need to be estimated is given by

$$Pr(Y_i = y_i | Y_i > c_l) = \frac{\lambda_i^{y_i}}{[e^{\lambda_i} - \sum_{k=0}^{c_l} \frac{\lambda_i^k}{k!}] y_i!}, \quad (8)$$

where $\lambda_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta})$, $i = 1, \dots, m$. Under the assumption of independent observations of (8), and in the context of the left truncated Poisson model, the joint likelihood and log-likelihood functions are

$$L(\boldsymbol{\beta}; y_i) = \prod_{i=1}^m \frac{\lambda_i^{y_i}}{y_i!} \left[\frac{1}{e^{\lambda_i} - \sum_{k=0}^{c_l} \frac{\lambda_i^k}{k!}} \right],$$

and

$$\begin{aligned} l(\boldsymbol{\beta}; y_i) &= \sum_{i=1}^m \left[y_i \ln \lambda_i - \ln(y_i!) - \ln \left(e^{\lambda_i} - \sum_{k=0}^{c_l} \frac{\lambda_i^k}{k!} \right) \right] \\ &= \sum_{i=1}^m \left[y_i \ln e^{\mathbf{x}_i^T \boldsymbol{\beta}} - \ln(y_i!) - \ln \left(e^{\lambda_i} - \sum_{k=0}^{c_l} \frac{\lambda_i^k}{k!} \right) \right] \\ &= \sum_{i=1}^m \left[y_i \mathbf{x}_i^T \boldsymbol{\beta} - \ln(y_i!) - \ln \left(e^{\lambda_i} - \sum_{k=0}^{c_l} \frac{\lambda_i^k}{k!} \right) \right]. \end{aligned} \quad (9)$$

It is easy to see that the maximum likelihood left truncated Poisson regression estimators can be obtained from solving the first derivative with respect to $\boldsymbol{\beta}$ of (9) as follows. The first partial derivatives with respect to β are

$$\frac{\partial l}{\partial \beta_h} = \sum_{i=1}^m x_{ih} \left[y_i - \frac{e^{\mathbf{x}_i^T \boldsymbol{\beta}} e^{\mathbf{x}_i^T \boldsymbol{\beta}} - \sum_{k=0}^{c_l} \frac{k e^{\mathbf{x}_i^T \boldsymbol{\beta} k}}{k!}}{e^{\mathbf{x}_i^T \boldsymbol{\beta}} - \sum_{k=0}^{c_l} \frac{e^{\mathbf{x}_i^T \boldsymbol{\beta} k}}{k!}} \right].$$

Setting the derivatives equal to zero, the first order condition of maximum likelihood is

$$\sum_{i=1}^m x_{ih} \left[y_i - \frac{e^{\mathbf{x}_i^T \boldsymbol{\beta}} e^{e^{\mathbf{x}_i^T \boldsymbol{\beta}}} - \sum_{k=0}^{c_l} \frac{k e^{\mathbf{x}_i^T \boldsymbol{\beta} k}}{k!}}{e^{e^{\mathbf{x}_i^T \boldsymbol{\beta}}} - \sum_{k=0}^{c_l} \frac{e^{\mathbf{x}_i^T \boldsymbol{\beta} k}}{k!}} \right] = 0. \quad (10)$$

The system (10) above is nonlinear concerning the parameters so that one has to use an iterative method such as Newton-Raphson, or Fisher scoring. To ensure the log-likelihood has been maximized, it would be sufficient to determine that the second partial derivatives are negative. The second partial derivatives with respect to $\boldsymbol{\beta}$ are

$$\begin{aligned} \frac{\partial^2 l}{\partial \beta_h \partial \beta_{h'}} &= - \sum_{i=1}^m x_{ih} x_{ih'} \left[\frac{\left(e^{e^{\mathbf{x}_i^T \boldsymbol{\beta}}} - \sum_{k=0}^{c_l} \frac{e^{\mathbf{x}_i^T \boldsymbol{\beta} k}}{k!} \right) \left[e^{2\mathbf{x}_i^T \boldsymbol{\beta}} e^{e^{\mathbf{x}_i^T \boldsymbol{\beta}}} + e^{\mathbf{x}_i^T \boldsymbol{\beta}} e^{e^{\mathbf{x}_i^T \boldsymbol{\beta}}} - \sum_{k=0}^{c_l} \frac{k^2 e^{\mathbf{x}_i^T \boldsymbol{\beta} k}}{k!} \right]}{\left(e^{e^{\mathbf{x}_i^T \boldsymbol{\beta}}} - \sum_{k=0}^{c_l} \frac{e^{\mathbf{x}_i^T \boldsymbol{\beta} k}}{k!} \right)^2} \right] \\ &+ \sum_{i=1}^m x_{ih} x_{ih'} \left[\frac{\left(e^{\mathbf{x}_i^T \boldsymbol{\beta}} e^{e^{\mathbf{x}_i^T \boldsymbol{\beta}}} - \sum_{k=0}^{c_l} \frac{k e^{\mathbf{x}_i^T \boldsymbol{\beta} k}}{k!} \right) \left(e^{\mathbf{x}_i^T \boldsymbol{\beta}} e^{e^{\mathbf{x}_i^T \boldsymbol{\beta}}} - \sum_{k=0}^{c_l} \frac{k e^{\mathbf{x}_i^T \boldsymbol{\beta} k}}{k!} \right)}{\left(e^{e^{\mathbf{x}_i^T \boldsymbol{\beta}}} - \sum_{k=0}^{c_l} \frac{e^{\mathbf{x}_i^T \boldsymbol{\beta} k}}{k!} \right)^2} \right]. \end{aligned}$$

The MLE $\hat{\boldsymbol{\beta}}$ is asymptotically normal with mean $\boldsymbol{\beta}$ and variance matrix $I(\boldsymbol{\beta})^{-1} = -E \left[\frac{\partial^2 l}{\partial \beta_h \partial \beta_{h'}} \right]$. In maximum likelihood estimation of truncated models, a misspecification of the underlying distribution leads to inconsistency due to the presence of the adjustment factor. Ignoring the left truncation data can lead to substantially different statistical estimates (Liu, Pitt, Wang, and Wu, 2012). They found that ignoring the left-truncation in the data leads the bias of the parameter estimates.

Right Truncated Poisson Model at $k = c_r$

Count data in which $y > c_r$ counts cannot be observed are called right truncated count data at $k = c_r$. “This type of truncation arise when there is difficulty in counting high numbers owing to inability to distinguish each individual or when the counting apparatus gives trouble for high count” Moore (1954). For example, if we want to model

the number of times (per week) an in-vehicle navigation system is used on the morning commute to work during weekdays, the data are right truncated at 5, which is the maximum number of uses in any given week. As another example, if you ask a group of smoking school pupils at what age they started smoking, you necessarily have truncated data, as individuals who start smoking after leaving school are not included in the study. Similarly, if a drug developed in 1990 is compared with a placebo to analyze the survival times of (dead) cancer patients, all times to death must be ≤ 23 years. Estimating a Poisson regression model without considering this truncation will cause biased estimates of the parameter vector β and erroneous inferences will be drawn (Liu, Pitt, Wang, and Wu, 2012). For the Poisson probability function, a model for count data truncated on the right at value $k = c_l$ can be posited as

$$\begin{aligned}
 Pr(Y_i = y_i | Y_i \leq c_r) &= \frac{Pr(Y_i = y_i / Y_i \leq c_r)}{Pr(Y_i \leq c_r)} \\
 &= \frac{Pr(Y_i = y_i)}{F_p(c_r)} \\
 &= \frac{Pr(Y_i = y_i) [F_p(c_r)]^{-1}}{F_p(c_r)} \\
 &= \frac{e^{-\lambda_i} \lambda_i^{y_i}}{y_i!} \left[\sum_{k=0}^{c_r} \frac{e^{-\lambda_i} \lambda_i^k}{k!} \right]^{-1} \\
 &= \frac{\lambda_i^{y_i}}{\left(\sum_{k=0}^{c_r} \frac{\lambda_i^k}{k!} \right) y_i!}, \quad i = 1, \dots, m,
 \end{aligned}$$

where m is the number of observation after truncation. In the case of the right truncated Poisson model, the first two moments of the distribution are readily shown to be

$$\begin{aligned}
 E[Y_i = y_i | x_i, y_i \leq c_r] &= \sum_{y_i=0}^{c_r} y_i \frac{\lambda_i^{y_i}}{\left(\sum_{k=0}^{c_r} \frac{\lambda_i^k}{k!} \right) y_i!} \\
 &= \frac{\lambda_i}{\sum_{k=0}^{c_r} \frac{\lambda_i^k}{k!}} \sum_{y_i=1}^{c_r} \frac{\lambda_i^{y_i-1}}{(y_i - 1)!},
 \end{aligned} \tag{11}$$

and

$$\begin{aligned}
E[Y_i^2 | X_i, Y_i \leq c_r] &= \sum_{y_i=0}^{c_r} y_i^2 \frac{\lambda_i^{y_i}}{\left(\sum_{k=0}^{c_r} \frac{\lambda_i^k}{k!} \right) y_i!} \\
&= \frac{1}{\sum_{k=0}^{c_r} \frac{\lambda_i^k}{k!}} \left[\lambda_i^2 \sum_{Y_i=2}^{c_r} \frac{\lambda_i^{y_i-2}}{(y_i-2)!} + \lambda_i \sum_{y_i=1}^{c_r} \frac{\lambda_i^{y_i-1}}{(y_i-1)!} \right],
\end{aligned}$$

giving the variance as

$$Var(Y_i | X_i, Y_i \leq c_r) = \frac{1}{\sum_{k=0}^{c_r} \frac{\lambda_i^k}{k!}} \left[\lambda_i^2 \sum_{y_i=2}^{c_r} \frac{\lambda_i^{y_i-2}}{(y_i-2)!} + \lambda_i \sum_{y_i=1}^{c_r} \frac{\lambda_i^{y_i-1}}{(y_i-1)!} \right] - \left[\frac{\lambda_i}{\sum_{k=0}^{c_r} \frac{\lambda_i^k}{k!}} \sum_{y_i=1}^{c_r} \frac{\lambda_i^{y_i-1}}{(y_i-1)!} \right]^2. \quad (12)$$

It can be seen that right truncated Poisson distribution results in a smaller mean and variance compared to the standard Poisson distribution (Cameron and Trivedi, 2001).

Maximum Likelihood Estimation

The right truncated Poisson model whose parameters need to be estimated is given by

$$Pr(Y_i = y_i | Y_i \leq c_r) = \frac{\lambda_i^{y_i}}{\left(\sum_{k=0}^{c_r} \frac{\lambda_i^k}{k!} \right) y_i!}, \quad (13)$$

where $\lambda_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta})$, $i = 1, \dots, n$. Under the assumption of independent observations of (13), and in the context of the right truncated Poisson model, the joint likelihood and log-likelihood functions are

$$L(\boldsymbol{\beta}; y_i) = \prod_{i=1}^m \frac{\lambda_i^{y_i}}{\left(\sum_{k=0}^{c_r} \frac{\lambda_i^k}{k!} \right) y_i!},$$

and

$$\begin{aligned} l(\boldsymbol{\beta}; y_i) &= \sum_{i=1}^m \left[y_i \ln \lambda_i - \ln(y_i!) - \ln \left(\sum_{k=0}^{c_r} \frac{\lambda_i^k}{k!} \right) \right] \\ &= \sum_{i=1}^m \left[y_i \mathbf{x}_i^T \boldsymbol{\beta} - \ln(y_i!) - \ln \left(\sum_{k=0}^{c_r} \frac{e^{\mathbf{x}_i^T \boldsymbol{\beta} k}}{k!} \right) \right]. \end{aligned}$$

To obtain the maximum likelihood right truncated Poisson regression estimators, we take the first derivative with respect to $\boldsymbol{\beta}$ and set equal to zero as follows. The first partial derivatives with respect to $\boldsymbol{\beta}$ are

$$\begin{aligned} \frac{\partial l}{\partial \beta_h} &= \sum_{i=1}^m \left[y_i \mathbf{x}_i - \frac{\sum_{k=0}^c \frac{\mathbf{x}_i k e^{\mathbf{x}_i^T \boldsymbol{\beta} k}}{k!}}{\sum_{k=0}^c \frac{e^{\mathbf{x}_i^T \boldsymbol{\beta} k}}{k!}} \right] \\ &= \sum_{i=1}^m x_{ih} \left[y_i - \frac{\sum_{k=0}^c \frac{k e^{\mathbf{x}_i^T \boldsymbol{\beta} k}}{k!}}{\sum_{k=0}^c \frac{e^{\mathbf{x}_i^T \boldsymbol{\beta} k}}{k!}} \right]. \end{aligned}$$

Setting the derivatives equal to zero, the first order condition for maximum likelihood is

$$\sum_{i=1}^m x_{ih} \left[y_i - \frac{\sum_{k=0}^c \frac{k e^{\mathbf{x}_i^T \boldsymbol{\beta} k}}{k!}}{\sum_{k=0}^c \frac{e^{\mathbf{x}_i^T \boldsymbol{\beta} k}}{k!}} \right] = 0. \quad (14)$$

The system defined by (14) above is nonlinear concerning the parameters so that one has to use an iterative method, such as Newton Raphson, or Fisher scoring. To ensure the log-likelihood has been maximized, it would be sufficient to determine that the second partial derivatives is negative. The second partial derivatives with respect to $\boldsymbol{\beta}$ are

$$\frac{\partial^2 l}{\partial \beta_h \partial \beta_{h'}} = - \sum_{i=1}^m x_{ih} x_{ih'} \left[\frac{\left(\sum_{k=0}^c \frac{e^{\mathbf{x}_i^T \boldsymbol{\beta} k}}{k!} \right) \left(\sum_{k=0}^c \frac{k^2 e^{\mathbf{x}_i^T \boldsymbol{\beta} k}}{k!} \right)}{\left[\sum_{k=0}^c \frac{e^{\mathbf{x}_i^T \boldsymbol{\beta} k}}{k!} \right]^2} \right] + \sum_{i=1}^m x_{ih} x_{ih'} \left[\frac{\left(\sum_{k=0}^c \frac{k e^{\mathbf{x}_i^T \boldsymbol{\beta} k}}{k!} \right)^2}{\left[\sum_{k=0}^c \frac{e^{\mathbf{x}_i^T \boldsymbol{\beta} k}}{k!} \right]^2} \right].$$

The MLE $\hat{\beta}$ is asymptotically normal with mean β and variance matrix $I(\beta)^{-1} = -E \left[\frac{\partial^2 l}{\partial \beta_h \partial \beta_{h'}} \right]$. If the distribution is misspecified as the truncated Poisson, then the conditional mean is misspecified and the MLE will be inconsistent (Cameron and Tivedi, 2001). Failure to account for truncation leads to biased estimates and can bias inferences (Kalbfleisch and Lawless, 1991).

Double Truncated Poisson Model

Double truncated Poisson data are a combination of the left truncated and right truncated Poisson data. This type of truncation occurs sometimes in botanical work where only quadrats with at least one individual in them are retained and high densities are difficult to count (Moore, 1954). As another example, observations on the number of items purchased when there is a limit per customer would exclude the zero class (since customers who do not buy the item will not be identified), and would exclude observations above the limit. Furthermore, a double truncated Poisson model has been considered for regression analysis of time series of counts (Kedem and Fokianos, 2002). From Cohen (1954) the probability mass function for the double truncated Poisson random variable y is given by

$$Pr(Y_i = y_i | x_i, c_l \leq y_i \leq c_r) = \frac{\lambda_i^{y_i}}{y_i! \sum_{k=c_l}^{c_r} \frac{\lambda_i^k}{k!}},$$

where $\lambda_i = \exp(\mathbf{x}_i^T \beta)$, and c_l and c_r are the lower end and the upper end of the interval in which y can only be observed. In the case of the double truncated Poisson model, the first two moments of the distribution are readily shown to be

$$E[Y_i = y_i | X_i, c_l \leq y_i \leq c_r] = \frac{\sum_{k=c_l}^{c_r} \frac{\lambda_i^k}{(k-1)!}}{\sum_{k=c_l}^{c_r} \frac{\lambda_i^k}{k!}}, \quad (15)$$

and

$$E[Y_i^2 = y_i | X_i, c_l \leq y_i \leq c_r] = \frac{\sum_{k=c_l}^{c_r} \frac{k\lambda_i^k}{(k-1)!}}{\sum_{k=c_l}^{c_r} \frac{\lambda_i^k}{k!}},$$

giving the variance as

$$Var(Y_i = y_i | X_i, c_l \leq y_i \leq c_r) = \frac{\sum_{k=c_l}^{c_r} \frac{k\lambda_i^k}{(k-1)!}}{\sum_{k=c_l}^{c_r} \frac{\lambda_i^k}{k!}} - \left(\frac{\sum_{k=c_l}^{c_r} \frac{\lambda_i^k}{(k-1)!}}{\sum_{k=c_l}^{c_r} \frac{\lambda_i^k}{k!}} \right)^2. \quad (16)$$

While the regular Poisson model typically encounters difficulty due to the assumed equality of mean and variance, the mean and variance of the doubly-truncated Poisson model are characteristic of underdispersion where the variance is less than the mean. Testing for overdispersion must now take this assumption into account (Gurmu and Trivedi, 1992). This assumed inequality also provides an intuitive reason as to why fitting a regular Poisson model to truncated data is a fundamental misspecification. The regular Poisson assumes greater variance than should be expected.

