A robust test of homogeneity in zero-inflated models for count data

by

Nadeesha R. Mawella

B.S., University of Peradeniya, 2010

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Statistics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2018

Abstract

Evaluating heterogeneity in the class of zero-inflated models has attracted considerable attention in the literature, where the heterogeneity refers to the instances of zero counts generated from two different sources. The mixture probability or the so-called mixing weight in the zero-inflated model is used to measure the extent of such heterogeneity in the population. Typically, the homogeneity tests are employed to examine the mixing weight at zero. Various testing procedures for homogeneity in zero-inflated models, such as score test and Wald test, have been well discussed and established in the literature. However, it is well known that these classical tests require the correct model specification in order to provide valid statistical inferences. In practice, the testing procedure could be performed under model misspecification, which could result in biased and invalid inferences. There are two common misspecifications in zero-inflated models, which are the incorrect specification of the baseline distribution and the misspecified mean function of the baseline distribution. As an empirical evidence, intensive simulation studies revealed that the empirical sizes of the homogeneity tests for zero-inflated models might behave extremely liberal and unstable under these misspecifications for both cross-sectional and correlated count data.

We propose a robust score statistic to evaluate heterogeneity in cross-sectional zero-inflated data. Technically, the test is developed based on the Poisson-Gamma mixture model which provides a more general framework to incorporate various baseline distributions without specifying their associated mean function. The testing procedure is further extended to correlated count data. We develop a robust Wald test statistic for correlated count data with the use of working independence model assumption coupled with a sandwich estimator to adjust for any misspecification of the covariance structure in the data. The empirical performances of the proposed robust score test and Wald test are evaluated in simulation studies. It is worth to mention that the proposed Wald test can be implemented easily with

minimal programming efforts in a routine statistical software such as SAS. Dental caries data from the Detroit Dental Health Project (DDHP) and Girl Scout data from Scouting Nutrition and Activity Program (SNAP) are used to illustrate the proposed methodologies.

A robust test of homogeneity in zero-inflated models for count data

by

Nadeesha R. Mawella

B.S., University of Peradeniya, 2010

A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Statistics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2018

Approved by:

Major Professor
Dr. Wei-Wen Hsu

Copyright

Abstract

Evaluating heterogeneity in the class of zero-inflated models has attracted considerable attention in the literature, where the heterogeneity refers to the instances of zero counts generated from two different sources. The mixture probability or the so-called mixing weight in the zero-inflated model is used to measure the extent of such heterogeneity in the population. Typically, the homogeneity tests are employed to examine the mixing weight at zero. Various testing procedures for homogeneity in zero-inflated models, such as score test and Wald test, have been well discussed and established in the literature. However, it is well known that these classical tests require the correct model specification in order to provide valid statistical inferences. In practice, the testing procedure could be performed under model misspecification, which could result in biased and invalid inferences. There are two common misspecifications in zero-inflated models, which are the incorrect specification of the baseline distribution and the misspecified mean function of the baseline distribution. As an empirical evidence, intensive simulation studies revealed that the empirical sizes of the homogeneity tests for zero-inflated models might behave extremely liberal and unstable under these misspecifications for both cross-sectional and correlated count data.

We propose a robust score statistic to evaluate heterogeneity in cross-sectional zero-inflated data. Technically, the test is developed based on the Poisson-Gamma mixture model which provides a more general framework to incorporate various baseline distributions without specifying their associated mean function. The testing procedure is further extended to correlated count data. We develop a robust Wald test statistic for correlated count data with the use of working independence model assumption coupled with a sandwich estimator to adjust for any misspecification of the covariance structure in the data. The empirical performances of the proposed robust score test and Wald test are evaluated in simulation studies. It is worth to mention that the proposed Wald test can be implemented easily with

minimal programming efforts in a routine statistical software such as SAS. Dental caries data from the Detroit Dental Health Project (DDHP) and Girl Scout data from Scouting Nutrition and Activity Program (SNAP) are used to illustrate the proposed methodologies.

Table of Contents

List of Figures x

List of Tables xi

Acknowledgements xiii

1 Introduction 1

2 Impact of misspecification on the homogeneity tests for zero-inflated models . . . 5

2.1 Introduction 5

2.2 Zero-Inflated model 6

2.3 Score test of homogeneity 7

2.4 Misspecifications 9

2.5 Numerical study 10

2.5.1 Misspecification of mean function of the null model 10

2.5.2 Misspecification of baseline distribution 12

2.6 Discussion 14

3 A robust score test of homogeneity for zero-inflated count data 15

3.1 Introduction 15

3.2 Robust homogeneity test for zero-inflated models 16

3.2.1 Kassahun's model 16

3.2.2 The proposed robust test 17

3.3 Simulation Study 20

3.4 Real Data Application 22

3.4.1 Dental Caries Data 22
3.4.2 Girl Scout Data 24
3.5 Discussion 25

4 A robust homogeneity test for correlated count data with excess zeros 28
4.1 Introduction 28
4.2 Zero-inflated model for correlated count data 30
4.3 The proposed robust test for correlated zero-inflated count data 30
4.4 Numerical Studies 32
4.5 Real Data Application 35

5 Discussion 41
5.1 Summary 41

Bibliography 43

A Score function and second derivatives of log likelihood function for Robust score test 47

B Sample R Code: Robust score test 51
B.1 Sample SAS Code: Robust Wald test 52

List of Figures

3.1 Observed vs fitted proportions for Girl Scout data by the NB model. 26

3.2 Observed vs fitted proportions by the NB model for intervention group and control group in Girl Scout data. 27

4.1 Observed vs fitted proportions by waves under the NB model for Detroit dental caries data. 40

List of Tables

2.1 Empirical sizes of the score test statistics at α=0.05 based on 1,000 samples generated from Poisson with true mean $\log(\lambda^*) = 0.75 - 1.45X_1 - 0.8X_2$. . . . 11

2.2 Empirical sizes of the score test statistics with well specified working mean function based on 1,000 samples generated from Poisson regression model with mean λ^* at α=0.05. 12

2.3 Empirical sizes of the score test statistics with well specified working mean function based on 1,000 samples generated from Negative Binomial regression model with mean λ^* at α=0.05. 13

3.1 Empirical sizes and power of the robust score test statistic based on 1,000 samples generated from zero-inflated Negative Binomial regression model with mean λ^*, at the nominal level 0.05. 21

3.2 Empirical sizes and power of the robust score test statistics based on 1,000 samples generated from zero-inflated Poisson regression model with mean λ^*, at the nominal level 0.05. 22

3.3 Comparison of score test statistics, degrees of freedoms and associated p-values of homogeneity tests for dental caries data. 23

3.4 Comparison of score test statistics, degrees of freedoms and associated p-values of homogeneity tests for Girl Scout data. 24

3.5 Fits of different count models for Girl Scout data. 26

4.1 At the 5% significance level, the empirical type I error rates of the classical Wald test under the misspecification of the conditional mean with data generated from a Poisson process with true mean $\log(\lambda^*) = 0.3 + 0.5X_1 + 0.2X_2$. 34

4.2 At the 5% significance level, the empirical type I error rates of the classical Wald test under the misspecification of the baseline distribution. 34
4.3 Empirical size and power of the robust Wald test statistics at the nominal level 0.05, based on 1,000 samples of correlated zero-inflated Negative Binomial data with mean λ^*. 36
4.4 Empirical size and power of the robust Wald test statistics at the nominal level 0.05, based on 1,000 samples of correlated zero-inflated Poisson data with mean λ^*. 37
4.5 Comparison of Wald test statistics and associated p-values of the homogeneity tests for longitudinal dental caries data. 38
4.6 Fitted constant mixing weight models for longitudinal dental caries data. . . 39

Acknowledgments

I would like to express my sincere gratitude and appreciation to my major professor, Dr. Wei-Wen Hsu. His mentorship, encouragement and valuable suggestions greatly helped me to advance my research training by developing professional qualities and improving my scientific writing skills. I would also like to thank my committee members Dr. Weixing Song, Dr. Gyuhyeong Goh, Dr. Christopher Vahl and Dr. Richard Rosenkranz for their willingness to serve on my committee and for their time, interest, and helpful comments on my research work. I would also like to acknowledge Dr. Todd Easton for his willingness to be the chairperson of the examining committee.

I would like to thank Kansas State University and the department of Statistics for giving me an opportunity to achieve my higher education. Many thanks to the faculty and staff of the department of statistics, at some point in my graduate career, each of them has helped me to become successful in the graduate program.

I would like to thank my family and friends for their support and belief in my success. Finally, I would like to thank my husband, Chinthaka. His support and understanding made this work possible; his friendship and love made it fulfilling.

Chapter 1

Introduction

Zero-Inflated models have gained a considerable attention in a variety of fields due to their capability of handling excess zeros in count data. They are useful extensions of the classical generalized linear models and have been widely used in many areas of science, finance, epidemiology, agriculture and biomedical sciences (for example, see Farewell and Sprott, 1988; Ridout et al., 1998; Böhning et al., 1999; Yip and Yau, 2005; Hu et al., 2011).

Zero-Inflated models provide a parametric framework to accommodate heterogeneity in a population, in which the heterogeneity is referred to zeros generated from two different sources. Specifically, zeros are assumed as being generated from a degenerate distribution at zero and a discrete distribution (or so-called baseline distribution). The mixing weight of the zero-inflated model measures the extent of this type of heterogeneity in the population.

In real applications of the zero-inflated models, the question typically interested to address is whether the mixing weight or so-called mixture probability adequately represents the inherent heterogeneity in the population. Intuitively, a zero mixing weight indicates that zeros are only generated from the non-degenerate distribution. Therefore, evaluating the hypothesis that mixing weight equals to zero is equivalent to evaluating the homogeneous model against the heterogeneous model.

In practice, as a goodness-of-fit, homogeneity tests are often used to evaluate the hypothesis of zero mixing weight. One popular homogeneity test in the literature for this class of models is score test. Score test has an advantage that the parameter estimation is required only for the model under the null hypothesis which makes the test easy to perform in some cases. Several studies in the literature have used score tests for testing heterogeneity under this class of models. For example, Van den Broek (1995) proposed a score test to evaluate whether a constant mixing weight equals to zero under Zero-Inflated Poisson models. A similar work under Zero-Inflated Binomial models was proposed by Deng and Paul (2000). As an extension, Jansakul and Hinde (2002) proposed a score test under the framework of Zero-Inflated Poisson model using an identity link function to relate mixing weight to covariates in order to improve the power of the test. However, the identity link function is seldom used in practice and requires an additional constraint on the likelihood function if the true mixing weight is between 0 and 1. Without using the identity link function, Todem et al. (2012) developed a more general score test using a novel transformation and allowing the mixing weight to depend on covariates to improve the power.