Maximum Likelihood Estimation

The double truncated Poisson model whose parameters need to be estimated is given by

$$Pr(Y_i = y_i | x_i; c_l \leq y_i \leq c_r) = \frac{\lambda_i^{y_i}}{y_i! \sum_{k=c_l}^{c_r} \frac{\lambda_i^k}{k!}}. \quad (17)$$

where $\lambda_i = \exp(\mathbf{x}_i^T \boldsymbol{\beta})$. Under the assumption of independent observations of (17), and in the context of the double truncated Poisson model, the joint likelihood and log-likelihood functions are

$$L(\boldsymbol{\beta}; y_i) = \prod_{i=1}^m \frac{\lambda_i^{y_i}}{y_i! \sum_{k=c_l}^{c_r} \frac{\lambda_i^k}{k!}},$$

and

$$l(\boldsymbol{\beta}; y_i) = \sum_{i=1}^m \left[y_i \mathbf{x}_i^T \boldsymbol{\beta} - \ln(y_i!) - \ln \left(\sum_{k=c_l}^{c_r} \frac{\lambda_i^k}{k!} \right) \right]. \quad (18)$$

The maximum likelihood double truncated Poisson regression estimators are found by solving the following first order condition of (18)

$$\sum_{i=1}^m x_{ih} \left[y_i - \frac{\sum_{k=c_l}^{c_r} \frac{k e^{\mathbf{x}_i^T \boldsymbol{\beta} k}}{k!}}{\sum_{k=c_l}^{c_r} \frac{e^{\mathbf{x}_i^T \boldsymbol{\beta} k}}{k!}} \right] = 0. \quad (19)$$

Here we note that there is no closed form solution for (19), hence estimation must be accomplished numerically using methods such as Newton-Raphson, or Fisher scoring. To ensure the log-likelihood has been maximized, it would be sufficient to determine that the second partial derivatives is negative. The second partial derivatives with respect to $\boldsymbol{\beta}$ are

$$\frac{\partial^2 l}{\partial \beta_h \partial \beta_{h'}} = - \sum_{i=1}^m x_{ih} x_{ih'} \left[\frac{\left(\sum_{k=c_l}^{c_r} \frac{e^{\mathbf{x}_i^T \boldsymbol{\beta} k}}{k!} \right) \left(\sum_{k=c_l}^{c_r} \frac{k^2 e^{\mathbf{x}_i^T \boldsymbol{\beta} k}}{k!} \right)}{\left[\sum_{k=c_l}^{c_r} \frac{e^{\mathbf{x}_i^T \boldsymbol{\beta} k}}{k!} \right]^2} \right] + \sum_{i=1}^m x_{ih} x_{ih'} \left[\frac{\left(\sum_{k=c_l}^{c_r} \frac{k e^{\mathbf{x}_i^T \boldsymbol{\beta} k}}{k!} \right)^2}{\left[\sum_{k=c_l}^{c_r} \frac{e^{\mathbf{x}_i^T \boldsymbol{\beta} k}}{k!} \right]^2} \right].$$

The MLE $\hat{\boldsymbol{\beta}}$ is asymptotically normal with mean $\boldsymbol{\beta}$ and variance matrix $I(\boldsymbol{\beta})^{-1} = -E \left[\frac{\partial^2 l}{\partial \beta_h \partial \beta_{h'}} \right]$. In maximum likelihood estimation of truncated models a misspecification of the underlying distribution can lead to inconsistency due to the presence of the adjustment factor (Cameron and Trivedi, 1998). The effect of truncating the distribution of a response variable is clear in which the mean of the truncated distribution moves away from the truncation point, and the variance is reduced. Therefore, ignoring truncation leads to strongly biased estimators and the bias will result in over-estimates of parameters being much higher than expected (Baud, Frachot, and Roncalli, 2002).

Truncated count data have a long history in genetic epidemiology, astronomic, demographic reliability testing, and many other studies. Therefore, misspecification of the

distribution of truncated data implies that the first conditional truncated moment will also be misspecified. Cameron and Trivedi (1998) verified that misspecification of correct distribution will result in inconsistent estimators of $\beta's$. Liu, Pitt, Wang, and Wu (2012) found that ignoring the left truncation in Poisson model leads to extend the bias of the parameter estimates. In addition, Kalbfleisch and Lawless (1991) showed that failure to account for truncation leads to biased estimates. Baud, Frachot, and Roncalli (2002) showed that ignoring truncation leads to strongly biased estimators and the bias will be over estimated and much higher than really expected. Finally, accounting for truncation of the dependent variable makes a substantial difference in the coefficient estimates, and subsequently, in benefit estimates, regardless of the choice of statistical model (Creel and Loomis, 1990). Although, there are many studies that investigated the truncated Poisson regression model, the majority of them considered only the left truncated Poisson regression model or the zero-truncated Poisson regression model as a special case.

In a forthcoming paper by Suaiee and Lalonde (2013), the effects of power, bias and standard error of parameter estimates, standard error of ignoring left, right and double truncated count data were studied through simulation. It was verified that the power of ignoring truncation is less than the power of considering truncation. Also, it was verified that the parameter estimates of truncation are asymptotically unbiased while they are biased in case of ignoring truncation and the bias is severe. Furthermore, standard errors of the estimates in the case of truncation are very reasonable and decrease as the sample size increases. In the case of ignoring truncation, the standard errors were underestimated.

Linear Mixed Model

Linear mixed models have attracted considerable attention over the years and are widely used in the biological and social sciences. The linear mixed effect models can accommodate data from different settings such as clustered, longitudinal, and repeated measurements. Mixed model analysis provides a general, flexible approach in these situations. Linear mixed effects models include both the fixed and the random effects in

which they contribute linearly to the response function. The general form of such models is:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\epsilon}, \quad (20)$$

where

\mathbf{y} is the $n \times 1$ vector of observations,

$\boldsymbol{\beta}$ is the $p \times 1$ vector of fixed effects,

\mathbf{u} is the $q \times 1$ vector of random effects,

$\boldsymbol{\epsilon}$ is the $n \times 1$ vector of random error terms,

\mathbf{X} is the $n \times p$ design matrix for the fixed effects relating observations \mathbf{Y} to $\boldsymbol{\beta}$,

\mathbf{Z} is the $n \times q$ design matrix for the random effects relating observations \mathbf{Y} to \mathbf{u} .

We assume that \mathbf{u} and $\boldsymbol{\epsilon}$ are uncorrelated random variables with zero means and covariance matrices \mathbf{G} and \mathbf{R} respectively, so

$$E[\mathbf{u}] = \mathbf{0}, \text{ } Var(\mathbf{u}) = \mathbf{G},$$

$$E[\boldsymbol{\epsilon}] = \mathbf{0}, \text{ } Var(\boldsymbol{\epsilon}) = \mathbf{R},$$

$$Cov(\boldsymbol{\epsilon}, \mathbf{u}) = \mathbf{0}.$$

Thus, the expectation and variance (\mathbf{V}) of the observation vector \mathbf{y} are given by

$$E[\mathbf{y}] = \mathbf{X}\boldsymbol{\beta},$$

$$Var(\mathbf{y}) = \mathbf{V} = \mathbf{ZGZ}^T + \mathbf{R}.$$

Understanding the \mathbf{V} matrix is a very important component in order to work with mixed models since it contains both sources of random variation and defines how these models differ computationally from Ordinary Least Squares (OLS). If we also assume the

random terms are normally distributed as

$$\mathbf{u} \sim N(\mathbf{0}, \mathbf{G}), \text{ and } \boldsymbol{\epsilon} \sim N(\mathbf{0}, \mathbf{R}),$$

then, the observation vector \mathbf{y} will also be normally distributed,

$$\mathbf{y} \sim N(\mathbf{X}\boldsymbol{\beta}, \mathbf{V}).$$

For the linear mixed model in (20), Henderson's (1975) mixed model equations (MME) can be used to find $\hat{\boldsymbol{\beta}}$ and $\hat{\mathbf{u}}$. The best linear unbiased estimator (BLUE) of $\boldsymbol{\beta}$, and the best linear unbiased predictor (BLUP) of \mathbf{u} , respectively, are

$$\begin{bmatrix} \mathbf{X}^T \mathbf{R}^{-1} \mathbf{X} & \mathbf{X}^T \mathbf{R}^{-1} \mathbf{Z} \\ \mathbf{Z}^T \mathbf{R}^{-1} \mathbf{X} & \mathbf{Z}^T \mathbf{R}^{-1} \mathbf{Z} + \mathbf{G}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^T \mathbf{R}^{-1} \mathbf{Y} \\ \mathbf{Z}^T \mathbf{R}^{-1} \mathbf{Y} \end{bmatrix}.$$

If the \mathbf{G} and \mathbf{R} matrices are known, generalized least squares can estimate any linear combination of the fixed $\boldsymbol{\beta}$. However, as usually we do not know these matrices, a complex iterative algorithm for fitting linear models must be used to estimate them.

For covariance matrix estimators, restricted maximum likelihood estimates (REML) are usually preferred to maximum likelihood estimates (MLEs) in linear mixed effects models. REML was suggested by Thompson (1962) and was described formally by Patterson and Thompson (1971).

Count Model With Random Effect

The standard Poisson regression model for count responses assumes observations to be independent (Hedeker and Gibbons, 2006). A random effects model should be used if the assumption of independence between the individual specific effects and the regressors

does not appear tenable. In this case, a random effect model will tend to be more efficient (Winkelmann, 2008). Specifically, in many cases the frequency data are clustered or longitudinal and the assumption of independence is not reasonable (Hedeker and Gibbons, 2006). For example, in studies of mental health service utilization, changes in utilization frequency over time is often the primary focus of the investigation. Likewise, in cross sectional studies where observations are clustered within geographic units, observations within the geographic units are more likely to be correlated than observations between geographic units. Goldstein (1991) proposed a multilevel log-linear model. Breslow (1984) described a Poisson model with normally distributed random effects, and Lawless and Willmot (1989) studied a Poisson regression model with random effects having an inverse Gaussian distribution. Siddiqui (1996) developed a Poisson mixed effects regression model for clustered count data, and compared models with normally distributed and gamma distributed random effects.

Mixed Effects Poisson Regression Model

Mixed Poisson distributions have been widely used in scientific fields for modeling nonhomogeneous populations. Historically, mixed Poisson distributions were driven by Greenwood and Yule in 1920 where they considered the negative binomial distribution as a mixture of a Poisson distribution with a Gamma mixing distribution. Based on the choice of the mixing distribution, various mixed Poisson distributions can be constructed. Since then, a large number of mixed Poisson distributions has showed up in the literature. However, only a few of them have been used in practice, the main reason being that often their form is complicated (Karlis and Xekalaki, 2005).

Let Y_{ij} be the number of independent events that occur during a fixed time period, realized for observation j in group i as y_{ij} , which takes any non-negative integer as its value for $i = 1, \dots, m$, $j = 1, \dots, n_i$. The Poisson distribution with mean and variance λ_{ij} is

described by

$$Y_{ij} \sim Poi(y_{ij}|\lambda_{ij}) = \frac{\exp(\lambda_{ij})\lambda_{ij}^{y_{ij}}}{y_{ij}!}, \quad y_{ij} = 0, 1, 2, \dots \quad (21)$$

The q -dimensional vector of random effects, \mathbf{u} , is restricted to be mean zero, and therefore is completely characterized by the variance matrix \mathbf{D} , a $(q \times q)$ symmetric positive semi-definite matrix,

$$\mathbf{u}_i \sim (\mathbf{0}, \mathbf{D}).$$

The systematic component is

$$\lambda_{ij} = \exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i),$$

where

\mathbf{x}_{ij} is the $(n_i \times p \times m)$ array of known fixed effects explanatory variables,

$\boldsymbol{\beta}$ is the p -dimensional vector of fixed effects coefficients,

\mathbf{z}_{ij} is the $(n_i \times q \times m)$ array of known random effects explanatory variables,

\mathbf{u} is the q -dimensional vector of random effects.

Parameter Estimation

Different estimation methods for the parameters in the mixed Poisson regression model (21) have been proposed over the years (Searle, Casella, and McCulloch, 1992), but some of the most commonly used methods today are Maximum Likelihood (ML) (Longford, 1993), Penalized Likelihood (PL), and Hierarchical Likelihood (HL).

Maximum Likelihood

For simplicity, consider the model in (21) with a single random effect u_i (Hedeker and Gibbons, 2006) such that

$$\begin{aligned}\lambda_{ij} &= \exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + u_i) \\ &= \exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \sigma \theta_i),\end{aligned}$$

where $\theta_i = \frac{u_i}{\sigma}$ such that $\theta_i \sim N(0, 1)$.

Thus the log-likelihood function corresponding to (21) as in (Hedeker and Gibbons, 2006) is

$$l(\mathbf{y}_i | \theta) = - \sum_{j=1}^{n_i} [\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \sigma \theta_i) + y_{ij}(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \sigma \theta_i) - \log(y_{ij}!)] .$$

The first derivatives with respect to $\boldsymbol{\beta}$ and σ are

$$\frac{\partial l(\mathbf{y}_i | \theta)}{\partial \boldsymbol{\beta}} = \sum_{j=1}^{n_i} (y_{ij} - \exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \sigma \theta_i)) \mathbf{x}_{ij},$$

and

$$\frac{\partial l(\mathbf{y}_i | \theta)}{\partial \sigma} = \sum_{j=1}^{n_i} (y_{ij} - \exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \sigma \theta_i)) \theta_i.$$

Parameters of the model can be estimated by setting the above first derivatives to zero and iteratively solved using the Fisher scoring or Newton-Raphson method.

To ensure the log-likelihood has been maximized, it would be sufficient to determine that the second partial derivatives is negative. The second derivatives with respect to $\boldsymbol{\beta}$ and σ are

$$\frac{\partial^2 l(\mathbf{y}_i | \theta)}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'} = - \sum_{j=1}^{n_i} (\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \sigma \theta_i)) \mathbf{x}_{ij} \mathbf{x}_{ij}',$$

and

$$\frac{\partial^2 l(\mathbf{y}_i|\theta)}{\partial \sigma \partial \sigma'} = - \sum_{j=1}^{n_i} (\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \sigma \theta_i)) \theta_i^2.$$

The MLE $\hat{\boldsymbol{\beta}}$ is asymptotically normal with mean $\boldsymbol{\beta}$ and variance matrix

$$I(\boldsymbol{\beta})^{-1} = -E \left[\frac{\partial^2 l}{\partial \beta_h \partial \beta_{h'}} \right].$$

One problem of the Maximum Likelihood Estimation is that the estimated variance components are negatively biased (Searle, Casella, and McCulloch, 1992). For example, in normally distributed data, if we knew the population mean μ , an unbiased estimate of the variance is $\frac{\sum (X_i - \mu)^2}{n}$, where n is the sample size. Since we usually do not know μ , the MLE of the variance is $\frac{\sum (X_i - (\bar{X}))^2}{n}$ so

$$\frac{\sum (X_i - (\bar{X}))^2}{n} \leq \frac{\sum (X_i - \mu)^2}{n}.$$

Thus, the MLE of the variance is an underestimate since the unbiased estimate of the variance is $\frac{\sum (X_i - (\bar{X}))^2}{n-1}$. MLEs, although widely used and often very convenient, have other problems. At a more practical level, in real world examples the likelihood function may take an intricate form, mainly when several parameters are estimated. This can lead to problems in actually finding the MLE because of multiple maxima (Gates, 1993). When a model has several parameters, there is often a choice of parameterization. The choice can be significant in interpreting the parameters, and in determining how easy it is to find MLEs (Morgan, 2000). Another type of problem arises when the MLE occurs at the boundary of the region of allowable variable such as zero variance for a random effect. This can ruin the nice asymptotic properties of the MLE (Catchpole and Morgan, 1994).

Penalized Quasilikelihood

This method of estimation has been called Penalized Quasilikelihood (PQL) because it obtains from optimizing a quasi-likelihood (involving only first and second derivatives)

with a penalty term on the random effects. The PQL method (Breslow and Clayton 1993) is based on a decomposition of the data into the mean and an error term, with a first-order Taylor series expansion of the mean which is a non-linear function of the linear predictor. It is analogous to iteratively reweighted least squares for linear models in that the model is linear in each iteration (Fitzmaurice, Davidian, Molenberghs and Verbeke 2008). PQL estimates are less accurate than ML because asymptotic results require an increasing number of observations for each random effect. The PQL by Breslow and Clayton (1993) is the most popular for the generalized linear mixed models. It approximates high dimensional integration using the Laplace approximation and the approximated likelihood function is assumed to have a Gaussian distribution. Subsequently, it suggests to apply linear mixed model restricted maximum likelihood (REML) estimation to the normal theory problem introduced in Harville (1977). Even though the penalized likelihood method has been widely used in various applications, it is known that estimating the variance components is quite challenging due to their non observability.