It is known that a valid score test requires the assumption of correct specification of the model under the null hypothesis. This assumption might not be satisfied in practice and the violation of this assumption could lead to unreliable statistical inferences. Misspecification in hypothesis testing is a general issue which has been pointed out in several studies under different classes of models. For example, Godfrey (1991) showed the impact of misspecification on the Lagrange multiplier test for the regression models. Bera and Yoon (1993) mentioned that the score test is not robust when the nuisance parameter is locally misspecified. Liang and Self (1996) also indicated that the nuisance parameter may be misspecified in likelihood functions, which affects the validity of the likelihood ratio test. DiRienzo and Lagakos (2001) showed that score test and Wald test under Cox's proportional hazards model will not be asymptotically valid when the model is misspecified. To our knowledge, the impact of misspecification of homogeneity tests under the class of zero-inflated models has not been discussed in the literature.

In this study, we evaluate two common types of misspecification on the homogeneity tests under zero-inflated models: (1) Misspecification of the mean function of the null model and (2) Misspecification of the baseline distribution of the null model. To address these issues of misspecification, we propose a new robust score test of homogeneity for cross-sectional zero-inflated data. The proposed test can be performed without specifying the mean function and hence avoids the impact of any possible misspecification of conditional mean under the null model. Technically, the test assumes a more general model as the baseline model, thus, it is also robust to the misspecification of the baseline distribution.

The proposed testing procedure is further extended for correlated count data with excess zeros as the zero-inflation may present in data simultaneously with correlation due to the hierarchical nature of the study design or the data collection procedure such as repeated measurements on subjects. To account for such data, a robust Wald test statistic is developed. Under the same model formulation proposed in this study, a working independence model assumption accompanying a sandwich estimator is used to accommodate the correlation in data. Despite the alternative approach of directly modeling the correlation structure in data using the random effect terms with a conditional model (Hall, 2000; Min and Agresti, 2005; Lee et al., 2006; Xiang et al., 2006), a working independence model approach is used due to the possible impacts on the mean structure by the former approach when the random effects are integrated out (Hsu et al., 2014). Instead of integrating over the distributions defined by random effects, the proposed approach along with the sandwich estimator of the variance-covariance matrix allows to execute the estimation process in a computationally tractable manner using standard software such as SAS.

In Chapter 2, we evaluate the impact of the two types of misspecification on the performances of homogeneity score test by intensive simulations. A new robust test for testing heterogeneity under zero-inflated models is introduced in Chapter 3. The performance of the proposed test is evaluated by simulation studies and then applied to dental caries data from the Detroit Dental Health Project (DDHP) and Girl Scout data from Scouting Nutrition and Activity Program (SNAP). In Chapter 4, we extend the proposed methodology to accommodate the dependency in count data along with the zero-inflation by introducing a robust Wald test for correlated count data with extra zeros. The empirical performance of the proposed Wald test is assessed though simulation studies and longitudinal dental caries data are used to illustrate the proposed Wald test. In Chapter 5, we discuss the advantages and the limitations of the proposed methodology and the possible extensions of the proposed test for future work.

Chapter 2

Impact of misspecification on the homogeneity tests for zero-inflated models

2.1 Introduction

In this chapter, we evaluate the impact of two common types of misspecification on homogeneity tests under zero-inflated models. The first type of misspecification occurs when the mean function is not well specified in the model under null hypothesis. It is often assumed that the mean of the null model depends on covariates through an appropriate link function, but it could be misspecified. The second type of misspecification occurs when the working baseline distribution is not correctly specified. For instance, if the working baseline distribution is assumed to be Poisson but the true underlying distribution is Negative Binomial distribution. As it is well known that count data often show over dispersion compared to the Poisson distribution, if the over dispersion was not incorporated by the baseline model appropriately, the homogeneity test for zero-inflated data may provide unreliable conclusions about the population. We evaluate the performances of score test under these two types of misspecification through intensive simulations.

This chapter is organized as follows. In Section 2.2, we describe the general formulation of zero-inflated model and in Section 2.3, the score test statistic for testing heterogeneity. Two common types of misspecification are discussed in Section 2.4. The simulations investigating the impact of misspecification are given in Section 2.5. Discussion based on the simulation studies is given in Section 2.6.

2.2 Zero-Inflated model

Suppose we randomly select a sample of n independent subjects with count responses $Y_i, i = 1, ..., n$ from a population represented by a zero-inflated model. The responses Y_i are assumed as generating from a mixture of degenerate distribution at 0 and a discrete distribution with probability mass function $g(\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is a vector of unknown parameters. The corresponding distribution of Y_i is a two component mixture with the probability mass function given below,

$$P(Y_i = y_i) = \begin{cases} \omega_i + (1-\omega_i)\, g_i(0;\boldsymbol{\theta}) & \text{if } y_i = 0 \\ (1-\omega_i)\, g_i(y_i;\boldsymbol{\theta}) & \text{if } y_i = 1, 2, 3, ..., \end{cases}$$

where ω_i is the mixing weight or so-called mixture probability.

In general, under zero-inflated models, mixing weight is considered as bounded between 0 and 1 such that $0 \leq \omega_i \leq 1$, for $i = 1, ..., n$. However, under the marginal representation of the model, mixing weight is constrained as $-g_i(0;\boldsymbol{\theta})/(1 - g_i(0;\boldsymbol{\theta})) \leq \omega_i \leq 1$, for $i = 1, ..., n$, which allows to accommodate both zero-inflation and zero-deflation (see Todem et al., 2012). Particularly, when $\omega_i{>}0$, there exists many zeros than the zeros expected from the discrete distribution $g(\boldsymbol{\theta})$, then the model is called the zero-inflated model. If $\omega_i{<}0$, there exists too few zeros than the expected zeros from the discrete distribution $g(\boldsymbol{\theta})$ which results in zero-deflated model. When $\omega_i = 0$, the model reduces to the homogeneous model which corresponds to discrete distribution $g(\boldsymbol{\theta})$. For different discrete distributions $g(\boldsymbol{\theta})$, there are several popular zero-inflated models such as Zero-Inflated Poisson model, Zero-Inflated

Negative Binomial model and Zero-Inflated Binomial model. Lambert (1992) proposed Zero-Inflated Poisson model where $g(\boldsymbol{\theta})$ is a Poisson distribution,

$$g_i(y_i;\boldsymbol{\theta}) = g_i(y_i;\lambda_i) = \frac{e^{-\lambda_i}\lambda_i^{y_i}}{y_i!}, \ \ y_i = 0, 1, 2, 3, ...,$$

and λ_i is the mean such that $\lambda_i = e^{X_i^T\boldsymbol{\beta}}$ for unknown parameter vector $\boldsymbol{\beta}$ and a matrix of covariates X. In this model, $\boldsymbol{\theta} = \boldsymbol{\beta}$.

For Zero-Inflated Negative Binomial model, the baseline distribution $g(\boldsymbol{\theta})$ is specified as,

$$g_i(y_i;\boldsymbol{\theta}) = g_i(y_i;\lambda_i,\alpha) = \frac{\Gamma(y_i+\alpha^{-1})}{y_i!\,\Gamma(\alpha^{-1})}\frac{\alpha^{y_i}\lambda_i^{y_i}}{(1+\alpha\lambda_i)^{y_i+\alpha^{-1}}}, \ \ y_i = 0, 1, 2, 3, ...,$$

where $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta})$ and α is the dispersion parameter, λ_i is the mean such that $\lambda_i = e^{X_i^T\boldsymbol{\beta}}$ for unknown parameter vector $\boldsymbol{\beta}$ and a matrix of covariates X. If $g(\boldsymbol{\theta})$ is a Binomial distribution, the resulting model is a Zero-Inflated Binomial model (See, for example, Farewell and Sprott, 1988).

2.3 Score test of homogeneity

We are often interested to evaluate whether the mixing weight is zero under the class of zero-inflated models. In the literature, many testing procedures assume constant mixing weight such that $\omega_i = \omega$ for all i. For example, Van den Broek (1995) proposed a score test under Zero-Inflated Poisson models assuming a constant mixing weight. Under Zero-Inflated Binomial model, a score test assuming a constant mixing weight was discussed in Deng and Paul (2000). For ω=0, heterogenous model reduces to the homogeneous model that corresponds to the baseline distribution. Thus, for testing whether the homogenous model is adequate, we are typically interested in evaluating the hypothesis $\omega = 0$.

For testing the hypothesis $\omega = 0$, score test is often used under this class of models as it requires parameter estimation only under the null hypothesis which is an advantage in practice. Under this class of models, assuming a sample of independent observations

$y_1, ..., y_n$, the associated log-likelihood function $l(\boldsymbol{y}, \boldsymbol{\theta}, \omega)$ with $\omega_i = \omega$ for all i is

$$l(\boldsymbol{y}, \boldsymbol{\theta}, \omega) = \sum_{i=1}^{n} \left[I_{(y_i=0)} \log\{\omega + (1-\omega) g_i(0;\theta_i)\} + I_{(y_i>0)} \log\{(1-\omega) g_i(y_i;\theta_i)\} \right]$$

To conduct the score test, we can follow the standard procedures in asymptotic theory. The score vector can be derived based on the log-likelihood function $l(\boldsymbol{y}, \boldsymbol{\theta}, \omega)$.

$$S(\boldsymbol{\theta}, \omega) = \begin{bmatrix} S_{\boldsymbol{\theta}}(\boldsymbol{\theta}, \omega) \\ S_{\omega}(\boldsymbol{\theta}, \omega) \end{bmatrix} = \begin{bmatrix} \dfrac{\partial l(\boldsymbol{y}, \boldsymbol{\theta}, \omega)}{\partial \boldsymbol{\theta}} \\ \dfrac{\partial l(\boldsymbol{y}, \boldsymbol{\theta}, \omega)}{\partial \omega} \end{bmatrix}$$

The expected Fisher information matrix $I(\boldsymbol{\theta}, \omega)$ can be partitioned as

$$I(\boldsymbol{\theta}, \omega) = \begin{bmatrix} I_{\boldsymbol{\theta}\boldsymbol{\theta}}(\boldsymbol{\theta}, \omega) & I_{\boldsymbol{\theta}\omega}(\boldsymbol{\theta}, \omega) \\ I_{\omega\boldsymbol{\theta}}(\boldsymbol{\theta}, \omega) & I_{\omega\omega}(\boldsymbol{\theta}, \omega) \end{bmatrix},$$

where the elements $I_{\boldsymbol{\theta}\boldsymbol{\theta}}(\boldsymbol{\theta}, \omega)$, $I_{\boldsymbol{\theta}\omega}(\boldsymbol{\theta}, \omega)$=$I_{\omega\boldsymbol{\theta}}(\boldsymbol{\theta}, \omega)^T$and $I_{\omega\omega}(\boldsymbol{\theta}, \omega)$ are

$$-E\left[\frac{\partial^2 l(\boldsymbol{y}, \boldsymbol{\theta}, \omega)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T}\right], \quad -E\left[\frac{\partial^2 l(\boldsymbol{y}, \boldsymbol{\theta}, \omega)}{\partial \boldsymbol{\theta} \partial \omega}\right] \quad \text{and} \quad -E\left[\frac{\partial^2 l(\boldsymbol{y}, \boldsymbol{\theta}, \omega)}{\partial \omega^2}\right], \text{ respectively.}$$