PQL is a method that approximates the data by the mean $E[\mathbf{y}_{ij}|\mathbf{u}_i]$ plus an error term ϵ_{ij} with variance equal to $Var(\mathbf{y}_{ij}|\mathbf{u}_i)$. It then uses a Taylor approximation.

$$\begin{aligned}
y_{ij} &\approx \lambda_{ij} + \epsilon_{ij} \\
&= f(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i) + \epsilon_{ij} \\
&= f\left(\mathbf{x}_{ij}^T \hat{\boldsymbol{\beta}} + \mathbf{z}_{ij}^T \hat{\mathbf{u}}_i\right) + f'\left(\mathbf{x}_{ij}^T \hat{\boldsymbol{\beta}} + \mathbf{z}_{ij}^T \hat{\mathbf{u}}_i\right) \mathbf{x}_{ij}^T (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) + f'\left(\mathbf{x}_{ij}^T \hat{\boldsymbol{\beta}} + \mathbf{z}_{ij}^T \hat{\mathbf{u}}_i\right) \mathbf{z}_{ij}^T (\mathbf{u}_i - \hat{\mathbf{u}}_i) + \epsilon_{ij}.
\end{aligned}$$

This gives

$$y_{ij} \approx \hat{\lambda}_i + \hat{\mathbf{V}}_i \mathbf{X}_i (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) + \hat{\mathbf{V}}_i \mathbf{Z}_i (\mathbf{u}_i - \hat{\mathbf{u}}_i) + \epsilon_{ij}, \quad (22)$$

where $\hat{\lambda}_i$ contains the values $\hat{\lambda}_{ij} = f\left(\mathbf{x}_{ij}^T \hat{\boldsymbol{\beta}} + \mathbf{z}_{ij}^T \hat{\mathbf{u}}_i\right)$, $\hat{\mathbf{V}}_i$ is the diagonal matrix with elements $V(\hat{\lambda}_{ij}) = f'\left(\mathbf{x}_{ij}^T \hat{\boldsymbol{\beta}} + \mathbf{z}_{ij}^T \hat{\mathbf{u}}_i\right)$, and \mathbf{X}_i and \mathbf{Z}_i contain the x_{ij} and z_{ij} respectively.

Reordering and and pre-multiplication with $\hat{\mathbf{V}}_i^{-1}$ gives

$$\mathbf{y}_i^* := \hat{\mathbf{V}}_i^{-1}(\mathbf{y}_i - \hat{\lambda}_i) + \mathbf{X}_i\hat{\beta} + \mathbf{Z}_i\hat{\mathbf{u}}_i + \epsilon_i^* \approx \mathbf{X}_i\beta + \mathbf{Z}_i\mathbf{u}_i + \epsilon_i^*,$$

where $\epsilon_i^* = \hat{\mathbf{V}}_i^{-1}\epsilon_i$ still has mean zero. This can be viewed as a linear mixed model for the pseudo data \mathbf{y}_i^* .

Algorithm for PQL

- Step 1: For given values of β and θ , empirical Bayes estimates for \mathbf{u}_i , which can be any initial values, and resulting pseudo data \mathbf{y}_i^* are computed.
- Step 2: For given pseudo data \mathbf{y}_i^* , model (22) is fit and estimates for β and θ are updated.

Iterate Steps A and B until convergence.

Another justification for PQL, in the linear mixed model, when obtaining the Best Linear Unbiased Prediction (BLUP) $\hat{\beta}$ and $\hat{\mathbf{u}}$ for given θ , Henderson's mixed model equations arise from maximizing the penalized log-likelihood

$$\ln f((\mathbf{f}|\mathbf{u})) - \frac{1}{2}\mathbf{u}^T \mathbf{D}^{-1}\mathbf{u}.$$

If we maximize this penalized log-likelihood for given θ for the generalized case, the Fisher's method of scoring or Newton Raphson both lead to (22) as well.

PQL uses an approximate likelihood and is exact only for linear mixed models. It is better the closer the y_i are to normal. Thus, it works better for larger means in Poisson regression, or larger denominators in binomial proportions. The PQL estimates are inconsistent.

There exist some bias corrections for PQL (Breslow and Lin, 1995), and (Lin and Breslow, 1996). Also, the accuracy may be improved by using a second order Taylor approximation.

Hierarchical Likelihood

The hierarchical likelihood (h-likelihood) has been proposed for generalized linear models with random effects by Lee and Nelder (1996) and further studied in Lee and Nelder (2001) and Lee, Nelder and Pawitan (2006) and for non linear mixed effects models by Noh and Lee (2008). This is very similar to the Penalized Likelihood used by McGilchrist and Aisbett (1991) and Therneau and Grambsch (2000) for frailty models. The main idea is to treat the random effects as parameters and estimate all the parameters by maximizing the likelihood function conditionally on the random effects minus a penalty term which takes large values if the random parameters are highly dispersed. The advantage of this approach is that it may avoid computing numerical integrals, Commenges et al (2010).

Consider the following nonlinear mixed model

$$\mathbf{y} = f(\boldsymbol{\eta}) + \mathbf{e},$$

$$\boldsymbol{\eta} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u},$$

where \mathbf{X} is the $n \times p$ model matrix for fixed effects $\boldsymbol{\beta}$, \mathbf{Z} is the $b \times q$ model matrix for random effects \mathbf{u} , $\mathbf{u} \sim (0, \mathbf{D})$ and $\mathbf{e} \sim (0, \boldsymbol{\phi})$ with diagonal matrix $\boldsymbol{\phi}$. The h-likelihood is the log-likelihood of the observations and the unobserved random effects. Estimators (denoted by MHLE) of both $\boldsymbol{\beta}$ and \mathbf{u} obtained by maximizing the h-likelihood. Lee and Nelder (1996) introduced the h-likelihood as

$$h = l(\boldsymbol{\beta}, \boldsymbol{\phi}; \mathbf{y}|\mathbf{u}) + l(\boldsymbol{\alpha}; \mathbf{u}),$$

where $l(\boldsymbol{\alpha}; \mathbf{u})$ is the logarithm of the density function for the random effects \mathbf{u} with parameter vector $\boldsymbol{\alpha}$ and $l(\boldsymbol{\beta}, \boldsymbol{\phi}; \mathbf{y}|\mathbf{u})$ is the logarithm of the density function for the conditional response $\mathbf{y}|\mathbf{u}$ with parameters $\boldsymbol{\beta}$ and $\boldsymbol{\phi}$. Note that both \mathbf{y} and \mathbf{u} can come

from any distribution from the exponential family, which means that it is not necessary to assume normal distribution. Fixed and random effects can be estimated by solving

$$\frac{dh}{d\boldsymbol{\beta}} = \mathbf{0},$$

and

$$\frac{dh}{d\mathbf{u}} = \mathbf{0}.$$

Because the maximum h-likelihood using $\hat{\boldsymbol{\beta}}$ can bias dispersion estimators, Lee and Nelder (1996, 2001) proposed the use of the adjusted profile h-likelihood to estimate dispersion parameters. Estimate the fixed parameter $\boldsymbol{\alpha}$ using adjusted profile likelihood

$$\begin{aligned} h_A(\boldsymbol{\alpha}) &= h - \frac{1}{2} \ln \left[\det \left(\frac{\mathbf{H}}{2\pi} \right) \right] \Big|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}, \mathbf{u}=\hat{\mathbf{u}}} \\ &= h(\hat{\mathbf{u}}(\boldsymbol{\alpha}), \boldsymbol{\alpha}) - \frac{1}{2} \ln \left[\det \left(\frac{\mathbf{H}(\hat{\mathbf{u}}, \boldsymbol{\alpha})}{2\pi} \right) \right]. \end{aligned}$$

Estimate variance components using adjusted profile likelihood

$$h_A = \left(h + \frac{1}{2} \ln |2\pi \mathbf{H}^{-1}| \right) \Big|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}, \mathbf{u}=\hat{\mathbf{u}}},$$

by solving $\frac{dh_A}{d\phi} = 0$.

Maximum Hierarchical Likelihood Estimation (MHLE) $\hat{\boldsymbol{\beta}}$ are asymptotically normal and the most efficient estimators. Also, the MHLE for the random effect \mathbf{u} is asymptotically the best unbiased predictor. Finally, MHLE covariance matrix estimates are asymptotically consistent (Lee and Nelder, 1996).

Truncated Mixed Count Models

Mixed Poisson models have been used extensively in many areas such as medicine, genetics, economics and in the physical and social sciences. Gupta and Gupta (2004) modeled the Poisson model with random effect to take care of overdispersion. They

introduced an additive and a multiplicative random effect Poisson model as follows:

1. Adding an unobserved random variable u . Thus the probability mass function of the modified random variable Y is given by

$$P(Y = y|u) = \frac{e^{-(\lambda+u)}(\lambda+u)^y}{y!}, \quad y = 0, 1, 2, \dots$$

where u is considered as a random variable.

2. Multiplying by an unobserved random variable u . Thus the probability mass function of the modified random variable Y is given by

$$P(Y = y|u) = \frac{e^{-(\lambda u)}(\lambda u)^y}{y!}, \quad y = 0, 1, 2, \dots$$

where u is considered as a random variable.

In both models, the random effect was modeled by the gamma distribution and the inverse Gaussian distribution. Maximum Likelihood (ML) estimation was used to obtain the parameter estimates of the models. It was verified that the fit by the additive Poisson-Gamma model is apparently better than the fit by the Poisson model and the other mixed Poisson models. In addition, Litierie et al. (2008) verified by Monte Carlo studies that the misspecification of the random effect distribution may lead to severe bias in the random effect component prediction.

In a truncated mixed model context, the zero truncated Poisson mixed regression model has the most attention by researchers. Saei and Chambers (2005) described a random components truncated Poisson model that can be applied to clustered and zero-truncated count data. They investigated two types of models, one involving a single random component to account for between subject heterogeneity, and a second involving two random components, with the second component used to account for within subject heterogeneity as follows:

They assumed that $Y \sim Poi(\lambda_{ij})$, with distribution function

$$Y_{ij} \sim Poi(y_{ij}|\lambda_{ij}) = \frac{\exp(\lambda_{ij})\lambda_{ij}^{y_{ij}}}{y_{ij}!}.$$

Two different models for λ were considered.

1. $\lambda_{ij} = \exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + u_{1j})$, where $u_{1j} \sim N(0, \phi_1)$
2. $\lambda_{ij} = \exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + u_{1j} + \Delta_j u_{2j})$, where $\Delta_j = I(j > 1)$, and $u_i = (u_{1j}, u_{2j}) \sim N(\mathbf{0}, \boldsymbol{\Phi})$.

Log-likelihood function of the model is given by

$$l = l_1 + l_2,$$

where l_1 is the log-likelihood function of truncated Poisson observations conditional on the value of the random component vector \mathbf{u} and l_2 is the logarithm of the probability density function of \mathbf{u} .

Penalized Likelihood (PL) estimates $\hat{\boldsymbol{\beta}}$ and $\hat{\mathbf{u}}$ were obtained by maximizing $l = l_1 + l_2$ with respect to $\boldsymbol{\beta}$ and \mathbf{u} respectively. These estimates were then used as an initial step in finding Residual Maximum Likelihood (REML) estimates of ϕ_j via Henderson (1975) algorithm. Misleading inferences were obtained if the random components are ignored in model specification. Results from a simulation shows that REML method leads to regression parameter estimates that are both unbiased and efficient. However, the REML estimators of the variance components in the model tend to be negatively biased.

In another study, Nasiri (2011) reviewed the application of mixed Poisson model to the health care events where the response variable came from zero-truncated count data. The Gamma distribution for the random effect was used, which then resulted in the zero-truncated negative binomial distribution with function

$$f_0(y) = P(Y = y|Y > 0) = \frac{\Gamma(k+y)}{y!\Gamma(k)} \left(\frac{P^k}{1-P^k} \right) (1-P)^y, \quad y = 1, 2, 3, \dots$$

Method of moments (MOM) was used to estimate the parameters of the model.

In addition to the zero truncated mixed Poisson model, other types of truncation could be the case. For example, the double truncated mixed Poisson model, a combination of left and right truncation, is often appropriate for data with clusters or repeated measures. Truncated mixed Poisson models that have been discussed fall short. Saei and Chambers (2005) and also Nasiri (2011), used only zero truncated. Also Saei and Chambers (2005) use a very specific type of random effect were distributed only normal. Finally, methods estimation that have been used are not good in terms of estimating variance components. Thus, based on the review of the literature, the gap of the subject of the study has not been studied, and this research study is undertaken to fill this gap.

CHAPTER III

DOUBLE TRUNCATED MIXED POISSON REGRESSION MODEL

Introduction

Poisson regression models are commonly used to analyze count data. A frequent problem with the use of these models is that the observed variation is greater than expected and mixed Poisson models are alternative models that provide a means of explaining the extra-Poisson variation. Mixed Poisson Regression models have extensive research and literature studies, and have been commonly used in fields such as epidemiology, medicine, genetics, economics, engineering, marketing and in the physical and social sciences. Recently, many researchers have incorporated random effects into a wide variety of regression models to account for correlated responses and multiple sources of variation. In a mixed model context, Van Duijn and Bockenholt (1995) presented a latent class Poisson model for analysing overdispersed repeated count data. Gao and Thiebaut (2009) added random effects to the left truncated Longitudinal Outcomes. In this study, we formulate a class of regression models based on a double truncated Poisson regression model. For count responses, the situation of double truncation often occurs in biomedical and sociological applications. This study can be viewed as an extension of a double truncated Poisson regression model that has been studied by Suaiee and Lalonde (2013) where random effects to the model were added. Two different distributions for the random effect were studied through simulation. The misspecification of these distributions was addressed. Comparison with Left truncated mixed Poisson model and regular mixed Poisson model was presented.

The difficulty of parameter estimation in mixed models is well known. Several approaches have been considered to evaluate the loglikelihood. Various authors have considered analytic approximations such as the Laplace approximation (e.g., Wolfinger, 1993; Steele, 1996) to motivate fitting algorithms or estimating equations. An Expectation Maximization (EM) based estimation approach for mixed effects models when the outcomes are truncated was proposed by Hughes (1999). In addition, the REML estimation method was used in generalized mixed models by McGilchrist (1994). The Lack of this method that the estimate of variance component could be negative. In this study, we adopt the h-likelihood estimation method that was proposed by Lee and Nelder, (1996) and was discussed in chapter two, where we drop the assumption of normality on the random effect.

The chapter is organized as follows: Based on the first research question, the author constructs the double truncated mixed Poisson regression model in the second section. In the third section, the author outlines the h-likelihood estimation method for the model with Normal random effect. Section four presents h-likelihood estimation method for the model with Gamma random effect. Both section three and section four answer the second research question. At the end of this chapter, the author gives a brief description of the simulation study that is performed in chapter four and it is answering the research questions three, four, and five.

The Double Truncated Mixed Poisson Regression Model

Double truncated Poisson model with mixed effects are often appropriate for data with clusters or repeated measures. The mixed Poisson distribution and negative binomial distribution have been widely used in situations where counts display overdispersion. A common model that is used for analyzing count response data is the Poisson regression model (McCullagh and Nelder, 1989). However, in our case, these data are truncated at c_l from the lower end and c_r from the upper end. Thus, this distribution of Poisson count

data is called the double truncated Poisson. Besides the generality of the double truncated Poisson model, it is more realistic in applications since in many instances the data are confined to an interval so that no data are observed outside the interval (c_l, c_r) , (Kedem and Fokianos, 2002).

Let $Y_{ij}(i = 1, \dots, m, j = 1, \dots, n_i)$ be the observations of the response variable. Let u_i be the unobserved random effect on the i^{th} individual. We consider the model

$$Pr(Y_{ij} = y_{ij} | u_i, c_l \leq y_{ij} \leq c_r) = \frac{\lambda_i^{y_{ij}}}{y_{ij}! \sum_{k=c_l}^{c_r} \frac{\lambda_i^k}{k!}}, \quad (23)$$

such that

$$\lambda_{ij} = \exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i).$$

In general matrix notation, we write it as

$$\boldsymbol{\lambda} = \exp(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}),$$

where

\mathbf{X} is a design matrix of fixed effects explanatory variables,

$\boldsymbol{\beta}$ is a vector of fixed effect coefficients,

\mathbf{Z} is a design matrix of random effects explanatory variables,

\mathbf{u} is a vector of random effects.

We assume two cases for random effect.

Case 1:

$$u_i \sim Normal(0, \sigma^2). \quad (24)$$

Assuming a normal distribution for the random effects is a routine assumption for mixed models, which is convenient from a computational point of view. However, it could

be rather restrictive. In general, incorrect distribution assumption for the random effects has unfavorable influence on statistical inferences (Agresti, Caffo, and Ohman-Strickland, 2004). Thus, this motivates the search for mixed models with more flexible distribution for the random effect. Hence, we have

Case 2:

$$u_i \sim \Gamma(\alpha_1, \alpha_2). \quad (25)$$

H-likelihood Estimation for Normal Random Effect

For the model (23), we have that

$$E[Y_{ij} = y_{ij} | u_i, c_l \leq y_{ij} \leq c_r] = \frac{\sum_{k=c_l}^{c_r} \frac{\lambda_i^k}{(k-1)!}}{\sum_{k=c_l}^{c_r} \frac{\lambda_i^k}{k!}}.$$

Under the assumption of (24), the h-likelihood (h) is defined by

$$h = l_1(\boldsymbol{\beta}; \mathbf{y} | \mathbf{u}) + l_2(\sigma^2, \mathbf{u}),$$

where $l_1(\boldsymbol{\beta}; \mathbf{y} | \mathbf{u})$ is the logarithm of the conditional Poisson density function for the response \mathbf{Y} given \mathbf{u} with parameter $\boldsymbol{\lambda} = \exp(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u})$, and $l_2(\sigma^2, \mathbf{u})$ is the logarithm of the Normal density function for the random effect \mathbf{u} . Thus,

$$\begin{aligned} l_1(\boldsymbol{\beta}; \mathbf{y} | \mathbf{u}) &= \sum_{ij} \left[y_{ij} \log(\lambda_{ij}) - \log(y_{ij}!) - \log \left(\sum_{k=c_l}^{c_r} \frac{\lambda_{ij}^k}{k!} \right) \right] \\ &= \sum_{ij} \left[y_{ij} (\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i) - \log \left(\sum_{k=c_l}^{c_r} \frac{\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i)^k}{k!} \right) - \log(y_{ij}) \right], \end{aligned}$$

and

$$l_2(\mathbf{u}) = - \sum_i \left[\frac{1}{2} \log(2\pi) + \frac{1}{2} \log(\sigma^2) + \frac{u_i^2}{2\sigma^2} \right].$$

Given the dispersion components σ^2 , the maximum h-likelihood estimators (MHLEs) of $\tau = (\boldsymbol{\beta}, \mathbf{u})$ with $\mathbf{u} = (u_1, \dots, u_m)^T$ are obtained by solving

$$\frac{\partial h}{\partial \beta_l} = \sum_{ij} \left[y_{ij} - \frac{1}{\sum_{k=c_l}^{c_r} \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{k!}} \sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{(k-1)!} \right] x_{ijl} = 0, \quad l = 1, \dots, p$$

and

$$\frac{\partial h}{\partial u_i} = \sum_j \left[y_{ij} - \frac{1}{\sum_{k=c_l}^{c_r} \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{k!}} \sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{(k-1)!} \right] z_{ij} - \frac{u_i}{\sigma^2} = 0, \quad i = 1, \dots, m$$

Parameters of the model can be estimated iteratively using the Fisher scoring or Newton-Raphson method. For the estimation of the dispersion parameter σ^2 given the estimator of $\hat{\tau}$, we use Lee and Nelder's (1996) adjusted profile h-likelihood, defined by

$$h_A = h - \frac{1}{2} \ln \left[\det \left(\frac{\mathbf{H}}{2\pi} \right) \right] \Big|_{\tau=\hat{\tau}},$$

such that

$$\mathbf{H} = \frac{\partial^2 h}{\partial \tau \partial \tau'} \Big|_{\tau=\hat{\tau}} = \begin{bmatrix} \frac{\partial^2 h}{\partial \beta \partial \beta'} & \frac{\partial^2 h}{\partial \beta \partial u'} \\ \frac{\partial^2 h}{\partial u \partial \beta'} & \frac{\partial^2 h}{\partial u \partial u'} \end{bmatrix}.$$

Here

$$\begin{aligned} \frac{\partial^2 h}{\partial \beta_l \partial \beta_{l'}} &= - \sum_{ij} x_{ijl} x_{ijl'} \left[\frac{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{k!} \right) \left(\sum_k \frac{k(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{(k-1)!} \right)}{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{k!} \right)^2} \right] \\ &\quad - \sum_{ij} x_{ijl} x_{ijl'} \left[\frac{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{(k-1)!} \right)^2}{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{k!} \right)^2} \right], \quad l = 1, \dots, p \end{aligned}$$

and

$$\begin{aligned} \frac{\partial^2 h}{\partial \beta_l \partial \beta_{l'}} &= - \sum_{ij} x_{ijl} z_{ij} \left[\frac{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{k!} \right) \left(\sum_k \frac{k(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{(k-1)!} \right)}{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{k!} \right)^2} \right] \\ &\quad - \sum_{ij} x_{ijl} z_{ij} \left[\frac{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{(k-1)!} \right)^2}{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{k!} \right)^2} \right], \quad l = 1, \dots, p \end{aligned}$$

and

$$\begin{aligned} \frac{\partial^2 h}{\partial \beta_l \partial \beta_{l'}} &= - \sum_{ij} z_{ij} x_{ijl} \left[\frac{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{k!} \right) \left(\sum_k \frac{k(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{(k-1)!} \right)}{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{k!} \right)^2} \right] \\ &\quad - \sum_{ij} z_{ij} x_{ijl} \left[\frac{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{(k-1)!} \right)^2}{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{k!} \right)^2} \right], \quad l = 1, \dots, p \end{aligned}$$

and

$$\begin{aligned} \frac{\partial^2 h}{\partial \beta_l \partial \beta_{l'}} = & - \sum_{ij} z_{ij} z_{ij'} \left[\frac{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{k!} \right) \left(\sum_k \frac{k(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{(k-1)!} \right)}{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{k!} \right)^2} \right] \\ & - \sum_{ij} z_{ij} z_{ij'} \left[\frac{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{(k-1)!} \right)^2}{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{k!} \right)^2} \right] - \frac{m}{\sigma^2}, \quad i = 1, \dots, m \end{aligned}$$

The maximum adjusted profile h-likelihood estimators for σ^2 are obtained by solving iteratively

$$\frac{\partial h_A}{\partial \sigma^2} \Big|_{\tau=\hat{\tau}} = 0,$$

where $\hat{\tau}$ is re-evaluated at each iteration.

H-likelihood Estimation for Gamma Random Effect

For the the model (23), we have that

$$E[Y_{ij} = y_{ij} | u_i, c_l \leq y_{ij} \leq c_r] = \frac{\sum_{k=c_l}^{c_r} \frac{\lambda_i^k}{(k-1)!}}{\sum_{k=c_l}^{c_r} \frac{\lambda_i^k}{k!}}.$$

Under the assumption of (25), the h-likelihood (h) is defined by

$$h = l_1(\boldsymbol{\beta}; \mathbf{y} | \mathbf{u}) + l_2(\boldsymbol{\theta}, \mathbf{u}),$$

where $l_1(\boldsymbol{\beta}; \mathbf{y} | \mathbf{u})$ is the logarithm of the conditional Poisson density function for the response \mathbf{Y} given \mathbf{u} with parameter $\boldsymbol{\lambda} = \exp(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u})$, and $l_2(\boldsymbol{\theta}, \mathbf{u})$ is the logarithm of

the Gamma density function for the random effect \mathbf{u} with parameter $\boldsymbol{\theta} = (\alpha_1, \alpha_2)$. Thus,

$$\begin{aligned} l_1(\boldsymbol{\beta}; y|\mathbf{u}) &= \sum_{ij} \left[y_{ij} \log(\lambda_{ij}) - \log(y_{ij}!) - \log \left(\sum_{k=c_l}^{c_r} \frac{\lambda_{ij}^k}{k!} \right) \right] \\ &= \sum_{ij} \left[y_{ij} (\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i) - \log \left(\sum_{k=c_l}^{c_r} \frac{\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i)^k}{k!} \right) - \log(y_{ij}) \right], \end{aligned}$$

and

$$l_2(\boldsymbol{\theta}, \mathbf{u}) = - \sum_i \left[\alpha_1 \log(\alpha_2) + \log(\Gamma(\alpha_1)) - (\alpha_1 - 1) \log(u_i) + \frac{u_i}{\alpha_2} \right].$$

Giving the dispersion components $\boldsymbol{\theta} = (\alpha_1, \alpha_2)$, the maximum h-likelihood estimators (MHLEs) of $\tau = (\boldsymbol{\beta}, \mathbf{u})$ with $\mathbf{u} = (u_1, \dots, u_m)^T$ are obtained by solving

$$\frac{\partial h}{\partial \beta_l} = \sum_{ij} \left[y_{ij} - \frac{1}{\sum_{k=c_l}^{c_r} \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{k!}} \sum_k \frac{((\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k)}{(k-1)!} \right] x_{ijl} = 0, \quad l = 1, \dots, p$$

and

$$\begin{aligned} \frac{\partial h}{\partial u_i} &= \sum_j \left[y_{ij} - \frac{1}{\sum_{k=c_l}^{c_r} \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{k!}} \sum_k \frac{((\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k)}{(k-1)!} \right] \\ &+ \left[\frac{(\alpha_1 - 1)}{u_i} - \frac{1}{\alpha_2} \right] = 0, \quad i = 1, \dots, m. \end{aligned}$$

Parameters of the model can be estimated iteratively using the Fisher scoring or Newton-Raphson method. For the estimation of the dispersion parameter θ given the estimator of $\hat{\tau}$, we use Lee and Nelder's (1996) adjusted profile h-likelihood, defined by

$$h_A = h - \frac{1}{2} \ln \left[\det \left(\frac{\mathbf{H}}{2\pi} \right) \right] |_{\tau=\hat{\tau}},$$

Such that

$$\mathbf{H} = \frac{\partial^2 h}{\partial \tau \partial \tau'} \Big|_{\tau=\hat{\tau}} = \begin{bmatrix} \frac{\partial^2 h}{\partial \beta \partial \beta'} & \frac{\partial^2 h}{\partial \beta \partial u'} \\ \frac{\partial^2 h}{\partial u \partial \beta'} & \frac{\partial^2 h}{\partial u \partial u'} \end{bmatrix}.$$

Here

$$\begin{aligned} \frac{\partial^2 h}{\partial \beta_l \partial \beta_{l'}} &= - \sum_{ij} x_{ijl} x_{ijl'} \left[\frac{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{k!} \right) \left(\sum_k \frac{k (\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{(k-1)!} \right)}{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{k!} \right)^2} \right] \\ &\quad - \sum_{ij} x_{ijl} x_{ijl'} \left[\frac{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{(k-1)!} \right)^2}{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{k!} \right)^2} \right], \quad l = 1, \dots, p \end{aligned}$$

and

$$\begin{aligned} \frac{\partial^2 h}{\partial \beta_l \partial \beta_{l'}} &= - \sum_{ij} x_{ijl} z_{ij} \left[\frac{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{k!} \right) \left(\sum_k \frac{k (\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{(k-1)!} \right)}{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{k!} \right)^2} \right] \\ &\quad - \sum_{ij} x_{ijl} z_{ij} \left[\frac{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{(k-1)!} \right)^2}{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{z}_{ij}^T \mathbf{u}_i))^k}{k!} \right)^2} \right], \quad l = 1, \dots, p \end{aligned}$$

and

$$\begin{aligned} \frac{\partial^2 h}{\partial \beta_l \partial \beta_{l'}} = & - \sum_{ij} z_{ij} x_{ijl} \left[\frac{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + z_{ij}^T \mathbf{u}_i))^k}{k!} \right) \left(\sum_k \frac{k(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + z_{ij}^T \mathbf{u}_i))^k}{(k-1)!} \right)}{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + z_{ij}^T \mathbf{u}_i))^k}{k!} \right)^2} \right] \\ & - \sum_{ij} z_{ij} x_{ijl} \left[\frac{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + z_{ij}^T \mathbf{u}_i))^k}{(k-1)!} \right)^2}{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + z_{ij}^T \mathbf{u}_i))^k}{k!} \right)^2} \right], \quad l = 1, \dots, p \end{aligned}$$

and

$$\begin{aligned} \frac{\partial^2 h}{\partial \beta_l \partial \beta_{l'}} = & - \sum_{ij} z_{ij} z_{ij'} \left[\frac{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + z_{ij}^T \mathbf{u}_i))^k}{k!} \right) \left(\sum_k \frac{k(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + z_{ij}^T \mathbf{u}_i))^k}{(k-1)!} \right)}{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + z_{ij}^T \mathbf{u}_i))^k}{k!} \right)^2} \right] \\ & - \sum_{ij} z_{ij} z_{ij'} \left[\frac{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + z_{ij}^T \mathbf{u}_i))^k}{(k-1)!} \right)^2}{\left(\sum_k \frac{(\exp(\mathbf{x}_{ij}^T \boldsymbol{\beta} + z_{ij}^T \mathbf{u}_i))^k}{k!} \right)^2} \right] - \frac{(\alpha_1 - 1)}{u_i^2}, \quad i = 1, \dots, m \end{aligned}$$

The maximum adjusted profile h-likelihood estimators for θ are obtained by solving iteratively

$$\frac{\partial h_A}{\partial \theta} \Big|_{\tau=\hat{\tau}} = 0,$$

where $\hat{\tau}$ is re-evaluated at each iteration.

In next chapter, we present the simulation study of the models from this chapter. A numerical study, based upon 10,000 replications of simulated data, will be presented to evaluate the performance of the proposed estimation procedure. All the procedures from data generation to the model fitting will be done by using R-Package. MaxLik library will be used to solve h-likelihood function based on the Newton-Raphson method. Also, gamlss.tr Library will be used to generate double truncated count data. The simulation

will consist of the following steps:

Step 1: Generate the covariate variables (x_1, x_2) from uniform (1,3).

Step 2: Generate the random effect u . we consider two distributions for generating u , Normal($0, \sigma^2$) and Gamma(α_1, α_2). The following situations in context of the model will be presented:

- Random Effect Distributed Normal

Table 1
Generated Double Truncated with Normal Random Effect

Model	Distribution	
	Correct	Incorrect
Double Truncated	Normal	Gamma
Left Truncated	Normal	Gamma
Untruncated	Normal	Gamma

- Random Effect Distributed Gamma

Table 2
Generated Double Truncated with Gamma Random Effect

Model	Distribution	
	Correct	Incorrect
Double Truncated	Gamma	Normal
Left Truncated	Gamma	Normal
Untruncated	Gamma	Normal

Step 3: Define $\lambda = \exp(0.1 + 0.4x_1 + u)$ and then generate y from double truncated Poisson with mean λ .

Step 4: Fit the model and get the parameter estimates using h-likelihood method.

Step 5: Repeat the previous steps 10,000 times, then calculate the mean and standard deviation of each parameter estimates as well as the bias of the estimates.

The same generated data will be fitted using left truncated mixed Poisson regression model and regular mixed Poisson regression model. Comparison between the three models will be done in terms of parameter estimations, the bias of the estimates, and standard deviations of the parameters. Misspecification of the random effect distribution will be discussed. Finally, a real data set will be used to illustrate our model.

CHAPTER IV

SIMULATION STUDY

This chapter presents the simulation study to evaluate the models that were presented in the previous chapter. The main purpose of this simulation is to examine the impact of the double truncated mixed Poisson (DTMP) regression model compared to the left truncated mixed Poisson (LTMP) regression model and the regular mixed Poisson (RMP) regression model when we have clustered double truncated count data. The performance of each model is examined based on the power of the significant covariate, Type I error rate of the non-significant covariate, parameter estimates and their standard errors as well as the bias of the estimates. Two different distributions for the random effects were considered for each model, a normal distribution and a gamma distribution. Finally, a numerical example is presented to demonstrate the model.

The chapter is organized as follows: In the first section, a simulation study for the DTMP regression model with Normal random effects is presented. Second section presents a simulation study for the DTMP regression model with Gamma random effects. Third section, presents the DTMP regression model with the distribution of random effects miss-specified as Normal. In the fourth section, the DTMP regression model with the distribution of random effects miss-specified as Gamma. Finally, fifth section presents the performance of the models on a real data set.

Simulation Study for the Double Truncated Mixed Poisson Regression Data with Normal Random Effects

Consider the model

$$y_{ij}|u_i \sim \text{Poisson}(\lambda_i),$$

$$\lambda_{ij} = \exp(\beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} + u_i),$$

$$u_i \sim \text{Normal}(\text{mean} = 0, \text{Variance} = \sigma^2).$$

The simulation consists of the following steps:

Step 1: Generate two covariate vectors $(\mathbf{x}_1, \mathbf{x}_2)$ from uniform (0,3) with length N_i each.

Step 2: Fifteen random effects u_i representing 15 clusters are generated each from Normal (0,0.2).

Step 3: Define $\lambda_{ij} = \exp(\beta_0 + \beta_1 x_{1ij} + u_i)$ and then generate the response variable y_{ij} from the double truncated Poisson distribution with mean λ_{ij} using the function `trun.tr` under the `gamless` library in R package with $\beta_0 = 0.1$ and $\beta_1 = 0.4$. The values of y_{ij} are truncated by 2 on the left and by 8 on the right. Here generating y_{ij} depends on x_{1ij} but not x_{2ij} , so we expect x_{1ij} to be significant and x_{2ij} to be not significant, and hence calculate the power and Type I error rate.

Step 4: Using the h-likelihood estimation method, fit the double truncated Poisson regression model with Normal random effects.

Step 5: Get the parameter estimates of the model $\hat{\beta}$ and \hat{u} . Based on these estimates and initial value of $\sigma = 0.9$, the adjusted profile h-likelihood method was used to estimate the dispersion parameter σ .

Step 6: Repeat the previous steps 10,000 times, then calculate the mean of each parameter estimate and the average of the standard errors as well as the bias for the regression coefficients. The Power and Type I error rate are also calculated as follow:

$$\text{Power} = \frac{\text{Number of times } \hat{\beta}_1 \text{ is significant}}{10000},$$

$$\text{Type I Error Rate} = \frac{\text{Number of times } \hat{\beta}_2 \text{ is significant}}{10000}.$$

The same generated data is fitted using the left truncated Poisson regression model with Normal random effects. In this model, we have double truncated count data and the right truncation was ignored and only the left truncation was considered. Also, the same generated data are fitted using the regular Poisson regression model with Normal random effects. In this model, we have double truncated count data and both the left and the right truncation were ignored.