Under the null hypothesis $\omega = 0$, the score test statistic is then

$$S_T = S_{\omega}(\hat{\boldsymbol{\theta}}, 0)^T \hat{V}^{-1} S_{\omega}(\hat{\boldsymbol{\theta}}, 0),$$

where $S_{\omega}(\hat{\boldsymbol{\theta}}, 0) = \left[\dfrac{\partial l(\boldsymbol{y}, \boldsymbol{\theta}, \omega)}{\partial \omega}\right]_{\omega=0, \boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = \sum_{i=1}^{n} \left\{ \dfrac{I_{(y_i=0)}}{g_i(0;\hat{\boldsymbol{\theta}})} - 1 \right\}$
and $\hat{V} = I_{\omega\omega}(\hat{\boldsymbol{\theta}}, 0) - I_{\boldsymbol{\theta}\omega}(\hat{\boldsymbol{\theta}}, 0)^T I_{\boldsymbol{\theta}\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}}, 0)^{-1} I_{\boldsymbol{\theta}\omega}(\hat{\boldsymbol{\theta}}, 0)$. Here $\hat{\boldsymbol{\theta}}$ is the maximum likelihood estimate of $\boldsymbol{\theta}$ under the null hypothesis. With the assumption of constant mixing weight, this score test statistic will have an asymptotic χ_1^2 distribution under the null hypothesis.

As examples, under the Zero-Inflated Poisson model assuming a constant mixing weight, the score test statistic is

$$S_T = \frac{\left\{ \sum_{i=1}^{n} \left[\frac{I_{(y_i=0)}}{e^{\hat{\lambda}_i}} - 1 \right] \right\}^2}{\sum_{i=1}^{n} (e^{\hat{\lambda}_i} - 1) - \hat{\lambda}^T X [X^T diag(\hat{\lambda}) X]^{-1} X^T \hat{\lambda}},$$

where $\hat{\lambda} = (\hat{\lambda_1}, \hat{\lambda_2}, ..., \hat{\lambda_n})$, $\hat{\lambda}_i$ is the estimated Poisson mean such that $\hat{\lambda}_i = e^{X_i^T \hat{\beta}}$ and $\hat{\beta}$ is the maximum likelihood estimate of β under the null hypothesis and X is a matrix of covariates (for details, see Van den Broek, 1995).

Under the Zero-Inflated Negative Binomial model assuming a constant mixing weight, the score function is

$$S_\omega(\hat{\boldsymbol{\theta}}, 0) = \sum_{i=1}^{n} \left[\frac{I_{(y_i=0)}}{(1+\hat{\alpha}\hat{\lambda}_i)^{-\hat{\alpha}^{-1}}} - 1 \right],$$

where $\hat{\boldsymbol{\theta}} = (\hat{\alpha}, \hat{\boldsymbol{\beta}})$, the maximum likelihood estimates under H_0. This involves estimating extra parameter α which is the dispersion parameter. This is a special case of the score test proposed by Jansakul and Hinde (2008) under the Zero-Inflated Negative Binomial model relating mixing weight to covariates, where the mixing weight is assumed as a constant.

2.4 Misspecifications

Two common types of misspecification on the homogeneity tests under zero-inflated models are (1) Misspecification of the mean function of the null model and (2) Misspecification of the baseline distribution of the null model.

The mean of the null model is often related to covariates through an appropriate link function. However, in practice, the mean function could be misspecified when the nuisance parameters are misspecified. For instance, if the true mean $\lambda^* = \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2)$ where X_1 and X_2 are two independent covariates, the working mean in the test may define as $\lambda = \exp(\beta_0 + \beta_1 X_1)$. Mean of the null model could be misspecified also when the link function is not appropriate to relate covariates to the mean.

Baseline distribution under the null model could be also misspecified in the testing procedure. For instance, if the true underlying distribution under the null hypothesis is Negative Binomial distribution, but the working baseline distribution is assumed as the Poisson distribution. But it is worth to mention that if the true baseline distribution is Poisson but assumed as Negative Binomial in the testing procedure, it is not a real misspecification. It is known that Poisson is a special case of Negative Binomial when the dispersion parameter goes to 0.

We investigate the impact of misspecification of null model on the validity of homogeneity score test under zero-inflated models. Specifically, we evaluate the size of the test results from a model which is misspecified in terms of the mean and the baseline distribution under the null hypothesis. Simulation studies are conducted separately for the two cases, first, to evaluate the size of the test when the mean is misspecified but the baseline distribution is correctly specified and secondly to evaluate the size of the test when the baseline distribution is misspecified but the mean is correctly specified.

2.5 Numerical study

We investigate the impact of misspecification on the validity of score test using simulation studies. Throughout the simulations, we perform the score test that assumes a constant mixing weight in the zero-inflated model. The validity of the score test is evaluated based on the agreement with the asymptotic χ_1^2 distribution under the null hypothesis. Empirical type I error rate of the test is evaluated under the two types of misspecification mentioned previously. As a comparison, type I error rate under the well-specified null model is also evaluated.

2.5.1 Misspecification of mean function of the null model

To evaluate the empirical type I error rate of score test under the misspecification of mean function, we generate data from a homogeneous Poisson distribution with sample sizes 50, 100, 200 and 1000. The true underlying Poisson mean is assumed as $\lambda^* = \exp\{0.75 - 1.45X_1 - 0.8X_2\}$ where X_1 and X_2 are two independent covariates with $X_1 \sim$ U(0,1) and $X_2 \sim$ Bin(n,0.6). All simulations are replicated 1000 times.

Three types of misspecification of Poisson mean are considered: (1) Excluding covariates which should be included in the mean function, (2) Including covariates which should not be included in the mean function and (3) Using a different functional form of the mean. Empirical type I error rates of the score tests are evaluated under different working mean functions which correspond to those misspecifications.

Table 2.1 presents the empirical type I error rates at 5% nominal level of the score test under the true model and different working models. We can clearly see that the size of the score test is stable around the nominal level ($\alpha = 0.05$) under the null model with correctly specified mean function. In contrast, the size can not be well maintained at the nominal level when the mean function is misspecified in the null model. For an over-fitted mean function of the null model, the size is stable compared to these models with under-fitted mean functions. However, the over-fitted mean function may result in unnecessary loss of power to detect the heterogeneity in the population because of the loss of efficiency. When a different functional form of the mean is used, for instance, if a non-linear mean function is used instead of a true linear mean function of the null model, the size of the test can not be maintained at the nominal level.

Table 2.1: *Empirical sizes of the score test statistics at α=0.05 based on 1,000 samples generated from Poisson with true mean* $\log(\lambda^*) = 0.75 - 1.45X_1 - 0.8X_2$.

Models for $\log(\lambda)$	n=50	n=100	n=200	n=1000
Correct Specification:				
$\log(\lambda) = \beta_0 + \beta_1 X_1 + \beta_2 X_2$	0.050	0.053	0.050	0.048
Over-fitted model:				
$\log(\lambda) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$	0.062	0.057	0.047	0.046
Misspecification:				
$\log(\lambda) = \beta_0$	0.177	0.292	0.520	0.986
$\log(\lambda) = \beta_0 + \beta_1 X_1$	0.074	0.123	0.161	0.633
$\log(\lambda) = \beta_0 + \beta_2 X_2$	0.077	0.116	0.193	0.675
$\log(\lambda) = \beta_0 + \beta_3 X_3$	0.160	0.271	0.511	0.985
$\log(\lambda) = \beta_0 + \beta_1 X_1 + \beta_3 X_3$	0.058	0.108	0.154	0.623
$\log(\lambda) = \beta_0 + \beta_2 X_2 + \beta_3 X_3$	0.069	0.109	0.177	0.663
$\log(\lambda) = \exp\left\{\frac{\beta_1 X_1}{1+e^{-(\beta_0+\beta_2 X_2)}}\right\}$	0.755	0.954	0.999	1.000
$\log(\lambda) = \exp\left\{\beta_0 + \beta_1 e^{-\beta_2 X_2}\right\}$	0.589	0.827	0.931	0.947

$X_1 \sim U(0,1)$, $X_2 \sim Bin(n, 0.6)$ and $X_3 \sim N(0,1)$ with $X_3 \in$ (-1,1).

2.5.2 Misspecification of baseline distribution

We evaluate the impact of the misspecification of baseline distribution on the validity of score test in two different scenarios: (1) Under the framework of Zero-Inflated Poisson model and (2) under the Zero-Inflated Negative Binomial model. We generate data with sample sizes 50, 100, 200, 500, 1000 and pre-specified mean functions $\lambda^* = \exp\{-0.15\}$, $\lambda^* = \exp\{-0.15+0.1X_1\}$, $\lambda^* = \exp\{-0.15+0.1X_1-0.25X_2\}$ and $\lambda^* = \exp\{-0.15+0.1X_1-0.25X_2+0.4X_3\}$.

In the first simulation study, we generate data from a Poisson distribution. For each simulated dataset, working model is considered with well-specified mean function but the working baseline distribution is considered as the Negative Binomial distribution. In other words, under the null hypothesis, the mean is considered as well-specified but the baseline distribution is not well-specified. Empirical type I error rates are presented in Table 2.2. We can see that all tests have well controlled type I error rates at 5% nominal level when the null model is well specified. When the baseline distribution is specified as Negative Binomial, interestingly, the empirical sizes tend to be conservative but relatively stable. This result is expected because it is known that Poisson is a special case of Negative Binomial. Therefore, the tests are conservative simply due to the efficiency issue.

Table 2.2: *Empirical sizes of the score test statistics with well specified working mean function based on 1,000 samples generated from Poisson regression model with mean λ^* at α=0.05.*

		n				
$\log(\lambda^*)$	Working Model	50	100	200	500	1000
-0.15	Poisson	0.043	0.045	0.048	0.053	0.050
	NB	0.026	0.024	0.028	0.025	0.025
$-0.15+0.1X_1$	Poisson	0.047	0.046	0.049	0.045	0.053
	NB	0.029	0.027	0.024	0.026	0.025
$-0.15+0.1X_1-0.25X_2$	Poisson	0.046	0.045	0.046	0.047	0.048
	NB	0.038	0.029	0.029	0.028	0.023
$-0.15+0.1X_1-0.25X_2+0.4X_3$	Poisson	0.045	0.051	0.049	0.050	0.052
	NB	0.039	0.038	0.031	0.028	0.026

$X_1 \sim U(0,1)$, $X_2 \sim Bin(n, 0.6)$ and $X_3 \sim N(0,1)$ with $X_3 \in$ (-1,1).