The simulation results for the three models are given in Table (3). The table presents the mean of the parameter estimates and the average of the estimated standard errors as well as the bias for the regression coefficients for each model. The power and Type I rate for each model are presented in Table (4).

From Table (3), in the case of mixed double truncated Poisson model and mixed left truncated Poisson model, the parameter estimate $\hat{\beta}_1$ is within one standard error of the true value and the bias is getting close to zero as cluster size increases. The standard error of $\hat{\beta}_1$ decreases with increasing cluster size. The regular mixed Poisson model is the worse in recovering the true parameter value β_1 compared to mixed double truncated Poisson model and mixed left truncated Poisson model. The standard errors of the estimated parameters $\hat{\beta}_1$ for the mixed double truncated Poisson model and the mixed left truncated Poisson model are generally larger than the estimates obtained in the regular mixed Poisson model. The parameter estimate $\hat{\beta}_2$ is very close to zero for the three models as expected since the generated response variable depends only on x_1 but not x_2 .

The dispersion parameter estimate $\hat{\sigma}$ appears to be asymptotically unbiased for the mixed double truncated Poisson model and the mixed left truncated Poisson model and the bias is getting close to zero as cluster size is increased. For the regular mixed Poisson model the estimate $\hat{\sigma}$ is biased. The results from these three simulations for different cluster sizes reveal that the mixed double truncated Poisson regression model and the mixed left truncated Poisson regression model assuming Normal random effects provide

similar and better recovery of the true parameter values compared to the regular mixed Poisson regression model with Normal random effects.

Table 3
Double Truncated Poisson Regression Model With Normal Random Effects

Sample Size	Cluster Size	Parameter	Double Truncated Poisson Model			Left Truncated Poisson Model			Regular Poisson Model		
			Estimate (S.E.)	Bias		Estimate (S.E.)	Bias		Estimate (S.E.)	Bias	
60	4	β_1	0.4064	0.1456	0.0064	0.3915	0.1415	-0.0085	0.1668	0.0881	-0.2332
		β_2	0.0003	0.1370	0.0003	0.0004	0.1332	0.0004	-0.001	0.0876	-0.001
		σ	0.9867		0.0867	0.9901		0.0901	0.2503		-0.5497
75	5	β_1	0.4024	0.1287	0.0024	0.3935	0.1253	-0.0065	0.1668	0.0784	-0.2332
		β_2	-0.0002	0.1207	-0.0002	-0.0009	0.1178	-0.0009	-0.0003	0.0778	-0.0003
		σ	0.9465		0.0465	0.9527		0.0527	0.2197		-0.5803
90	6	β_1	0.4026	0.1167	0.0026	0.3924	0.1140	-0.0076	0.1667	0.0713	-0.2333
		β_2	-0.0007	0.1094	-0.0007	0.0002	0.1072	0.0002	-0.0001	0.0709	-0.0001
		σ	0.9165		0.0165	0.9288		0.0288	0.1816		-0.6184
105	7	β_1	0.3996	0.1078	-0.004	0.3914	0.1049	-0.0086	0.1667	0.0659	-0.2333
		β_2	-0.0007	0.1011	-0.0007	0.0002	0.0986	0.0002	-0.0004	0.0655	-0.0004
		σ	0.8963		-0.0037	0.9035		0.0035	0.1446		-0.6554

Table 4
Power and Type I Error Rate for Models with Normal Random Effects

Sample Size	Cluster Size		Double Truncated Poisson Model	Left Truncated Poisson Model	Regular Poisson Model
60	4	Power	0.8134	0.8122	0.4699
		Type I Rate	0.0534	0.0471	0.0031
75	5	Power	0.8843	0.8947	0.6009
		Type I Rate	0.0534	0.0491	0.0017
90	6	Power	0.9398	0.9383	0.7118
		Type I Rate	0.0526	0.0477	0.0037
105	7	Power	0.9669	0.9672	0.7973
		Type I Rate	0.0523	0.0473	0.0035

From Table (4), the Power of mixed double truncated Poisson model with Normal random effects and mixed left truncated Poisson model with Normal random effects are very similar and increasing by increasing the cluster size. The power of the regular mixed Poisson model with Normal random effects is very low compared to the other models as shown in Figure (1). The Type I error rates for double and left truncated mixed Poisson model are close to nominal value of 0.05 for all different cluster sizes as expected . Meanwhile, the Type I error rate for the regular mixed Poisson model is very conservative around zero as shown in Figure (2).

Figure (3) shows a sample of the histograms of the estimated random effect values for the three models. The histograms seem to be Normally distributed as expected.

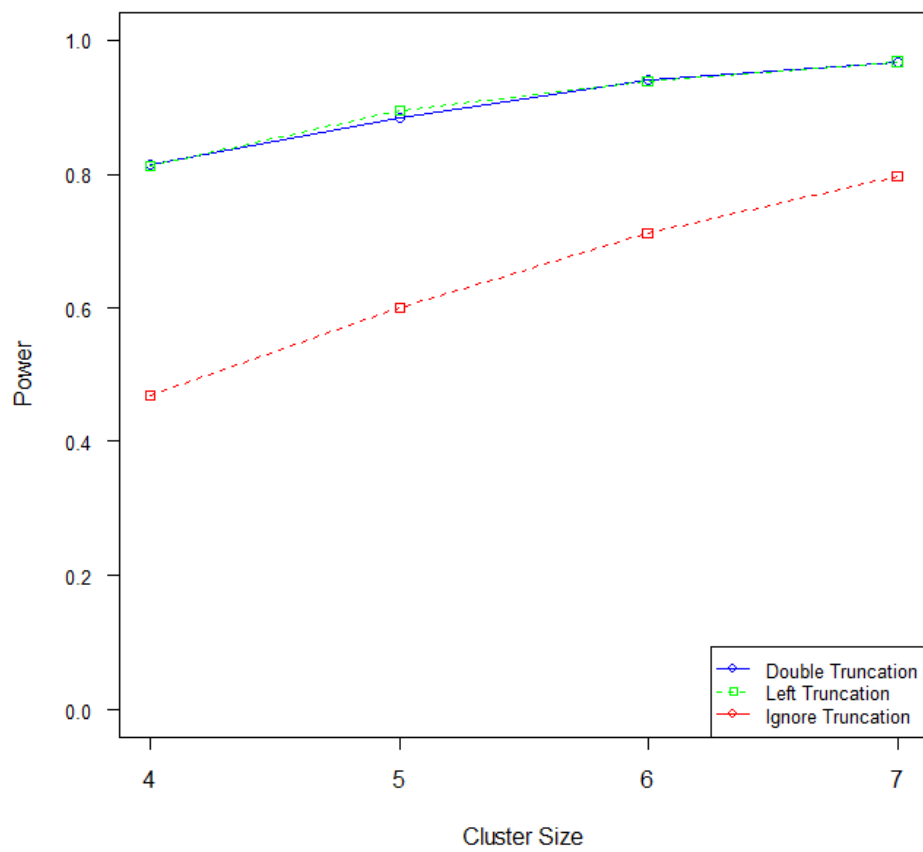


Figure 1 Power of Double Truncated Poisson Model with Normal Random Effects

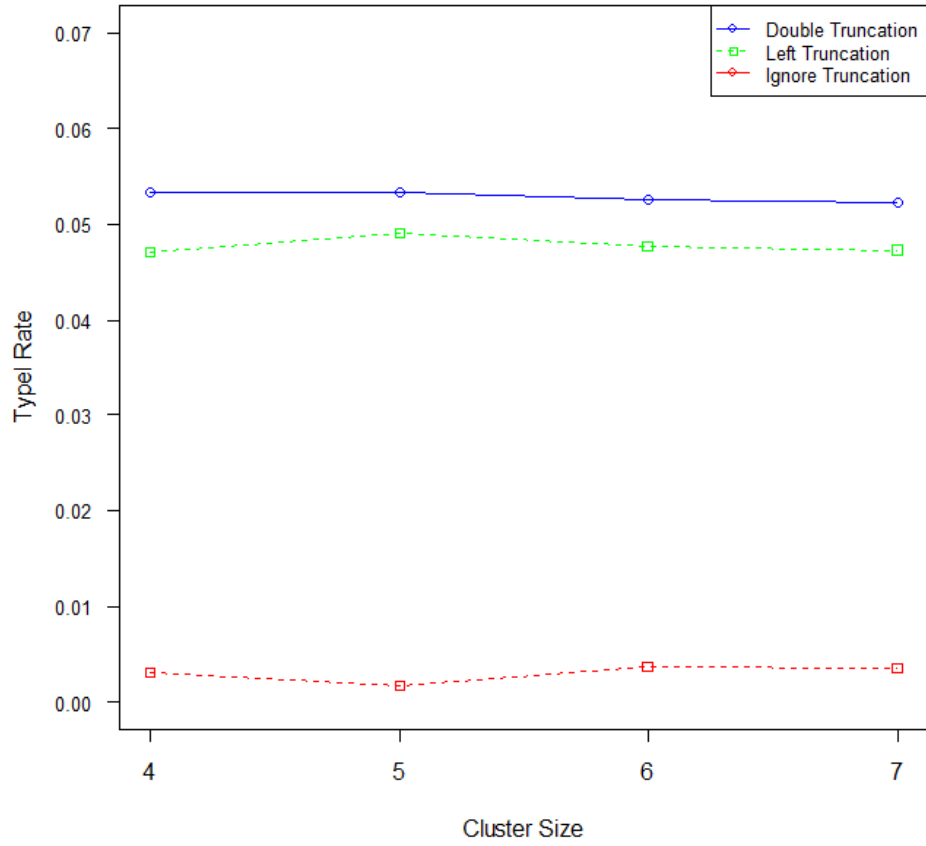


Figure 2 Type I rate of Double Truncated Poisson Model with Normal Random Effects

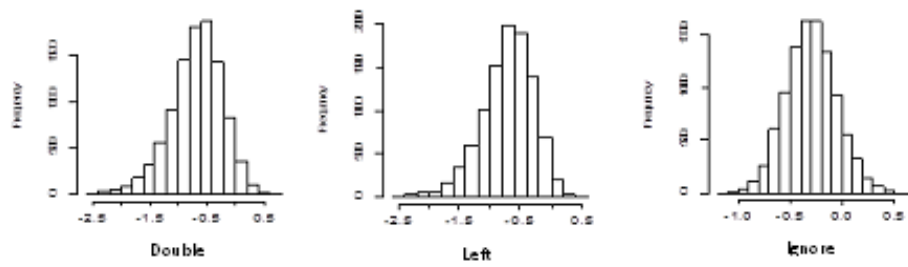


Figure 3 Histogram of the estimated Random effect

Simulation Study for the Double Truncated Mixed Poisson Regression Data with Gamma Random Effects

Consider the model

$$y_{ij}|u_i \sim \text{Poisson}(\lambda_{ij}),$$

$$\lambda_{ij} = \exp(\beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} + u_i),$$

$$u_i \sim \text{gamma}(\alpha_1, \alpha_2).$$

The simulation consists of the following steps:

Step 1: Generate two covariate vectors $(\mathbf{x}_1, \mathbf{x}_2)$ from Uniform (0,3) with length N_i each.

Step 2: Fifteen random effects u_i representing 15 clusters are generated each from Gamma(shape=1, scale=0.2).

Step 3: Define $\lambda_{ij} = \exp(\beta_0 + \beta_1 x_{1ij} + u_i)$ and then generate the response variable y_{ij} from the double truncated Poisson distribution with mean λ_{ij} using the function `trun.tr` under the `gamless` library in R package with $\beta_0 = 0.1$ and $\beta_1 = 0.4$. The values of y_{ij} are truncated by 2 on the left and by 8 on the right. Here generating y_{ij} depends on x_{1ij} but not x_{2ij} , so we expect x_{1ij} to be significant and x_{2ij} to be not significant, and hence calculate the power and Type I error rate.

Step 4: Using the h-likelihood estimation method, fit the double truncated Poisson regression model with Gamma random effects.

Step 5: Get the parameter estimates of the model $\hat{\beta}$ and \hat{u} . Based on these estimates and initial values of $\alpha_1 = 3$ and $\alpha_2 = 0.02$, the adjusted profile h-likelihood method was used to estimate the dispersion parameters α_1 and α_2 .

Step 6: Repeat the previous steps 10,000 times, then calculate the mean of each parameter's estimates and the average of the standard errors as well as the bias for the regression coefficients. The Power and Type I error rate are also calculated as follow:

$$\text{Power} = \frac{\text{Number of times } \hat{\beta}_1 \text{ is significant}}{10000},$$

$$\text{Type I Error Rate} = \frac{\text{Number of times } \hat{\beta}_2 \text{ is significant}}{10000}.$$

The same generated data is fitted using the left truncated Poisson regression model with Gamma random effects. In this model, we have double truncated count data. The right truncation was ignored and only the left truncation was considered. Also, the same generated data are fitted using the regular Poisson regression model with Gamma random effects. In this model, we have double truncated count data and both the left and the right truncation were ignored.

The simulation results for the three models are given in Table (5). The table presents the mean of the parameter estimates and the average of the estimated standard errors as well as the bias for the regression coefficients for each model. The power and Type I error rate for each model are presented in Table (6).

From Table (5), in the case of mixed double truncated Poisson model and mixed left truncated Poisson model, the parameter estimate $\hat{\beta}_1$ is within one standard error of the true value and the bias is getting close to zero as cluster size increases. The standard error of $\hat{\beta}_1$ decreases with increasing cluster size. The regular mixed Poisson model is the worse in recovering the true parameter value β_1 compared to mixed double truncated Poisson model and mixed left truncated Poisson model. The standard errors of the estimated parameters $\hat{\beta}_1$ for the mixed double truncated Poisson model and the mixed left truncated Poisson model are generally larger than the estimates obtained in the regular mixed Poisson model. The parameter estimate $\hat{\beta}_2$ is very close to zero for the three models as expected since the generated response variable depends only on x_1 but not x_2 .

The dispersion parameter estimate $\hat{\alpha}_1$ is almost unbiased for the DTMP regression model when the cluster size is 5 and as the cluster size increases $\hat{\alpha}_1$ becomes negatively biased. $\hat{\alpha}_2$ appears to be asymptotically unbiased and the bias is getting close to zero as the cluster size increases. For the LTMP regression model, the dispersion parameters $\hat{\alpha}_1 = 2.9010$ with bias=-0.0990 and $\hat{\alpha}_2 = 0.0216$ with bias=0.0016 for cluster size of 4. For larger cluster size 5, 6 or 7, the LTMP model failed to converge. The regular mixed

Poisson model is similar to the DTMP model in terms of recovering the true dispersion parameters for cluster size of 5 and the model failed to converge for larger cluster sizes. The results from these three simulations reveal that the DTMP regression model assuming Gamma random effects provides the best recovery of the true parameter values compared to the LTMP regression model and the regular mixed Poisson regression model.

Table 5
Double Truncated Poisson Regression Model With Gamma Random Effects

Sample Size	Cluster Size	Parameter	Double Truncated Poisson Model			Left Truncated Poisson Model			Regular Poisson Model		
			Estimate (S.E.)		Bias	Estimate (S.E.)		Bias	Estimate (S.E.)		Bias
60	4	β_1	0.4082	0.1362	0.0082	0.3891	0.1313	-0.0109	0.1818	0.0861	-0.2182
		β_2	-0.0003	0.1278	-0.0003	-0.0009	0.1233	-0.0009	0.00001	0.0855	0.00001
		α_1	3.3399		0.3399	2.9010		-0.0990	3.3201		0.3201
		α_2	0.0241		0.0041	0.0216		0.0016	0.0244		0.0044
75	5	β_1	0.4015	0.1205	0.0015				0.1824	0.0761	-0.2176
		β_2	-0.0013	0.1131	-0.0013				0.0009	0.0761	0.0009
		α_1	3.0807		0.0807				3.0609		0.0609
		α_2	0.0226		0.0026				0.0228		0.0028
90	6	β_1	0.4026	0.1095	0.0026						
		β_2	-0.00019	0.1027	-0.00019						
		α_1	2.8416		-0.1594						
		α_2	0.0212		0.0012						
105	7	β_1	0.4006	0.1007	0.0006						
		β_2	0.0009	0.0948	0.0009						
		α_1	2.6352		-0.3648						
		α_2	0.0201		0.0009						

Table 6
Power and Type I Error Rate for Models with Gamma Random Effects

Sample Size	Cluster Size		Double Truncated Poisson Model	Left Truncated Poisson Model	Regular Poisson Model
60	4	Power	0.8708	0.8630	0.6002
		Type I Rate	0.0478	0.0456	0.0036
75	5	Power	0.9283		0.7292
		Type I Rate	0.0518		0.0035
90	6	Power	0.9658		
		Type I Rate	0.0475		
105	7	Power	0.9831		
		Type I Rate	0.0457		

From Table (6), The Power of the DTMP regression model with Gamma random effects is a little bit higher than the power of the LTMP regression model with Gamma random effects for cluster size of 4, and it is much higher than the power of regular mixed Poisson model with Gamma random effects as shown in Figure (4). The Type I error rates for the double and left truncated mixed Poisson models are close to the nominal value of 0.05 for cluster size of 4 as expected. Meanwhile, the Type I error rate for the regular mixed Poisson model is very conservative around zero as shown in Figure (5).