In the second simulation study, we generate data from a Negative Binomial distribution. For each simulated dataset, working model is considered with the well-specified mean function but the baseline distribution in the working model is assumed as the Poisson distribution. Table 2.3 reports the empirical sizes of the score test when the baseline distribution is misspecified. It is worth to note that, if the working baseline distribution is assumed to be Poisson but the true underlying distribution is Negative Binomial distribution, the type I error rates can not be well-maintained at the 5% nominal level, type I error rate increases as the sample size increases.

Table 2.3: *Empirical sizes of the score test statistics with well specified working mean function based on 1,000 samples generated from Negative Binomial regression model with mean λ^* at α=0.05.*

		n				
$\log(\lambda^*)$	Working Model	50	100	200	500	1000
-0.15	NB	0.045	0.049	0.046	0.052	0.048
	Poisson	0.562	0.814	0.978	1.000	1.000
$-0.15 + 0.1X_1$	NB	0.047	0.051	0.052	0.054	0.052
	Poisson	0.554	0.872	0.987	1.000	1.000
$-0.15 + 0.1X_1 - 0.25X_2$	NB	0.045	0.047	0.045	0.045	0.050
	Poisson	0.447	0.759	0.962	1.000	1.000
$-0.15 + 0.1X_1 - 0.25X_2 + 0.4X_3$	NB	0.043	0.051	0.050	0.049	0.048
	Poisson	0.372	0.753	0.967	1.000	1.000

$X_1 \sim U(0,1)$, $X_2 \sim Bin(n, 0.6)$ and $X_3 \sim N(0,1)$ with $X_3 \in$ (-1,1).

2.6 Discussion

Score tests for evaluating heterogeneity under zero-inflated models primarily rely on the assumption of well-specified model under the null hypothesis. In this chapter, we evaluated the impact of misspecification on the validity of score test under zero-inflated models. Simulation results indicate that the empirical type I error rates can not be well maintained at the nominal level when the mean function is incorrectly specified under the null model. Additionally, score test tends to be very unstable under severe misspecification such as when the baseline distribution is assumed as Poisson but the true underlying distribution is Negative Binomial. Therefore, a close attention to the model under the null hypothesis is needed when performing the homogeneity score test under Zero-Inflated models.

Chapter 3

A robust score test of homogeneity for zero-inflated count data

3.1 Introduction

Score test has been suggested by many authors as a homogeneity test under the class of zero-inflated models due to the fact that it only requires the model fit under the null model (for example, see Van den Broek, 1995; Deng and Paul, 2000; Jansakul and Hinde, 2002; Todem et al., 2012). However, the general limitation of score test is that the score test relies on the assumption of a well-specified null model. As discussed in Chapter 2, when the null models are not correctly specified, score tests could lead to unreliable statistical inferences.

In this chapter, we propose a new score test of homogeneity which is robust to the misspecifications under zero-inflated models. Technically, the test is developed under the framework of Poisson-Gamma mixture model which provides a general framework to incorporate the baseline distributions under the Zero-Inflated Poisson (ZIP) or Zero-Inflated Negative Binomial (ZINB) models. Our test can be performed without specifying the mean function and hence avoid the impact of any possible misspecification of mean under the null model. The proposed test assumes a more general model as the baseline model, thus, it is robust to the misspecification of the baseline distribution.

The rest of this chapter is organized as follows. In Section 3.2, we introduce the new robust test for evaluating heterogeneity under zero-inflated models. The performance of the proposed test is evaluated by simulation studies under the possible misspecification of the mean and baseline distribution in Section 3.3. The proposed test is then applied to dental caries data from the Detroit Dental Health Project (DDHP) and Girl Scout data from Scouting Nutrition and Activity Program (SNAP) by Rosenkranz et al.(2010). We compare the performance of the proposed test with the existing homogeneity tests in Section 3.4. Discussion based on the simulation study and the real data analysis is included in Section 3.5.

3.2 Robust homogeneity test for zero-inflated models

We are interested in evaluating the hypothesis of zero mixing weight in order to evaluate the homogeneous model against the heterogeneous model under the framework of zero-inflated models. Simulation results in Chapter 2 indicated that the homogeneity score test may result in unreliable statistical inferences under the misspecification of mean and the baseline distribution of the null model (see Table 2.1, 2.2, 2.3). To address the issue of misspecification, we develop a robust homogeneity score test for zero-inflated models. The proposed test is developed under the framework of Poisson-Gamma mixture model based on the modeling framework discussed by Kassahun et al. (2014).

3.2.1 Kassahun's model

A more general model for zero-inflated data which can accommodate zero-inflation, over-dispersion and correlation was proposed by Kassahun et al. (2014) which is called the Zero-inflated over-dispersed hierarchical Poisson model. Let Y_{ij} be the j^{th} outcome measured for subject i. The model of Kassahun et al. (2014) is,

$$Y_{ij}|\theta_{ij} \sim \begin{cases} 0 & \text{with prob } \pi_{ij} \\ \text{Poisson}(\lambda_{ij} = \theta_{ij} K_{ij}) & \text{with prob } 1 - \pi_{ij}, \end{cases}$$

where π_{ij} is the mixing weight, $\theta_{ij} \sim \text{Gamma}(\alpha, \beta)$ such that α and β are shape and scale parameters and $K_{ij} = \exp(X_{ij}^T \eta + Z_{ij}^T b_i)$ such that η is a vector of unknown coefficients with X_{ij} and Z_{ij} are vectors of known covariates. Random effects $b_i \sim N(0, D)$, where D is the associated covariance matrix.

3.2.2 The proposed robust test

In this study, we consider the case for independent data but the test can be extended to handle correlated data. Based on the Kassahun's model, we consider that Y_i $(i = 1, ..., n)$ are independent observations from a mixture of degenerate distribution at 0 and a Poisson distribution with a random mean Λ_i such that,

$$Y_i|\Lambda_i \sim \begin{cases} 0 & \text{with prob } \omega_i \\ \text{Poisson}(\Lambda_i) & \text{with prob } 1 - \omega_i, \end{cases}$$

where ω_i is the mixture probability and $\Lambda_i \sim \text{Gamma}(\alpha,\ \beta)$ with α and β are shape and scale parameters respectively. The marginal distribution $f_Y(y)$ is given by,

$$f_Y(y) = \int_0^\infty f_{Y,\Lambda}(y, \lambda)\, d\lambda = \int_0^\infty f_{Y|\Lambda}(y|\lambda)\, f_\Lambda(\lambda)\, d\lambda.$$

Thus, the zero-inflated distribution can be re-expressed as

$$P(Y_i = y_i) = \begin{cases} \omega_i + (1-\omega_i)\; f_{Y_i}(0) & \text{if } y_i = 0 \\ (1-\omega_i)\; f_{Y_i}(y_i) & \text{if } y_i = 1, 2, 3, ..., \end{cases} \tag{3.1}$$

where

$$f_{Y_i}(y_i) = \frac{\Gamma(y_i + \alpha)}{y_i!\,\Gamma(\alpha)} \left(\frac{\beta}{1+\beta}\right)^{y_i} \left(\frac{1}{1+\beta}\right)^{\alpha}, \;\; y_i = 0, 1, 2, 3, \ldots$$

It is worth to mention that the baseline distribution is a Negative Binomial.

Assuming a sample of independent observations $y_1, ..., y_n$, the log-likelihood function $l(\boldsymbol{y}, \alpha, \beta, \omega)$ corresponds to probability mass function (3.1) with $\omega_i = \omega$ for all i is,

$$l(\boldsymbol{y}, \alpha, \beta, \omega) = \sum_{i=1}^{n} \Big[\mathrm{I}_{(y_i=0)} \log\big\{\omega + (1-\omega)(1+\beta)^{-\alpha}\big\} + \mathrm{I}_{(y_i>0)} \big\{ \log(1-\omega) + \log\Gamma(y_i+\alpha) - \log\Gamma(\alpha) \\ - \log\Gamma(y_i+1) + y_i \log\beta - (y_i+\alpha)\log(1+\beta)\big\} \Big]. \tag{3.2}$$

The score vector can be derived by the first-order derivatives of log-likelihood function and is given by,

$$S(\alpha, \beta, \omega) = \begin{bmatrix} S_\alpha(\alpha, \beta, \omega) \\ S_\beta(\alpha, \beta, \omega) \\ S_\omega(\alpha, \beta, \omega) \end{bmatrix} = \begin{bmatrix} \dfrac{\partial l(\boldsymbol{y}, \alpha, \beta, \omega)}{\partial \alpha} \\ \dfrac{\partial l(\boldsymbol{y}, \alpha, \beta, \omega)}{\partial \beta} \\ \dfrac{\partial l(\boldsymbol{y}, \alpha, \beta, \omega)}{\partial \omega} \end{bmatrix},$$

where

$$\frac{\partial l(\boldsymbol{y}, \alpha, \beta, \omega)}{\partial \alpha} = \sum_{i=1}^{n} \left\{ \mathrm{I}_{(y_i=0)} \frac{-(1-\omega)(1+\beta)^{-\alpha}\log(1+\beta)}{[\omega + (1-\omega)(1+\beta)^{-\alpha}]} - \mathrm{I}_{(y_i>0)}[\mathrm{dg}(y_i+\alpha) - dg(\alpha) - \log(1+\beta)] \right\}$$

$$\frac{\partial l(\boldsymbol{y}, \alpha, \beta, \omega)}{\partial \beta} = \sum_{i=1}^{n} \left\{ \mathrm{I}_{(y_i=0)} \frac{-\alpha(1-\omega)(1+\beta)^{-\alpha-1}}{[\omega + (1-\omega)(1+\beta)^{-\alpha}]} - \mathrm{I}_{(y_i>0)} \left[\frac{y_i}{\beta} - \frac{(y_i+\alpha)}{(1+\beta)}\right] \right\}$$

$$\frac{\partial l(\boldsymbol{y}, \alpha, \beta, \omega)}{\partial \omega} = \sum_{i=1}^{n} \left\{ \mathrm{I}_{(y_i=0)} \frac{[1 - (1+\beta)^{-\alpha}]}{[\omega + (1-\omega)(1+\beta)^{-\alpha}]} - \mathrm{I}_{(y_i>0)} \frac{1}{(1-\omega)} \right\},$$

where $\mathrm{dg}(y_i+\alpha) = \dfrac{\partial \log\Gamma(y_i+\alpha)}{\partial\alpha}$ and $\mathrm{dg}(\alpha) = \dfrac{\partial \log\Gamma(\alpha)}{\partial\alpha}$ (see, details in Appendix A).

The expected Fisher information matrix $I(\alpha,\beta,\omega)$ can be partitioned as

$$I(\alpha,\beta,\omega) = \begin{bmatrix} I_{\alpha\alpha} & I_{\alpha\beta} & I_{\alpha\omega} \\ I_{\beta\alpha} & I_{\beta\beta} & I_{\beta\omega} \\ I_{\omega\alpha} & I_{\omega\beta} & I_{\omega\omega} \end{bmatrix},$$

where the elements $I_{\alpha\alpha}$, $I_{\beta\alpha}$=$I_{\alpha\beta}^T$, $I_{\omega\alpha}$=$I_{\alpha\omega}^T$, $I_{\omega\beta}$=$I_{\beta\omega}^T$, $I_{\beta\beta}$ and $I_{\omega\omega}$ are such that,

$$I_{\alpha\alpha} = -\mathrm{E}\left[\frac{\partial^2 l(\boldsymbol{y},\alpha,\beta,\omega)}{\partial\alpha^2}\right], \quad I_{\beta\alpha} = I_{\alpha\beta}^T = -\mathrm{E}\left[\frac{\partial^2 l(\boldsymbol{y},\alpha,\beta,\omega)}{\partial\alpha\partial\beta}\right],$$
$$I_{\omega\alpha} = I_{\alpha\omega}^T = -\mathrm{E}\left[\frac{\partial^2 l(\boldsymbol{y},\alpha,\beta,\omega)}{\partial\alpha\partial\omega}\right], I_{\beta\beta} = -\mathrm{E}\left[\frac{\partial^2 l(\boldsymbol{y},\alpha,\beta,\omega)}{\partial\beta^2}\right] \quad \text{and} \quad I_{\omega\omega} = -\mathrm{E}\left[\frac{\partial^2 l(\boldsymbol{y},\boldsymbol{\theta},\omega)}{\partial\omega^2}\right],$$
respectively (see, details in Appendix A).

This requires second derivatives of the log-likelihood function (3.2) with respect to α,β and ω and the facts $\mathrm{E}[\mathrm{I}_{(y_i=0)}] = [\omega+(1-\omega)(1+\beta)^{-\alpha}]$ and $\mathrm{E}[\mathrm{I}_{(y_i>0)}] = (1-\omega)[1-(1+\beta)^{-\alpha}]$. Under the null hypothesis $H_0: \omega = 0$, the proposed robust score test is then,

$$S = S_\omega^T \ \hat{V_\omega}^{-1} \ S_\omega \ ,$$

where

$$S_\omega = \sum_{i=1}^{n} \left[\mathrm{I}_{(y_i=0)}\left\{(1+\hat\beta)^{\hat\alpha}-1\right\}\right]$$

and

$$\hat{V} = I_{\omega\omega}(\hat\alpha,\hat\beta,0) - I_{\omega\alpha\beta}(\hat\alpha,\hat\beta,0)' I_{\alpha\beta}^{-1}(\hat\alpha,\hat\beta,0) I_{\omega\alpha\beta}(\hat\alpha,\hat\beta,0) \quad \text{such that,}$$

$$I_{\omega\alpha\beta}(\hat\alpha,\hat\beta,0) = \begin{bmatrix} I_{\omega\alpha}(\hat\alpha,\hat\beta,0) \\ I_{\omega\beta}(\hat\alpha,\hat\beta,0) \end{bmatrix} \quad \text{and} \quad I_{\alpha\beta}(\hat\alpha,\hat\beta,0) = \begin{bmatrix} I_{\alpha\alpha}(\hat\alpha,\hat\beta,0) & I_{\alpha\beta}(\hat\alpha,\hat\beta,0) \\ I_{\beta\alpha}(\hat\alpha,\hat\beta,0) & I_{\beta\beta}(\hat\alpha,\hat\beta,0) \end{bmatrix},$$

where $I_{\omega\omega}(\hat\alpha,\hat\beta,0), I_{\omega\alpha}(\hat\alpha,\hat\beta,0), I_{\omega\beta}(\hat\alpha,\hat\beta,0), I_{\alpha\alpha}(\hat\alpha,\hat\beta,0), I_{\beta\beta}(\hat\alpha,\hat\beta,0), I_{\alpha\beta}(\hat\alpha,\hat\beta,0)$ and $I_{\beta\alpha}(\hat\alpha,\hat\beta,0)$ are elements of the Fisher information matrix evaluated under H_0. Under the null hypothesis, this score test statistic will have an asymptotic χ_1^2 distribution.

The proposed test assumes the mean of the baseline distribution is a random variable which can be described by some density. Since the over-dispersion could be involved in data which may not be adequately described by the null model under existing testing procedures, the proposed test would be able to capture the potential over-dispersion in data appropriately. Our test avoids the possible misspecification of mean function because the mean is considered as a random variable rather than assuming as a function of covariates through a link function. The baseline distribution in the proposed test is the Poisson-Gamma mixture model and technically, the Poisson-Gamma mixture model can incorporate the Negative Binomial distribution and the Poisson distribution. Hence, the proposed test would be robust to the misspecification of baseline distributions under the Zero-Inflated Negative Binomial and Zero-Inflated Poisson models.

3.3 Simulation Study

Empirical performance of the proposed test is evaluated by investigating the type I error rates of the test under the misspecification of mean and baseline distribution. To evaluate the size of the test, data are generated with sample sizes 50, 100, 200, 400, 800 and 1500 under several mean functions of the baseline distribution: $\log(\lambda^*) = 0.6$, $\log(\lambda^*) = 0.6 + 0.45X_1$ and $\log(\lambda^*) = 0.6 + 0.45X_1 - 0.2X_2$, where X_1 and X_2 are two independent covariates. The covariate X_1 is a continuous variable with uniformly distributed values on (0,1) and X_2 is a truncated normal random variable with values on (-1,1).

In the first simulation study, data are generated from zero-inflated Negative Binomial distribution with the dispersion parameter α=0.8 and various true mean functions. Simulation is conducted with 1000 Monte Carlo samples and empirical type I error rate of the proposed robust test is evaluated at the nominal level 0.05. The power of the test is evaluated under the non-zero constant mixing weights ($\omega^* = 0.1$, 0.2) and results are given in Table 3.1.

From Table 3.1, we can clearly see that the proposed robust test maintains the size well around the nominal level 0.05. The power of the test increases as the sample size increases or the mixing weight (ω^*) increases. Clearly, the test is robust to any mean function used

Table 3.1: *Empirical sizes and power of the robust score test statistic based on 1,000 samples generated from zero-inflated Negative Binomial regression model with mean λ^*, at the nominal level 0.05.*

		n					
$\log(\lambda^*)$	ω^*	50	100	200	400	800	1500
0.6	$\omega^* = 0$	0.075	0.069	0.058	0.052	0.054	0.052
	$\omega^* = 0.1$	0.127	0.161	0.184	0.221	0.282	0.461
	$\omega^* = 0.2$	0.139	0.223	0.317	0.478	0.691	0.880
$0.6 + 0.45\ X_1$	$\omega^* = 0$	0.072	0.066	0.058	0.051	0.047	0.048
	$\omega^* = 0.1$	0.155	0.167	0.198	0.218	0.349	0.524
	$\omega^* = 0.2$	0.168	0.275	0.378	0.546	0.762	0.944
$0.6 + 0.45\ X_1 - 0.2X_2$	$\omega^* = 0$	0.073	0.071	0.053	0.047	0.047	0.046
	$\omega^* = 0.1$	0.119	0.137	0.154	0.209	0.294	0.467
	$\omega^* = 0.2$	0.185	0.278	0.366	0.530	0.758	0.941
$0.6 + 0.2e^{-0.2X_2}$	$\omega^* = 0$	0.078	0.074	0.079	0.063	0.049	0.051
	$\omega^* = 0.1$	0.119	0.135	0.160	0.189	0.246	0.283
	$\omega^* = 0.2$	0.129	0.172	0.279	0.377	0.504	0.710

$X_1 \sim U(0,1)$ and $X_2 \sim N(0,1)$ with $X_2 \in$ (-1,1).

in the data generating process and performs well in all cases.

Second simulation study is conducted for data generated from zero-inflated Poisson distribution under the same mean functions and the results are given in Table 3.2.

In the Table 3.2, the size of the test tends to be conservative but remains stable as the sample size increases. This result is not surprising as the proposed test has developed under the framework of Poisson-Gamma mixture model which can provide a more general framework to incorporate Poisson distribution. Therefore, the tests are conservative simply due to the efficiency issue. The power of the test increases as the sample size increases and as the mixing weight increases from 0.1 to 0.2.

Table 3.2: *Empirical sizes and power of the robust score test statistics based on 1,000 samples generated from zero-inflated Poisson regression model with mean λ^*, at the nominal level 0.05.*

		n					
$\log(\lambda^*)$	ω^*	50	100	200	400	800	1500
0.6	$\omega^* = 0$	0.045	0.031	0.031	0.033	0.043	0.051
	$\omega^* = 0.1$	0.062	0.074	0.093	0.154	0.411	0.754
	$\omega^* = 0.2$	0.169	0.332	0.527	0.737	0.787	0.793
$0.6 + 0.45X_1$	$\omega^* = 0$	0.029	0.028	0.026	0.031	0.023	0.025
	$\omega^* = 0.1$	0.096	0.158	0.254	0.505	0.771	0.889
	$\omega^* = 0.2$	0.351	0.605	0.793	0.865	0.886	0.893
$0.6 + 0.45X_1 - 0.2X_2$	$\omega^* = 0$	0.020	0.024	0.023	0.022	0.020	0.024
	$\omega^* = 0.1$	0.108	0.169	0.276	0.496	0.774	0.890
	$\omega^* = 0.2$	0.355	0.604	0.812	0.888	0.902	0.903
$0.6 + 0.2e^{-0.2X_2}$	$\omega^* = 0$	0.038	0.036	0.039	0.031	0.035	0.032
	$\omega^* = 0.1$	0.091	0.181	0.252	0.301	0.451	0.642
	$\omega^* = 0.2$	0.126	0.256	0.494	0.707	0.827	0.848

$X_1 \sim U(0,1)$ and $X_2 \sim N(0,1)$ with $X_2 \in$ (-1,1).

3.4 Real Data Application

3.4.1 Dental Caries Data

To illustrate the proposed test we use dental caries data from the Detroit Dental Health Project (DDHP) which was designed to assess dental caries severity of children (Tellez et al., 2006). The target population of this study is low-income African-American children under age of six and their main caregivers who resided in Detroit, Michigan. Although the study is longitudinal in nature, we use cross-sectional data of 897 children surveyed in the first wave of examinations conducted between 2002 and 2003. In our analysis, the outcome variable is considered as DS which represents the number of decayed tooth surfaces and sugar intake and age are used as covariates. We compare the results of proposed test with the results from existing testing procedures proposed by Van den Broek (1995) and Jansakul and Hinde (2008) assuming a constant mixing weight.