Figure (6) shows a sample of the histograms of the estimated random effect values for the three models. The histograms seem to be Gamma distributed as expected.

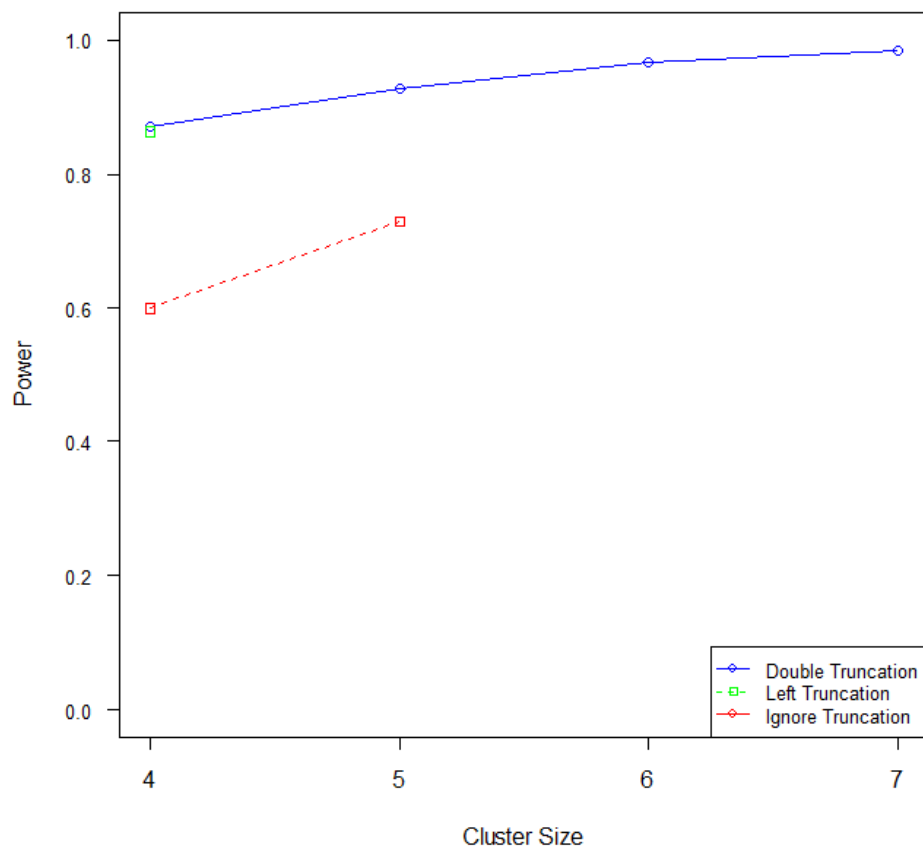


Figure 4 Power of Double Truncated Poisson Model with Gamma Random Effects

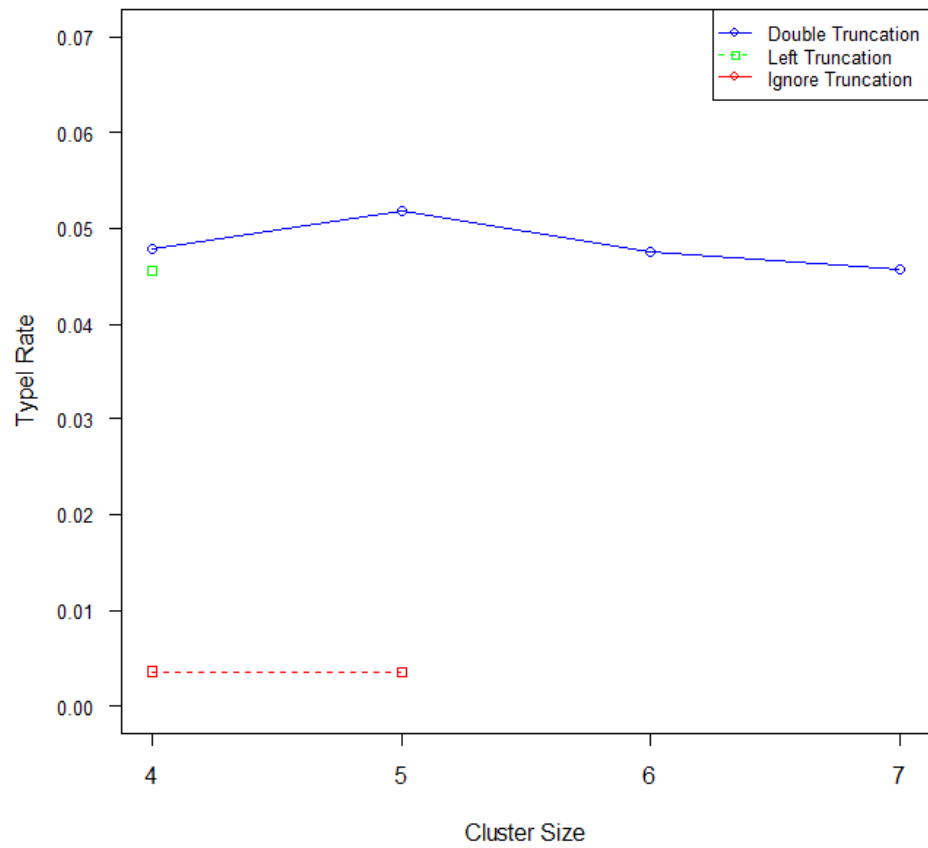


Figure 5 Type I rate of Double Truncated Poisson Model with Gamma Random Effects

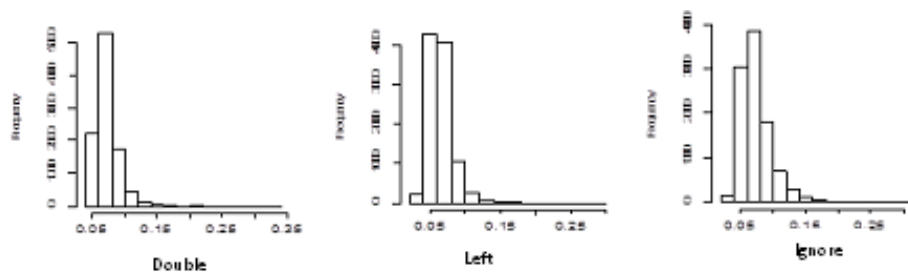


Figure 6 Histogram of the estimated Random effect

Simulation Study for the Double Truncated Mixed Poisson Regression Data with Mis-specified Normal Random Effects

In this section, the same steps 1-6 from the first section are repeated. In this case for step 2 instead of generating the random effects from a Normal distribution, the random effects are generated from $\text{Gamma}(\text{shape}=1, \text{scale}=0.2)$.

The simulation consists of the following steps:

Step 1: Generate two covariate vectors $(\mathbf{x}_1, \mathbf{x}_2)$ from Uniform (0,3) with length N_i each.

Step 2: Fifteen random effects u_i representing 15 clusters are generated each from $\text{Gamma}(\text{shape}=1, \text{scale}=0.2)$.

Step 3: Define $\lambda_{ij} = \exp(\beta_0 + \beta_1 x_{1ij} + u_i)$ and then generate the response variable y_{ij} from the double truncated Poisson distribution with mean λ_{ij} using the function `trun.tr` under the `gamless` library in R package with $\beta_0 = 0.1$ and $\beta_1 = 0.4$. The values of y_{ij} are truncated by 2 on the left and by 8 on the right.

Step 4: Using the h-likelihood estimation method, fit the double truncated Poisson regression model with Normal random effects.

Step 5: Get the parameter estimates of the model $\hat{\beta}$ and $\hat{\mathbf{u}}$. Based on these estimates and initial values of $\sigma = 0.9$, the adjusted profile h-likelihood method was used to estimate the dispersion parameter σ .

The purpose of this section is to examine the effect of mis-specification of the random effect distribution on the model parameter estimation. The results of the simulation are shown in Table (7). The table presents the mean and the average of the estimated standard errors as well as the bias for the regression coefficients for each model. The power and Type I error rate for each model are presented in Table (8).

Table 7
Double Truncated Poisson Regression Model With Miss-Specified Normal Random Effects

Sample Size	Cluster Size	Parameter	Double Truncated Poisson Model			Left Truncated Poisson Model			Regular Poisson Model		
			Estimate (S.E.)		Bias	Estimate (S.E.)		Bias	Estimate (S.E.)		Bias
60	4	β_1	0.4065	0.1364	0.0065	0.3875	0.1312	-0.0125	0.1821	0.0860	-0.2179
		β_2	0.0004	0.1278	0.0004	-0.0002	0.1235	-0.0002	-0.0009	0.0855	-0.0009
		σ	0.8493		-0.0507	0.8552		-0.0448	0.1883		-0.7117
75	5	β_1	0.4082	0.1206	0.0082	0.3897	0.1167	-0.0103	0.1823	0.0766	-0.2177
		β_2	-0.0001	0.1131	-0.0001	0.0009	0.1096	0.00091	-0.0003	0.0761	-0.0003
		σ	0.7986		-0.1014	0.8099		-0.0901	0.1266		-0.7734
90	6	β_1	0.4037	0.1095	0.0037	0.3868	0.1058	-0.0132	0.1826	0.0698	-0.2174
		β_2	-0.0006	0.1025	-0.0006	-0.0003	0.0993	-0.0003	0.0005	0.0693	0.0005
		σ	0.7579		-0.1421	0.7730		-0.1270	0.0880		-0.8120
105	7	β_1	0.4034	0.1009	0.0034	0.3879	0.0976	-0.0121	0.1823	0.0644	-0.2177
		β_2	0.0005	0.0946	0.0005	0.0011	0.0976	0.0011	-0.0005	0.0639	-0.0005
		σ	0.7159		-0.1841	0.7423		-0.1577	0.0865		-0.8135

From Table (7), in the case of the DTMP regression model, the parameter estimate $\hat{\beta}_1$ appears to be asymptotically unbiased and the bias is getting close to zero as cluster size increases. For the case of the LTMP regression model, the parameter estimate $\hat{\beta}_1$ is close to the true value but the bias is larger than the bias in case of the mixed double truncated Poisson model. The standard error of $\hat{\beta}_1$ decreases with increasing cluster size for both double and left truncated mixed Poisson model. The regular mixed Poisson model is the worse in recovering the true parameter value β_1 compared to mixed double truncated Poisson model and mixed left truncated Poisson model. The standard errors of the estimated parameters $\hat{\beta}_1$ for the mixed double truncated Poisson model and the mixed left truncated Poisson model are generally larger than the estimates obtained in the

regular mixed Poisson model. The parameter estimate $\hat{\beta}_2$ is very close to zero for the three models as expected since the generated response variable depends only on x_1 but not x_2 .

The dispersion parameter estimate $\hat{\sigma}$ is biased for the mixed double truncated Poisson model and the mixed left truncated Poisson model. The bias enlarges as cluster size increases. For the regular mixed Poisson model the estimate $\hat{\sigma}$ is highly biased.

The effects of mis-specification of the Normal random effects mainly occur on the dispersion parameter estimate for the three models. Note that the dispersion parameter estimate $\hat{\sigma}$ is biased and the bias negatively increases as the cluster size increases. For the parameter estimate $\hat{\beta}_1$, there is no major change between the correct models and the mis-specified models.

Table 8
Power and Type I Error Rate for Models with Misspecified Normal Random Effects

Sample Size	Cluster Size		Double Truncated Poisson Model	Left Truncated Poisson Model	Regular Poisson Model
60	4	Power	0.8686	0.8640	0.5969
		Type I Rate	0.0482	0.0443	0.0050
75	5	Power	0.9313	0.9313	0.7307
		Type I Rate	0.0482	0.0463	0.0045
90	6	Power	0.9649	0.9643	0.8315
		Type I Rate	0.0497	0.0472	0.0031
105	7	Power	0.9851	0.9840	0.8925
		Type I Rate	0.0510	0.0462	0.0044

From Table (8), the power of the DTMP regression model with mis-specified Normal random effects and LTMP regression model with mis-specified Normal random effects are very similar and increasing by increasing the cluster size. The power of the regular mixed

Poisson model with mis-specified Normal random effects is very low compared to the other models as shown in Figure (7). The Type I error rates for double and left truncated mixed Poisson model are close to nominal value of 0.05 for all different cluster sizes as expected. Meanwhile, the Type I error rate for the regular mixed Poisson model is very conservative around zero as shown in Figure (8).

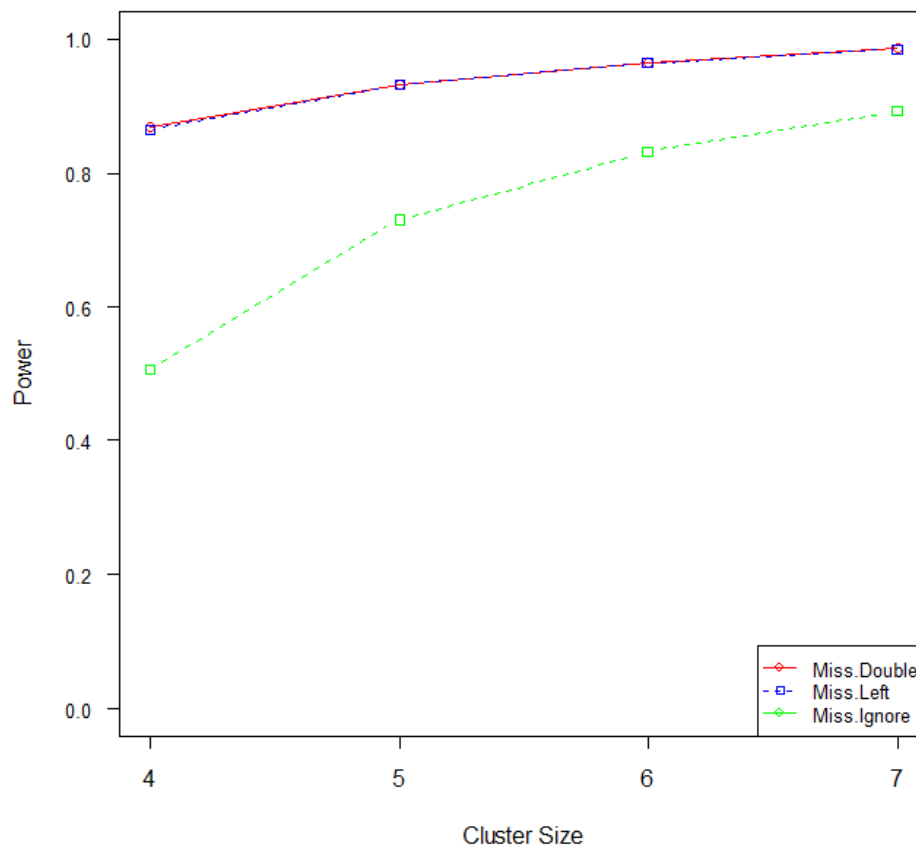


Figure 7 Power of Double Truncated Poisson Model with Miss-specified Normal Random Effects

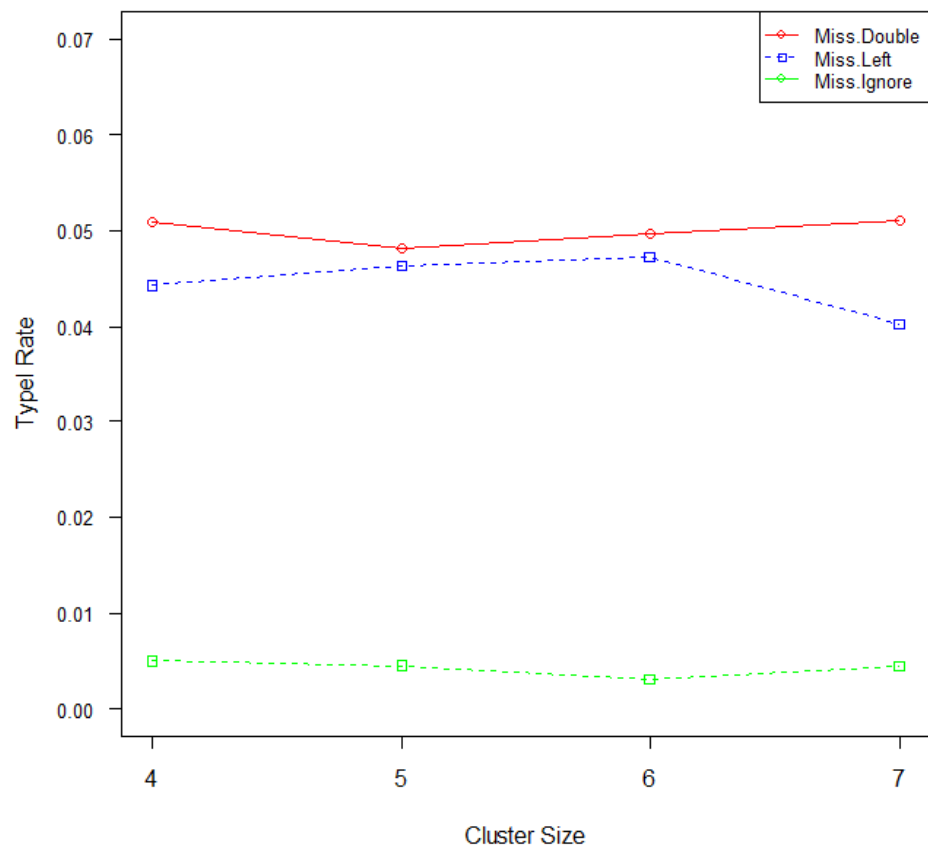


Figure 8 Type I rate of Double Truncated Poisson Model with Miss-specified Normal Random Effects