Score test proposed by Van den Broek (1995) is performed specifying the distribution under the null model as Poisson distribution and the mean of the null model as a function of covariates Age (child's age in years), SI (the child's sugar intake), and interaction Age$*$SI. The score test proposed by Jansakul and Hinde (2008) is conducted specifying the distribution under the null model as Negative Binomial distribution and with the same mean function. Our proposed robust test does not require the mean of the null model as a function of covariates and the test can be performed without specifying the baseline distribution specifically as Poisson distribution or Negative Binomial distribution. The results are given in Table 3.3.

Table 3.3: *Comparison of score test statistics, degrees of freedoms and associated p-values of homogeneity tests for dental caries data.*

	Van den Broek test	Jansakul and Hinde test	Robust test
df	1	1	1
Test statistic	30098.21	0.0106	19.0774
p-value	<0.001	0.9178	<0.001

The proposed test statistic and the score test proposed by Van den Broek (1995) reject the null hypothesis at 5% significance level, supporting the hypothesis of heterogeneity. The test statistic of Jansakul and Hinde (2008) also fails to reject the homogeneity hypothesis. Their test averages the mixing weights over the space of covariates under the assumption of constant mixing weight which may lose the power of the test when both deflation and inflation are present in the data. These data had been studied by Todem et al. (2012) and they revealed that there was an inflation of zero dental caries for younger children and deflation of zero dental caries for older children. When inflation and deflation at zero appear to be of the same magnitude, the test is not powerful enough to capture the heterogeneity in data (Todem et al., 2012). Interestingly, even under this scenario, our proposed test still can have power to detect the heterogeneity. It is worth to mention that rejecting the hypothesis of homogeneity under the Van den Broek (1995) test does not give evidence that the zero-inflated poisson model provides the best fit for the data.

3.4.2 Girl Scout Data

The use of proposed score test is illustrated with girl scout data from the study of Scouting Nutrition and Activity Program (SNAP) by Rosenkranz et al. (2010). The objective of the study is to evaluate the effectiveness of an intervention program designed to improve the physical activity and nutrition environment in Girl Scout troops. In this study, seven Girl Scout troops were randomized to intervention (3 troops with 34 girls) and to control (4 troops with 42 girls). In the intervention, troop leaders were trained to implement policies promoting physical activity and healthful eating opportunities at troop meetings. At each troop meeting during seven meetings from October, 2007 to April, 2008, a trained research assistant observed and counted Health and Nutrition Promotions implemented by troop leaders. Research assistants were blind to the condition of each troop.

For our analysis, we focus on the Nutrition Promotion activities. The number of nutrition promotions implemented by troop leaders in every 5 minutes at the troop meeting are considered as the count outcome variable. We conduct our proposed score test along with the score tests proposed by Van den Broek (1995) in which the baseline distribution is Poisson and by Jansakul and Hinde (2008) in which the baseline distribution is Negative Binomial, assuming a constant mixing weight. For Van den Broek's test and Jansakul and Hinde's test, the mean is considered as $\lambda = \exp(\beta_0 + \beta_1 X_1)$ where X_1 is an indicator variable for the intervention (1=intervention group, 0=control group).

Table 3.4: *Comparison of score test statistics, degrees of freedoms and associated p-values of homogeneity tests for Girl Scout data.*

	Van den Broek test	Jansakul and Hinde test	Robust test
df	1	1	1
Test statistic	56.8558	0.0302	0.057
p-value	<0.001	0.8620	0.8110

Score test proposed by Van den Broek (1995) rejects the null hypothesis at 5% significance level. Our proposed robust test and the score test proposed by Jansakul and Hinde (2008) fail to reject the null hypothesis, which are in favor of the homogeneous model under the null hypothesis. To evaluate these test results, we compare several count models by Akaike Information Criteria (AIC). The results are given in Table 3.5. Mean function in each count model is considered as $\lambda = \exp(\beta_0 + \beta_1 X_1)$ where X_1 is an indicator variable for the intervention. Compared to other count models, Negative Binomial (NB) model shows the smallest AIC. Additionally, the observed proportion of the count outcome and the fitted proportions by the Negative Binomial model indicate that NB model fits the data well (see, Figure 3.1). This result is consistent with the results of our robust test and Jansakul and Hinde's test. We also conduct a stratification analysis to investigate whether both zero inflation and deflation are present in data. Figure 3.2 presents the observed proportions and the predicted proportions for the intervention group and the control group. Moreover, we examined the mixing weights under both groups. The mixing weight for the intervention group was not significant (Wald test, p-value=0.7930) and it was same for the control group (Wald test, p-value=0.9913). Thus, there is no evidence of the presence of zero-inflation or zero-deflation in the data.

It is interested to mention that the coefficient of X_1 in NB is significant (estimated coefficient=1.965, p-value$<$ 0.001, Table 3.5) indicating that the intervention program has a significant effect on the Nutrition Promotions implemented by girl scout leaders.

3.5 Discussion

The proposed robust score test can address the misspecification issue under the class of zero-inflated models and performs reasonably well under the misspecifications. The test does not require a specification of mean function of the baseline distribution or the specification of baseline distribution specifically as Poisson or Negative Binomial when performing homogeneity score test. This test might not work well if the true baseline model is Binomial. It

Table 3.5: *Fits of different count models for Girl Scout data.*

Model	Estimated coefficient of intervention effect $\hat{\beta}_1$	$\hat{\omega}$	AIC
Poisson	1.965	-	354.86
	(< 0.001)		
NB	1.9646	-	271.61
	(< 0.001)		
ZIP	1.7754	0.4872	304.02
	(< 0.001)	(< 0.001)	
ZINB	2.0225	-2.6896	273.61
	(< 0.001)	(0.8609)	

Note: p-value of the Wald test is given in the parentheses.

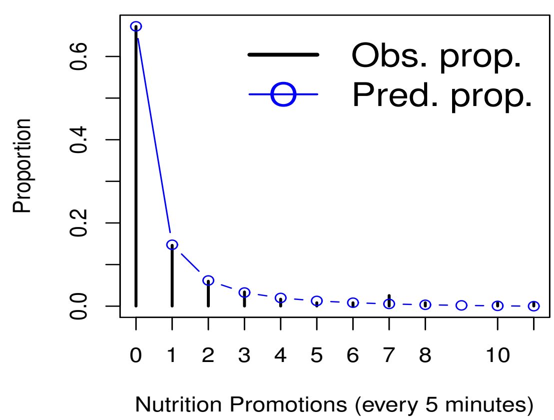

Figure 3.1: *Observed vs fitted proportions for Girl Scout data by the NB model.*

would be of interest to study the extensions of the test when the true underlying distribution is Binomial or any other count distribution. Further, this test can be extended to incorporate correlated count data. Since the true models are often unknown *a priori*, the robust test approach would be an efficient approach to detect heterogeneity under zero-inflated models.

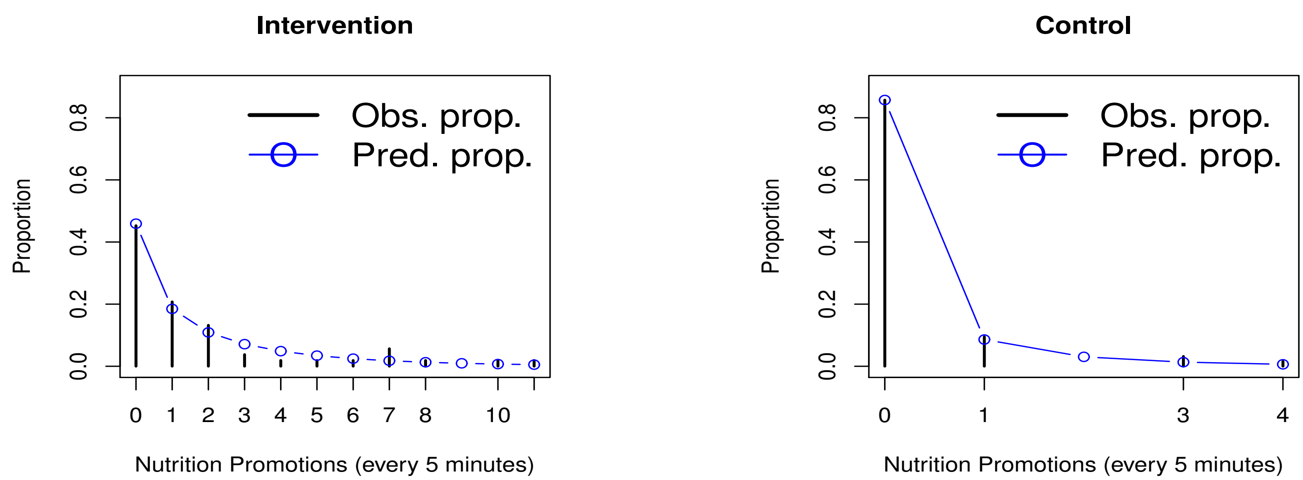

Figure 3.2: *Observed vs fitted proportions by the NB model for intervention group and control group in Girl Scout data.*

Chapter 4

A robust homogeneity test for correlated count data with excess zeros

4.1 Introduction

Count data with excess zeros often exhibit correlation due to the hierarchical nature of the study design or due to the data collection procedure with repeated measurements on subjects. These data are commonly encountered in Biomedical and Health Care applications where the count outcomes often represent repeated measurements on the patients or the outcomes of the patients who are typically clustered within the physicians or hospitals (for example, see Wang et al., 2002; Hur et al., 2002; Yau and Lee, 2001; Moulton et al., 2002; Min and Agresti, 2005).

Zero-Inflated models are often used in practice to accommodate these data with excess zeros. In literature, under the class of zero-inflated models, two approaches are commonly used to incorporate correlation in the data. In the first approach, random effects are included into the model and use the full likelihood for the parameter estimation (Hall, 2000; Wang et al., 2002; Hur et al., 2002; Yau et al., 2003; Lee et al., 2006) and in the second approach, a quasi-likelihood function derived from a working independence model is used with a sandwich estimator of the variance-covariance matrix (Moulton et al., 2002; Hsu et al., 2014).