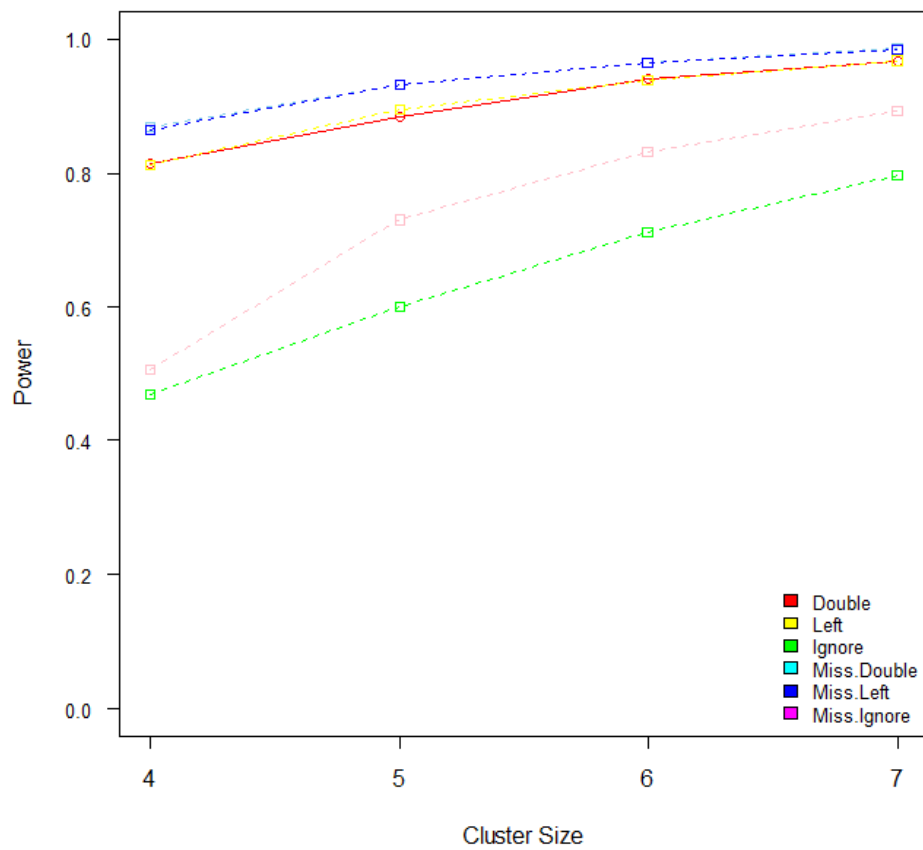


Figure 9 Power of Double Truncated Poisson Model with Mis-specified Normal Random Effects

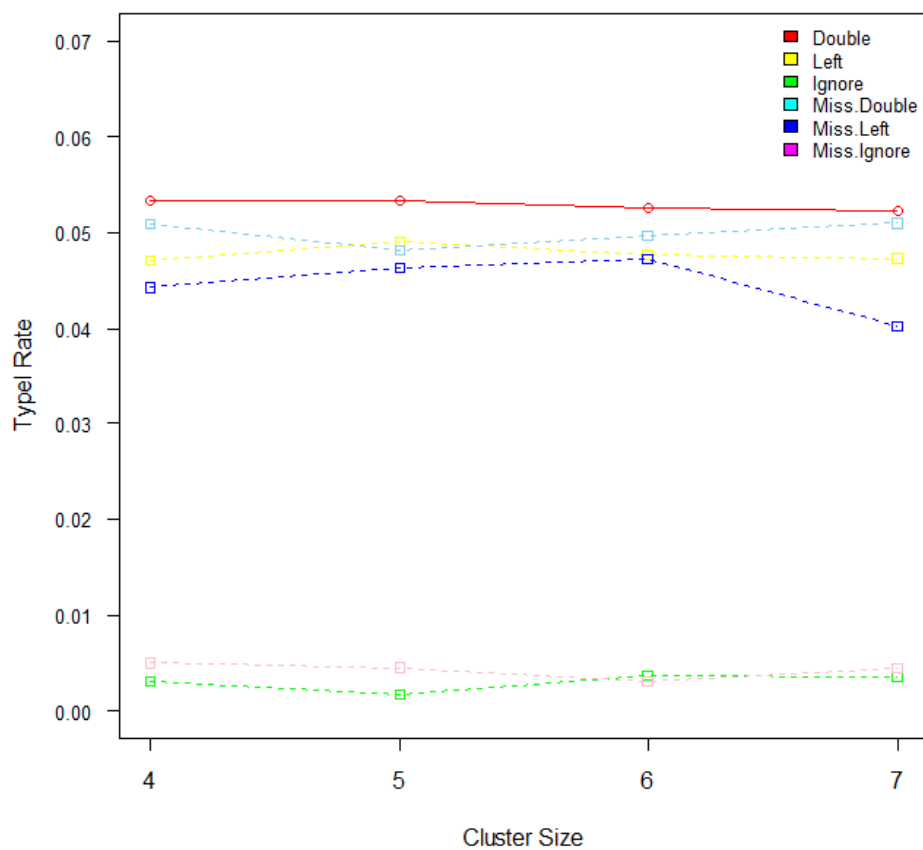


Figure 10 Type I rate of Double Truncated Poisson Model with Mis-specified Normal Random Effects

It can be seen from figure (9) that the mis-specification of Normal random effects affects the power of the models. In the case of the random effects distribution mis-specified as a Gamma distribution instead of a Normal distribution, the power of the double and left truncated mixed Poisson models and also the regular mixed Poisson model are larger than the power of the models when we have correct Normal distribution for the random effects as shown in Figure (9). For Type I error rates, there is no major change between the models with correct Normal random effects and the models with mis-specified Normal random effects. The Type I error rates for the double and left truncated mixed Poisson models with mis-specified Normal random effects are still around 0.05, while the Type I error rates for the regular mixed Poisson model with mis-specified Normal random effects are still around 0 for all different cluster sizes as shown in Figure (10).

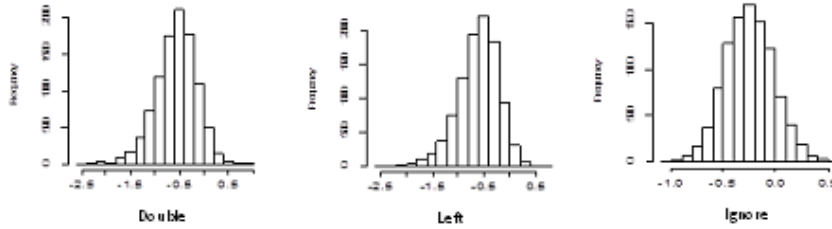


Figure 11 Histogram of the estimated Mis-specified Random effect

Figure (11) shows a sample of the histograms of the estimated random effect values for the three models. The histograms show that even though the random effects were generated from a Gamma distribution, fitting the mixed models with Normal random effects forces the estimated random effect values to be Normally distributed.

Simulation Study for the Double Truncated Mixed Poisson Regression Data with Mis-specified Gamma Random Effects

In this section, the same steps 1-6 from the second section are repeated. In this case for step 2 instead of generating the random effects from a Gamma distribution, the random effects are generated from $\text{Normal}(\text{mean}=0, \text{sd}=0.2)$.

The simulation consists of the following steps:

Step 1: Generate two covariate vectors $(\mathbf{x}_1, \mathbf{x}_2)$ from Uniform (0,3) with length N_i each.

Step 2: Fifteen random effects u_i representing 15 clusters are generated each from Normal (0,0.2).

Step 3: Define $\lambda_{ij} = \exp(\beta_0 + \beta_1 x_{1ij} + u_i)$ and then generate the response variable y_i from the double truncated Poisson distribution with mean λ_{ij} using the function `trun.tr` under the `gamless` library in R package with $\beta_0 = 0.1$ and $\beta_1 = 0.4$. The values of y_i are truncated by 2 on the left and by 8 on the right.

Step 4: Using the h-likelihood estimation method, fit the double truncated Poisson regression model with Gamma random effects.

Step 5: Get the parameter estimates of the model $\hat{\beta}$ and $\hat{\mathbf{u}}$. Based on these estimates and initial values of $\alpha_1 = 3$ and $\alpha_2 = 0.02$, the adjusted profile h-likelihood method was used to estimate the dispersion parameters α_1 and α_2 .

The purpose of this section is to examine the effect of mis-specification of the random effect distribution on the model parameter estimation. The results of the simulation are shown in Table (9). The table presents the mean and the average of the estimated standard errors as well as the bias for the regression coefficients for each model. The power and Type I rate for each model are presented in Table (10).

From Table (9), in the case of the DTMP regression model and the LTMP regression model, the parameter estimate $\hat{\beta}_1$ is within one standard error of the true value and the bias is getting close to zero as cluster size is increased. The standard error of $\hat{\beta}_1$ decreases with increasing cluster size. The regular mixed Poisson model is the worse in recovering the true parameter value β_1 compared to the DTMP regression model and the LTMP regression model. The standard errors of the estimated parameter $\hat{\beta}_1$ for the DTMP regression model and the LTMP regression model are generally larger than the estimates obtained in the regular mixed Poisson model. The parameter estimate $\hat{\beta}_2$ is very close to zero for the three models as expected since the generated response variable depends only on x_1 but not x_2 .

Table 9
Double Truncated Poisson Regression Model With Mis-Specified Gamma Random Effects

Sample Size	Cluster Size	Parameter	Double Truncated Poisson Model			Left Truncated Poisson Model			Regular Poisson Model		
			Estimate (S.E.)		Bias	Estimate (S.E.)		Bias	Estimate (S.E.)		Bias
60	4	β_1	0.4063	0.1456	0.0063	0.3905	0.1415	-0.0095	0.1666	0.0880	-0.2334
		β_2	0.0017	0.1369	0.0017	0.0015	0.1334	0.0015	0.0004	0.0875	0.0004
		α_1	3.1836		0.1836	2.7650		-0.2350	3.1758		0.1758
		α_2	0.0231		0.0031	0.0107		-0.0093	0.0234		0.0034
75	5	β_1	0.4090	0.1288	0.0090				0.1665	0.0784	-0.2335
		β_2	0.0048	0.1208	0.0048				9.8e-05	0.0778	
		α_1	2.9040		-0.0940				2.8886		-0.1114
		α_2	0.0215		0.0015				0.0217		0.0017
90	6	β_1	0.4050	0.1163	0.0050						
		β_2	-0.0003	0.1097	-0.0003						
		α_1	2.6578		-0.3422						
		α_2	0.0201		0.0010						

The dispersion parameter estimate $\hat{\alpha}_1$ is biased for the DTMP regression model for cluster size 4, 5 or 6. $\hat{\alpha}_2$ is asymptotically unbiased and the bias is close to zero for cluster size 4, 5 or 6. The model failed to converge for larger cluster size. For the LTMP regression model, the dispersion parameters $\hat{\alpha}_1 = 2.7650$ with bias=-0.2350 and $\hat{\alpha}_2 = 0.0231$ with bias=0.0069 for cluster size of 4. For larger cluster size 5, 6 or 7, the LTMP model failed to converge. The regular mixed Poisson model is similar to the double truncated mixed Poisson model in terms of recovering the true dispersion parameters for cluster size of 4 or 5 and the model failed to converge for larger cluster size.

The effect of mis-specification of the Gamma random effects mainly occur when the cluster size is increased. All three models fail to converge when the cluster size is 7 or

larger. Also, we can note that the dispersion parameter $\hat{\alpha}_1$ is affected by mis-specification of the random effect distribution. For the parameter estimate $\hat{\beta}_1$, there is no major change between the correct models and the mis-specified models.

Table 10
Power and Type I Error Rate for Models with Misspecified Gamma Random Effects

Sample Size	Cluster Size		Double Truncated Poisson Model	Left Truncated Poisson Model	Regular Poisson Model
60	4	Power	0.8118	0.8071	0.4686
		Type I Rate	0.0548	0.0525	0.0027
75	5	Power	0.8952		0.5991
		Type I Rate	0.0512		
90	6	Power	0.9510		
		Type I Rate	0.0531		

From Table (10), the power of the DTMP regression model with mis-specified Gamma random effects and the LTMP regression model with mis-specified Gamma random effects are very similar and increasing by increasing the cluster size. The power of the regular mixed Poisson regression model with mis-specified Gamma random effects is very low compared to the other models as shown in Figure (12). The Type I error rates for the double and left truncated mixed Poisson model are close to the nominal value of 0.05 for all different cluster sizes as expected. Meanwhile, the Type I error rate for the regular mixed Poisson model is very conservative around zero as shown in Figure (13).

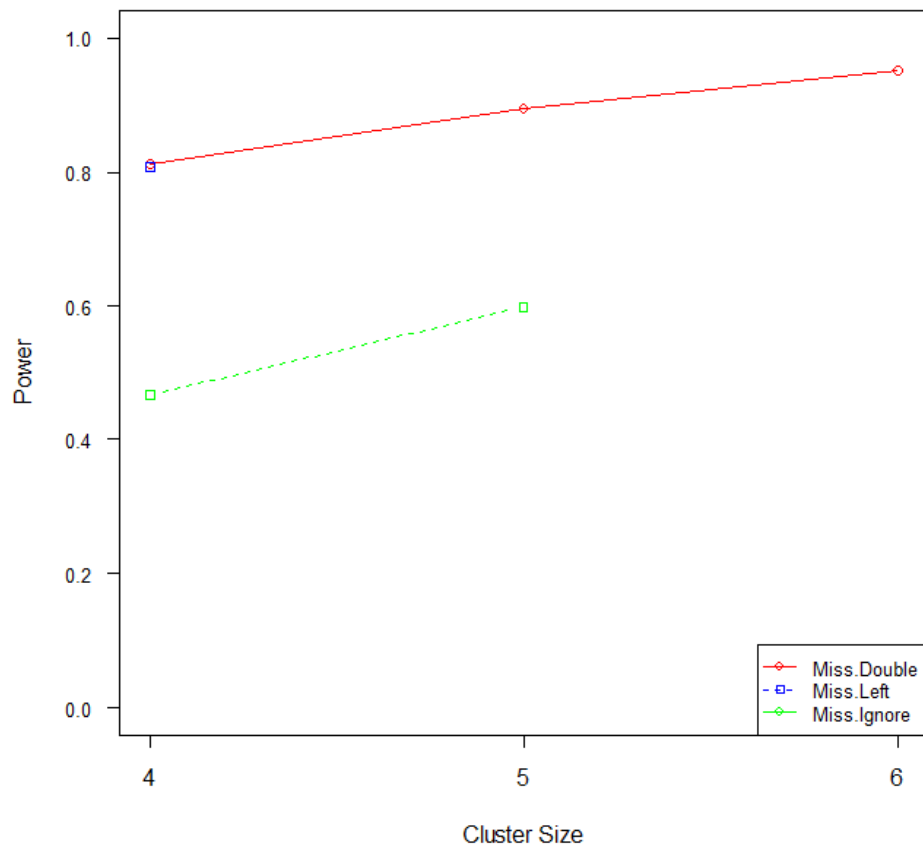


Figure 12 Power of Double Truncated Poisson Model with Mis-specified Gamma Random Effects

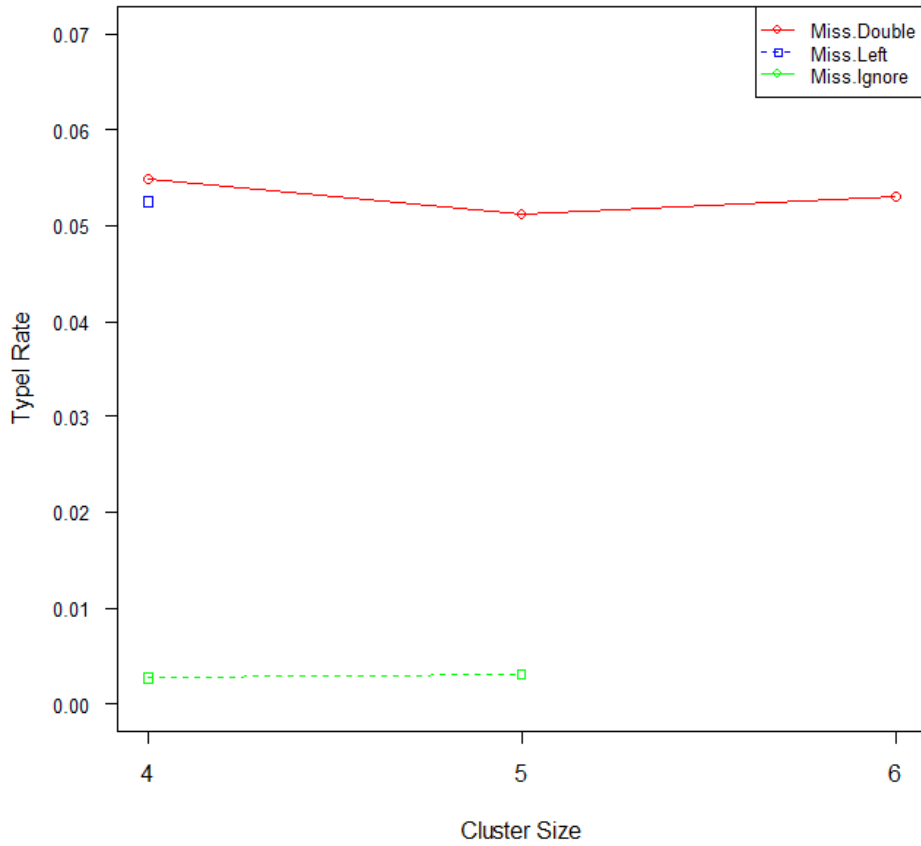


Figure 13 Type I rate of Double Truncated Poisson Model with Mis-specified Gamma Random Effects

It can be seen from figure (14) that the mis-specification of Gamma random effects affects the power of the models. In the case of the random effects distribution mis-specified as a Normal distribution instead of a Gamma distribution, the power of the double and left truncated mixed Poisson model and also the regular mixed Poisson models is lower than the power of the models when we have correct Gamma distribution for the random effects as shown in Figure (14). For Type I error rates, there is no major change between the models with correct Gamma random effects and the models with mis-specified Gamma random effects. The Type I error rates for the double and left truncated mixed Poisson model with mis-specified Gamma random effects are still around 0.05, while the Type I

error rates for regular mixed Poisson model with mis-specified Gamma random effects are still around 0 for all different cluster sizes as shown in Figure (15).