In many applications of Zero-Inflated models, as a goodness of fit, homogeneity tests are used to test for the zero-inflation or deflation in data. In these tests, the mixture probability or the mixing weight that represents the extent of heterogeneity is examined at zero. Particularly, a zero mixing weight indicates that zero counts are generated from a homogeneous population. There is a considerable literature on testing for heterogeneity on cross-sectional zero-inflated data (Van den Broek, 1995; Deng and Paul, 2000; Jansakul and Hinde, 2002; Todem et al., 2012). For longitudinal zero-inflated count data, there is a homogeneity score test proposed by Xiang et al. (2006) and a homogeneity Wald test proposed by Hsu et al. (2014). However, in the classical testing procedure, it is required to have a correct model specification to provide valid statistical inferences. To our knowledge, there is no homogeneity test which is robust to the misspecifications under the zero-inflated models. In this chapter, we propose a homogeneity test which is robust to a misspecified conditional mean or an incorrect baseline distribution in the testing procedure. As an addition to the commonly used score test, we propose a Wald test which can be easily performed in practice with minimal programming efforts. Technically, the proposed Wald test is developed under the framework of Poisson-Gamma mixture model which can provide a more general framework to incorporate various baseline distributions without specifying the mean function. We accommodate the correlation in data using a quasi likelihood approach under the working independence model coupled with the sandwich estimator of the variance-covariance matrix to adjust for any misspecification of variance-covariance matrix while conducting the test.

The rest of this chapter is organized as follows. In Section 4.2, we describe the zero-inflated model for correlated count data and the quasi likelihood approach which is used in the proposed methodology. In Section 4.3, we propose a robust homogeneity Wald test for evaluating heterogeneity under zero-inflated models. We conduct numerical studies to evaluate the finite sample properties of the proposed Wald test with an application to the longitudinal dental caries data from Detroit Dental Health Project in Section 4.4.

4.2 Zero-inflated model for correlated count data

Suppose that Y_{ij} is a random variable which represents a count response of the j^{th} subject in the i^{th} cluster such that $i = 1, ..., n$ are independent clusters and $j = 1, ..., m_i$ are correlated observations within the cluster. The outcome Y_{ij} is assumed as drawn from a mixture of a degenerate distribution at 0 and a non-degenerate distribution $h_{ij}(y_{ij}; \boldsymbol{\theta})$ such that,

$$P(Y_{ij} = y_{ij}) = \begin{cases} \omega_{ij} + (1 - \omega_{ij})\, h_{ij}(0; \boldsymbol{\theta}) & \text{if } y_{ij} = 0 \\ (1 - \omega_{ij})\, h_{ij}(y_{ij}; \boldsymbol{\theta}) & \text{if } y_{ij} = 1, 2, 3, ..., \end{cases}$$

where ω_{ij} is the mixing weight which is constrained as $-h_{ij}(0; \boldsymbol{\theta})/(1 - h_{ij}(0; \boldsymbol{\theta})) \leq \omega_{ij} \leq 1$, for $i = 1, ..., n$, which accommodates both zero-inflation and zero-deflation (see Todem et al., 2012). In practice, $h_{ij}(.)$ is often defined as the Poisson distribution or Negative Binomial distribution (see for example, Farewell and Sprott, 1988; Lambert, 1992; Van den Broek, 1995; Jansakul and Hinde, 2008) and $\boldsymbol{\theta}$ is the finite parameter vector which dominates the baseline distribution $h_{ij}(.)$.

4.3 The proposed robust test for correlated zero-inflated count data

We are specifically interested in the two-sided hypotheses,

$$\mathrm{H}_0 : \omega_{ij} = 0 \ \text{ for all } i, j \qquad \text{vs} \qquad \mathrm{H}_1 : \omega_{ij} \neq 0 \ \text{ for some } i, j$$

Under the assumption of constant mixing weight ω for all i, j, we consider a working independence model in which Y_{ij}s are independent random observations from a mixture of degenerate distribution at 0 and a Poisson distribution such that,

$$Y_{ij} \sim \begin{cases} 0 & \text{with prob. } \omega \\ \text{Poisson}(\Lambda_{ij}) & \text{with prob. } 1-\omega, \end{cases}$$

where $\Lambda_{ij} \sim \text{Gamma}(\alpha, \beta)$ with α and β are shape and scale parameters.
By integrating over the gamma random effects, the marginal distribution $f_{Y_{ij}}(y_{ij})$ under the working independence model is

$$f_{Y_{ij}}(y_{ij}) = \frac{\Gamma(y_{ij}+\alpha)}{y_{ij}!\,\Gamma(\alpha)} \left(\frac{\beta}{1+\beta}\right)^{y_{ij}} \left(\frac{1}{1+\beta}\right)^{\alpha}, \;\; y_{ij} = 0, 1, 2, 3, ...$$

The quasi-likelihood function $L_{quasi}(y_i, \boldsymbol{\zeta})$ for cluster i under working independence model is

$$L_{quasi}(y_i, \boldsymbol{\zeta}) = \prod_{j=1}^{m_i} P(Y_{ij} = y_{ij}),$$

where $y_i = (y_{i1}, y_{i2}, ..., y_{im_i})$ and $\boldsymbol{\zeta} = (\omega, \alpha, \beta)^t$ is a finite parameter vector.

It is well known that with the correct specification of the marginal distribution $f_{Y_{ij}}(y_{ij})$, $\hat{\boldsymbol{\zeta}}$ would be a consistent estimate of $\boldsymbol{\zeta}^*$ which is the true value of $\boldsymbol{\zeta}$. Further, $\hat{\boldsymbol{\zeta}}$ is asymptotically normally distributed such that,

$$\sqrt{n}(\hat{\boldsymbol{\zeta}} - \boldsymbol{\zeta}^*) \xrightarrow{d} N(0, \Omega),$$

where $\Omega = \lim_{n\to\infty} n[I(\boldsymbol{\zeta}^*)]^{-1}\Upsilon(\boldsymbol{\zeta}^*)[I(\boldsymbol{\zeta}^*)]^{-1}$ with $[I(\boldsymbol{\zeta}^*)] = -\text{E}\left[\sum_{i=1}^{n}\sum_{j=1}^{m_i} \frac{\partial S_{ij}(\boldsymbol{\zeta})}{\partial \boldsymbol{\zeta}}\right]$, $\Upsilon(\boldsymbol{\zeta}^*) = \sum_{i=1}^{n} \text{E}\left[\sum_{j=1}^{m_i} S_{ij}(\boldsymbol{\zeta})\right]\left[\sum_{j=1}^{m_i} S_{ij}(\boldsymbol{\zeta})\right]^t$ and $S_{ij}(\boldsymbol{\zeta}) = \left[\frac{\partial \log(P(Y_{ij} = y_{ij}))}{\partial \boldsymbol{\zeta}}\right]$ evaluated at $\boldsymbol{\zeta} = \boldsymbol{\zeta}^*$.
A consistent estimator of the asymptotic variance-covariance matrix that can adjust for any misspecification of the true association structure is called the Huber sandwich estimator which was proposed by Huber (1967). In particular, the variance-covariance estimation is

combined with its corresponding empirical version in a sandwich form such that,

$$\hat{V}(\hat{\boldsymbol{\zeta}}) = [\hat{I}(\hat{\boldsymbol{\zeta}})]^{-1}\hat{\Upsilon}(\hat{\boldsymbol{\zeta}})[\hat{I}(\hat{\boldsymbol{\zeta}})]^{-1},$$

where $\hat{I}(\hat{\boldsymbol{\zeta}}) = -\left[\sum_{i=1}^{n}\sum_{j=1}^{m_i}\frac{\partial S_{ij}(\boldsymbol{\zeta})}{\partial \boldsymbol{\zeta}}\right]$ evaluated at $\boldsymbol{\zeta} = \hat{\boldsymbol{\zeta}}$, $\hat{\Upsilon}(\hat{\boldsymbol{\zeta}}) = \sum_{i=1}^{n}\left[\sum_{j=1}^{m_i} S_{ij}(\hat{\boldsymbol{\zeta}})\right]\left[\sum_{j=1}^{m_i} S_{ij}(\hat{\boldsymbol{\zeta}})\right]^t$ and $S_{ij}(\hat{\boldsymbol{\zeta}}) = \left[\frac{\partial \log(P(Y_{ij} = y_{ij}))}{\partial \boldsymbol{\zeta}}\right]$ evaluated at $\boldsymbol{\zeta} = \hat{\boldsymbol{\zeta}}$. The estimated variance-covariance matrix $\hat{V}(\hat{\boldsymbol{\zeta}})$ is the Huber sandwich estimator which yields consistent estimate of the asymptotic variance-covariance matrix Ω (for details, see Freedman, 2006).

Based on these settings, the hypotheses can be rewritten as a linear combination of the parameter vector $\boldsymbol{\zeta}$ such that,

$$\mathrm{H}_0 : C\boldsymbol{\zeta} = 0 \qquad \text{vs} \qquad \mathrm{H}_1 : C\boldsymbol{\zeta} \neq 0$$

where C is a contrast matrix such that C=[1 0 0] and $\boldsymbol{\zeta} = (\omega, \alpha, \beta)^t$.

A Wald test statistic of homogeneity that accommodates correlated count data can be defined as,

$$W_n = (C\hat{\boldsymbol{\zeta}})^t\{C\hat{V}(\hat{\boldsymbol{\zeta}})C^t\}^{-1}(C\hat{\boldsymbol{\zeta}}),$$

where $\hat{\boldsymbol{\zeta}}$ are consistent estimates of the parameter vector $\boldsymbol{\zeta}$ and $\hat{V}(\hat{\boldsymbol{\zeta}})$ is the Huber sandwich estimator of the variance-covariance matrix.

Under the null hypothesis, this Wald test statistic follows a χ^2 distribution with rank(C) degrees of freedom. Hence, under the assumption of constant mixing weight, the test statistic follows a χ^2_1 distribution.

4.4 Numerical Studies

We generate correlated count data from zero-inflated Negative Binomial and zero-inflated Poisson distributions separately. True means of the Poisson process and the Negative Binomial process take the forms $\log(\lambda^*_{ij}) = \beta^*_0$, $\log(\lambda^*_{ij}) = \beta^*_0 + \beta^*_1 x_{1i}$, $\log(\lambda^*_{ij}) = \beta^*_0 + \beta^*_1 x_{1i} + \beta^*_2 x_{2i}$,

$\log(\lambda_{ij}^*) = \beta_0^* + \beta_1^* x_{1i} + \beta_2^* x_{2i} + \beta_3^* x_{3i}$ and $\log(\lambda_{ij}^*) = \beta_0^* + \beta_2^* x_{2i} e^{-\beta_2^* x_{2i}}$ where $\beta_0^*, \beta_1^*, \beta_2^*, \beta_3^*$ are regression parameters, $i = 1, ..., n$ and $j = 1, ..., m$. The covariates x_{1i}, x_{2i} and x_{3i} are constant across all $j = 1, ..., m$ and generated from a uniform distribution on (0,1), Binomial distribution with success probability 0.3 and a standard normal distribution respectively. To introduce correlation between the data from the same cluster i, we use a Bernoulli random variable coupled with the independent zero-inflated Negative Binomial or zero-inflated Poisson distributions. Specifically, we define $Y_{ij} = (1 - b_{ij})Y_{ij}^{ind} + b_{ij}u_i$ where b_{ij} is an independent Bernoulli variable with success probability $p^* = 0.7$ and u_i is an independent random variable with the same distribution as Y_{ij}^{ind}. The observations from the same cluster i share the same u_i and induce correlation in data. For each parameter setting, we consider $n = 50, 100, 200, 500$ and 1000 clusters with $m = 3$ observations within each cluster.