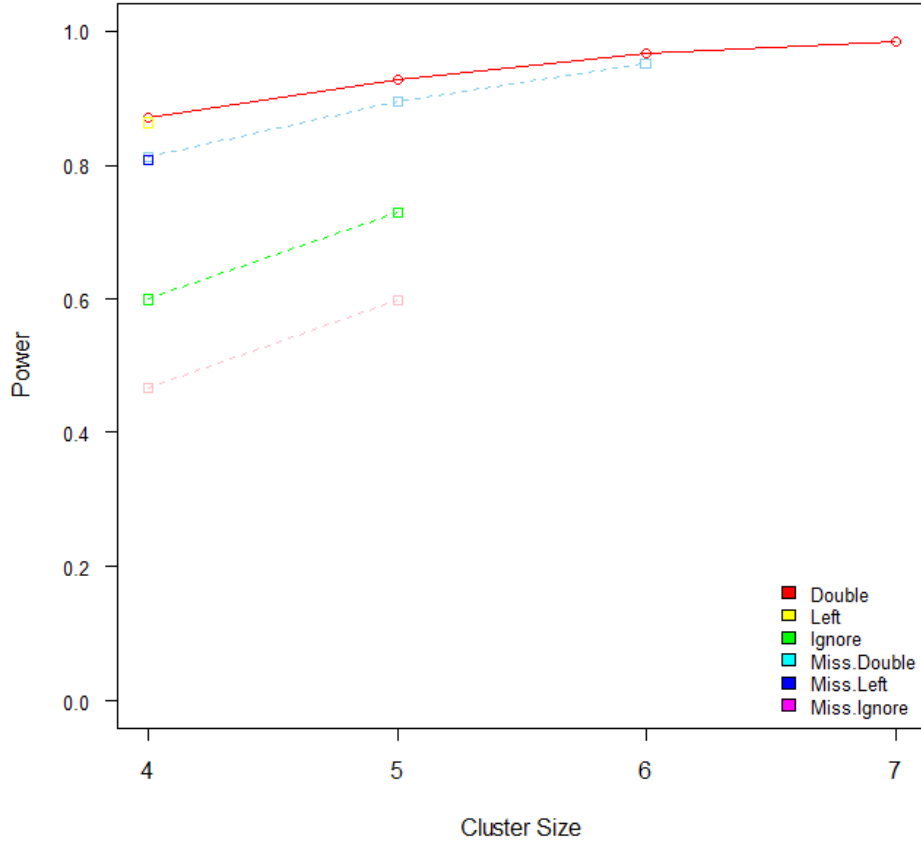


Figure 14 Power of Double Truncated Poisson Model with Mis-specified Gamma Random Effects

Figure (16) shows a sample of the histograms of the estimated random effect values for the three models. The histograms show that even though the random effects were generated from Normal distribution, fitting the mixed models with Gamma random effects force the estimated random effect values to be Gamma distributed.

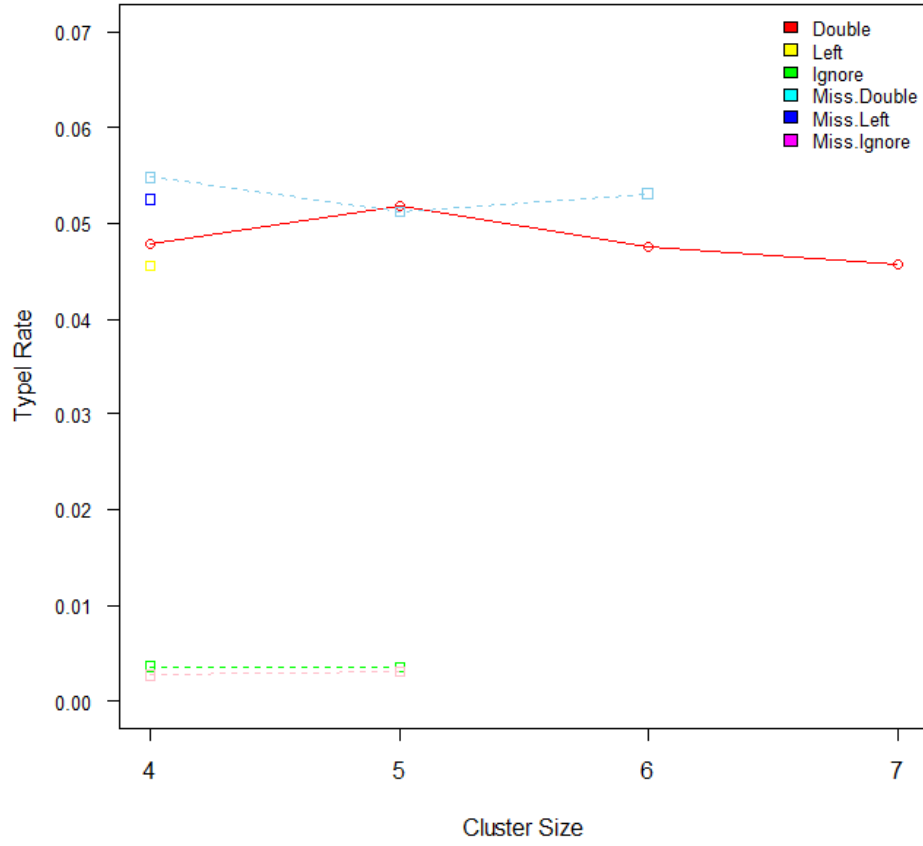


Figure 15 Type I rate of Double Truncated Poisson Model with Mis-specified Gamma Random Effects

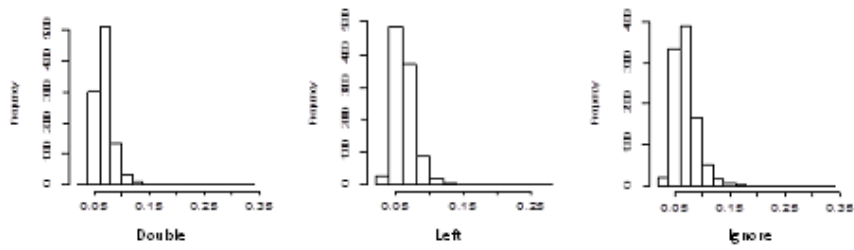


Figure 16 Histogram of the estimated Misspecified Random effect

Real Data Example

To illustrate the use of the mixed double truncated Poisson regression model on actual data, we utilized data presented from a transitional housing facility (THF) in the greater Rocky Mountain Region. The THF is a temporary facility helping family units find stable housing and employment while living within the shelter. The THF aims to help these first time and episodic (experiencing episodes of homelessness) homeless families regain stable housing and employment. Here we present an observation dataset of families living in THF from 2006 to 2010 with a total of 926 observations.

For the purpose of this study, we selected three independent variables in addition to the response variable described as:

LOS_NIGHT(Response Variable): The number of nights a family occupied a room within the shelter for each month (min=1, max=31).

CASE_HOURS: The number of hours a client spent with a case worker within the THF.

Employed: The number of family members currently employed.

Single_Parents: Dichotomous variable taking a value of 1 for a single parent family (1=Single Parent, 0=not single parent).

The double truncated mixed Poisson regression model is appropriate to model the number of nights a family occupied a room within the shelter for each month since the response variable values are between 1 and 31. Families are assumed to be clustered within months, so we have 12 clusters represent 12 months.

To evaluate the performance of the double truncated mixed Poisson regression model, we fit the double, left and regular mixed Poisson regression models assuming Normal and Gamma distributions for the random effects.

Double Truncated Poisson Model with Normal Random Effects

We have the model

$$y_{ij}|u_i \sim \text{Poisson}(\lambda_{ij}),$$

$$\lambda_{ij} = \exp(\beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \beta_3 x_{3ij}),$$

$$u_i \sim N(0, \sigma_u^2),$$

where

y_{ij} : represents the response variable LOS_NIGHT.

x_{1ij} : represents the CASE_HOURS variable.

x_{2ij} : represents the EMPLOYED variable.

x_{3ij} : represents the SINGLE_PARENTS variable.

u_i : represents clusters (months).

Table (11) lists the parameter estimates, standard errors and p-value for the regression coefficients as well as the log-likelihood values from the double truncated mixed Poisson model, the left truncated mixed Poisson model and the regular mixed Poisson model assuming a Normal distribution for the random effects. The question of interest is to compare goodness of fit of the three models and select the model that provides a best fit of the data. One way to select the best model is to use the likelihood ratio test. The log-likelihood values are -3075.034 for the double truncated mixed Poisson model, -3361.742 for the left truncated mixed Poisson model, and -0.3361.852 for the regular mixed Poisson model. By the likelihood ratio test, the double truncated mixed Poisson model improves model fit compared to the left truncated mixed Poisson model and the regular mixed Poisson model where $\chi^2 = 573.416$. Therefore, the double truncated Poisson model with Normal random effects provides the best fit of the data. Note that accounting for double truncation in the model changes the estimates of the regression coefficients and their standard errors compared to those obtained by the other two models.

A double truncation interval of $(2,8)$ was the only interval considered in this dissertation. Different truncation values will be of interest in future research.

In this dissertation, a single random effect has been considered in the double truncated mixed Poisson models. A further extension of these models include multiple random effects in order to accommodate two stages of clustering will be of interest in future research.

Finally, The Negative Binomial distribution has also been used to model count data. A double truncated mixed Negative Binomial model would be another way to study double truncated clustered count data. This model will also be of interest in future research.

- Greenwood, M. and Yule, G. U. (1920). *An Inquiry into the Nature of Frequency Distributions Representative of Multiple Happenings with Particular Reference to the Occurrence of Multiple Attacks of Disease or of Repeated Accidents* . Journal of the Royal Statistical Society, 83, 225-279.
- Goldstein, H. (1991). *Nonlinear Multilevel Models with an Application to Discrete Response Data*. Biometrika, 78, 45-51.
- Grogger, J. T. and Carson, R. T. (1991). *Models For Truncated Counts*. Journal of Applied Econometrics, 6, 225-238.
- Gupta, P. and Gupta, R. (2004). *Modeling Count Data by Random Effect Poisson Model*. Indian Journal of Statistics, 66, 548-565.
- Gurmu, S. (1991). *Tests for Detecting Overdispersion in the Positive Poisson Regression Model*. Journal of Business and Economic Statistics, 9, 215-222.
- Gurmu, S. and Trivedi, P. K. (1992). *Overdispersion Tests for Truncated Poisson Regression Models*. Journal of Econometrics, 9, 94-113.
- Harville, D. A. (1977). *Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems*. Journal of the American Statistical Association, 72, 320-338.
- Hedeker, D. and Gibbons, R. D. (2006). *Longitudinal Data Analysis*. John Wiley and Sons, Inc., Hoboken, New Jersey.

APPENDIX A

HISTOGRAMS OF THE ESTIMATED RANDOM EFFECTS

Appendix A presents the histograms of the estimated Normal and Gamma random effects for the double, left, and regular mixed Poisson models.


```
hist(uhat6)
hist(uhat7)
hist(uhat8)
hist(uhat9)
x11()
par(mfrow=c(2,3))
hist(uhat10)
hist(uhat11)
hist(uhat12)
hist(uhat13)
hist(uhat14)
hist(uhat15)
}
```



```

return(list(alpha1hat,alpha2hat))
}
#####
# Start the simulation
#####
Left.Simulation=function(m,r){
a=0
b=0
b1=0
b2=0
SE1=0
SE2=0
uhat1=0
uhat2=0
uhat3=0
uhat4=0
uhat5=0
uhat6=0
uhat7=0
uhat8=0
uhat9=0
uhat10=0
uhat11=0
uhat12=0
uhat13=0
uhat14=0
uhat15=0
Dis.Left1=0
Dis.Left2=0
for(i in 1:r){
n.clus <- m
n.per.clus <- 5
n <- n.clus*n.per.clus
x1=runif(n,0,3)
x2=runif(n,0,3)
v=rep(1,n)
X <- matrix(c(v,x1,x2),n,3)
Z <- diag(n.clus)%x%rep(1, n.per.clus)
u=rgamma(n.clus,shape=1,scale=0.12)
#u=rnorm(n.clus,0,0.2)
lambda <- exp(0.1+0.4*x1 + Z%*%u)
library(gamlss.tr)
test1<-trun.r(par=c(1,8), family="P0", type="both")
y=test1(n,lambda)
Par.Est=fix.Left.Trunc(y,X,Z,m=m,alpha1=2,alpha2=0.1,c1=1)
betahat=Par.Est[[1]][,1]

```



```

cat("Left Truncated u.hat11",mean(uhat11),"\n")
cat("Left Truncated u.hat12",mean(uhat12),"\n")
cat("Left Truncated u.hat13",mean(uhat13),"\n")
cat("Left Truncated u.hat14",mean(uhat14),"\n")
cat("Left Truncated u.hat15",mean(uhat15),"\n")
cat("alfahat1",mean(Dis.Left1),"\n")
cat("alfahat2",mean(Dis.Left2),"\n")
x11()
par(mfrow=c(3,3))
hist(uhat1)
hist(uhat2)
hist(uhat3)
hist(uhat4)
hist(uhat5)
hist(uhat6)
hist(uhat7)
hist(uhat8)
hist(uhat9)
x11()
par(mfrow=c(2,3))
hist(uhat10)
hist(uhat11)
hist(uhat12)
hist(uhat13)
hist(uhat14)
hist(uhat15)
}
#####
# Funtion to estimate the parameters beta and u
# fo regular Poisson model with gamma random effect
#####
fix.Ordinal.Trunc=function(y,X,Z,m,alpha1,alpha2)
{
  beta=c(0.1,0.4,0.4)
  u=rep(0.06,m)
  h.lik=function(beta,u)
  {
    log.lik1=function(beta)
    {
      lambda=exp(X%*%beta+Z%*%u)
      eta=log(lambda)
      sum(y*eta-lambda-log(factorial(y)))
    }
    log.lik2=function(u)
    {
      term1=-m*(alpha1*log(alpha2)+log(gamma(alpha1)))

```



```

cat("Ordinal Truncated u.hat14",mean(uhat14),"\n")
cat("Ordinal Truncated u.hat15",mean(uhat15),"\n")
cat("alfahat1",mean(Dis.Ordinal1),"\n")
cat("alfahat2",mean(Dis.Ordinal2),"\n")
x11()
par(mfrow=c(3,3))
hist(uhat1)
hist(uhat2)
hist(uhat3)
hist(uhat4)
hist(uhat5)
hist(uhat6)
hist(uhat7)
hist(uhat8)
hist(uhat9)
x11()
par(mfrow=c(2,3))
hist(uhat10)
hist(uhat11)
hist(uhat12)
hist(uhat13)
hist(uhat14)
hist(uhat15)

}

```

APPENDIX C

TRANSITIONAL HOUSING FACILITY DATA SET

20 8 1 1 10
20 8 1 1 10
6 4 0 1 10
5 4 2 0 10
1 1 2 0 10
22 6 1 1 10
8 4 0 1 10
31 8 2 0 10
31 10 2 0 10
2 2 2 0 10
2 4 2 0 10
31 8 0 1 10
31 10 1 1 10
15 6 1 1 11
11 4 2 0 11
7 3 1 1 11
5 3 1 1 11
30 10 1 1 11
18 6 1 1 11
30 10 2 0 11
30 8 3 1 11
30 10 1 0 11
12 7 0 1 11
30 10 1 1 11
22 8 1 1 11
30 8 1 1 11
26 8 1 1 11
19 6 1 1 11
18 6 2 0 11
20 8 1 1 11
20 8 2 0 11
30 10 2 0 11
30 10 2 0 11
25 8 1 1 11
10 4 1 1 11
5 4 1 1 11
25 6 1 1 11
30 10 2 0 11
30 10 2 0 11
26 8 1 1 11
8 4 2 0 11
3 3 0 1 11
24 8 1 1 11
24 8 2 0 11
20 6 0 1 11
14 6 0 1 11

31 10 1 1 12
22 8 0 1 12
31 10 1 1 12
1 2 2 0 12
31 10 2 0 12
10 2 2 0 12
20 6 1 1 12