(1) Impact of misspecification on the validity of classical Wald test

To evaluate the empirical type I error rate of the classical Wald test under the misspecification of the conditional mean, we generate correlated data from a Poisson distribution with sample sizes 50, 200, 500 and 1000. The true underlying Poisson mean under the working independence model is assumed as $\lambda^* = \exp\{0.3 + 0.5X_1 + 0.2X_2\}$. Empirical type I error rates of the classical Wald tests are evaluated under the correctly specified conditional mean and in the situations where the conditional mean is not well-specified. Empirical type I error rates are presented in Table 4.1. When the mean is well specified tests have well controlled type I error rates at 5% nominal level, however, when it is not well-specified, the classical Wald tests result in inflated type I error rates.

Table 4.1: *At the 5% significance level, the empirical type I error rates of the classical Wald test under the misspecification of the conditional mean with data generated from a Poisson process with true mean* $\log(\lambda^*) = 0.3 + 0.5X_1 + 0.2X_2$.

		n			
	Working mean	50	200	500	1000
misspecified:	$\log(\lambda) = \beta_0 + \beta_1 X_1$	0.080	0.095	0.110	0.190
correctly specified:	$\log(\lambda) = \beta_0 + \beta_1 X_1 + \beta_2 X_2$	0.043	0.052	0.051	0.049
misspecified:	$\log(\lambda) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$	0.071	0.083	0.090	0.089

$X_1 \sim U(0,1)$, $X_2 \sim Bin(n, 0.3)$ and $X_3 \sim N(0,1)$.

To evaluate the impact of misspecification of the baseline distribution, we generate correlated data from a Negative Binomial distribution with sample sizes 50, 200, 500 and 1000 and mean functions $\lambda^* = \exp\{1.5 + 0.35X_1\}$, $\lambda^* = \exp\{1.5 + 0.35X_1 + 0.2X_2\}$ and $\lambda^* = \exp\{1.5 + 0.35X_1 + 0.2X_2 + 0.3X_3\}$. For each simulated dataset, working model is considered with well specified mean function but the working baseline distribution is considered as the Poisson distribution. Empirical type I error rates are presented in Table 4.2. We can see that the test has a well controlled type I error rate at 5% nominal level when the null model is well specified. However, when it is not well-specified, the classical Wald test result in inflated type I error rate.

Table 4.2: *At the 5% significance level, the empirical type I error rates of the classical Wald test under the misspecification of the baseline distribution.*

True baseline distribution: NB process	n			
Working baseline distribution: Poisson process	50	200	500	1000
$\log(\lambda) = \beta_0 + \beta_1 X_1$	0.910	0.976	0.996	0.998
$\log(\lambda) = \beta_0 + \beta_1 X_1 + \beta_2 X_2$	0.890	0.979	0.990	0.995
$\log(\lambda) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$	0.081	0.440	0.830	0.962

$X_1 \sim U(0,1)$, $X_2 \sim Bin(n, 0.3)$ and $X_3 \sim N(0,1)$.

(2) Empirical size and power of the proposed robust test

We conduct simulation studies to evaluate the proposed Wald test in terms of the type I error rate and the power to detect heterogeneity in the population under the misspecifications. The robust Wald test is conducted assuming the constant mixing weight under the alternative and all simulations are replicated 1000 times. We evaluate the type I error rate of the proposed Wald test at 5% nominal level when the true mean function takes various linear and non-linear forms. Size of the test is evaluated separately for the cases where the true baseline distribution is Poisson or Negative Binomial. We further study the empirical power of the test to detect the heterogeneity in the population and evaluate the power of the test when the true mixing weight is constant and when it depends on covariates. Results are given in Table 4.3 and Table 4.4 for the correlated zero-inflated Negative Binomial and correlated zero-inflated Poisson data respectively. This quasi-likelihood approach can be easily implemented in SAS NLMIXED and the parameter estimation can be done routinely with the use of 'EMPIRICAL' option to obtain the sandwich estimator along with the options 'QPOINTS=1' and 'NOAD' for using the working independence model (Hsu et al., 2014).

From Table 4.3, we can see that the proposed Wald test maintains the size well around the nominal level 0.05. The power of the test increases as the sample size increases or the mixing weight (ω^*) increases. The test is capable enough to detect the heterogeneity in the population when the true mixing weight is constant or when it depends on covariates. Further, the test is robust to any linear or non-linear mean function used in the data generating process.

From Table 4.4, when the true baseline distribution is Poisson, the size of the proposed robust test tends to be conservative but stable as the sample size increases. The power of the test increases as the sample size increases. In both situations where the true mixing weight is covariate dependent or constant, the test performs well. The proposed Wald test is robust to any linear or non-linear mean function used in the data generating process.

4.5 Real Data Application

The proposed Wald test is applied to longitudinal data from Detroit Dental Health Project (DDHP) which focused on the oral health of low-income African-American children and their caregivers (Sohn et al., 2007). For this study, the eligible participants were 1021

Table 4.3: *Empirical size and power of the robust Wald test statistics at the nominal level 0.05, based on 1,000 samples of correlated zero-inflated Negative Binomial data with mean* λ^*.

		n				
$\log(\lambda^*)$	ω^*	50	100	200	500	1000
1.5	$\omega^* = 0$	0.065	0.058	0.052	0.048	0.051
	$\omega^* = 0.1$	0.680	0.822	0.874	0.910	0.964
	$\omega^* = 0.2$	0.774	0.858	0.890	0.950	0.983
	$\omega^* = 0.1 - 0.2X_1$	0.451	0.572	0.820	0.970	0.999
$1.5 + 0.35X_1$	$\omega^* = 0$	0.047	0.049	0.051	0.054	0.048
	$\omega^* = 0.1$	0.670	0.850	0.890	0.930	0.970
	$\omega^* = 0.2$	0.760	0.870	0.920	0.970	0.990
	$\omega^* = 0.1 - 0.2X_1$	0.413	0.630	0.872	0.975	0.993
$1.5 + 0.35X_1 - 0.1X_2$	$\omega^* = 0$	0.058	0.060	0.056	0.057	0.055
	$\omega^* = 0.1$	0.770	0.881	0.960	0.981	0.997
	$\omega^* = 0.2$	0.883	0.962	0.978	0.989	0.998
	$\omega^* = 0.1 - 0.2X_1$	0.530	0.593	0.820	0.980	0.996
$0.3 + 0.5e^{-0.2X_2}$	$\omega^* = 0$	0.060	0.057	0.056	0.048	0.051
	$\omega^* = 0.1$	0.671	0.802	0.880	0.974	0.986
	$\omega^* = 0.2$	0.680	0.820	0.951	0.979	0.998
	$\omega^* = 0.1 - 0.2X_1$	0.552	0.665	0.783	0.960	0.991

$X_1 \sim U(0,1)$ and $X_2 \sim Bin(n, 0.3)$

families with a child under age of 6 and his/her main caregiver. Dental examinations were conducted and the participants were followed up in three waves in 2002-2003 (Wave1), 2004-2005 (Wave2) and 2007 (Wave3). The data consist of oral health related characteristics and general health related information of the participants. Some participants have not completed the follow-up examinations and it was identified that there was no systematic difference between the participants who completed the follow-up examinations and who dropped out. For our analysis, data collected on 1021 children are considered. As the outcome variable, we consider the DS score which represents the number of decayed tooth surfaces of each child. Dental caries scores collected across children are independent but the scores collected on the same child across the waves are correlated. Hence, the data are longitudinal in nature. As

Table 4.4: *Empirical size and power of the robust Wald test statistics at the nominal level 0.05, based on 1,000 samples of correlated zero-inflated Poisson data with mean* λ^*.

		n				
$\log(\lambda^*)$	ω^*	50	100	200	500	1000
0.3	$\omega^* = 0$	0.035	0.038	0.042	0.036	0.041
	$\omega^* = 0.1$	0.730	0.871	0.910	0.953	0.974
	$\omega^* = 0.2$	0.762	0.890	0.962	0.980	0.990
	$\omega^* = 0.1 - 0.2X_1$	0.571	0.734	0.890	0.972	0.990
$0.3 + 0.5X_1$	$\omega^* = 0$	0.023	0.028	0.034	0.027	0.053
	$\omega^* = 0.1$	0.740	0.830	0.920	0.960	0.980
	$\omega^* = 0.2$	0.860	0.950	0.970	0.981	0.990
	$\omega^* = 0.1 - 0.2X_1$	0.530	0.831	0.913	0.980	0.991
$0.3 + 0.5X_1 - 0.1X_2$	$\omega^* = 0$	0.038	0.035	0.048	0.052	0.049
	$\omega^* = 0.1$	0.652	0.892	0.960	0.985	0.992
	$\omega^* = 0.2$	0.861	0.940	0.989	0.996	0.998
	$\omega^* = 0.1 - 0.2X_1$	0.590	0.622	0.910	0.960	0.991
$0.3 + 0.5e^{-0.2X_2}$	$\omega^* = 0$	0.040	0.038	0.043	0.039	0.042
	$\omega^* = 0.1$	0.833	0.951	0.987	0.992	0.994
	$\omega^* = 0.2$	0.930	0.968	0.989	0.995	0.998
	$\omega^* = 0.1 - 0.2X_1$	0.370	0.532	0.811	0.982	0.996

$X_1 \sim U(0,1)$ and $X_2 \sim Bin(n, 0.3)$

covariates, we consider the child's age at the baseline year (Age), child's sugar intake (SI) and indicator variables for the follow-up examinations.

The proposed Wald test does not require the mean of the baseline model as a function of covariates and the test can be conducted without specifying the baseline distribution as Poisson or Negative Binomial. As the true baseline distribution is unknown in practice, for comparison, we also conduct the classical homogeneity Wald test under the Zero-Inflated Negative Binomial (ZINB) model with constant mixing weight and using the sandwich estimator of the variance-covariance matrix. The baseline mean λ_{ij} is defined as $\log(\lambda_{ij}) = x_{ij}^{'}\beta$ where x_{ij} is a vector of covariates including Age, SI, Age* SI, Wave2, Wave3, Age*Wave2, Age*Wave3, where Wave2 and Wave3 are indicator variables for the Waves and β is a vector of regression parameters. The overdispersion parameter under the Negative Binomial distiribution is κ. The results of the proposed Wald test and the classical Wald test under ZINB model are given in Table 4.5. Parameter estimates of the model with corresponding standard errors are given in Table 4.6. The overdispersion parameter κ under the ZINB model is significant ($\hat{\kappa}$=1.3989, p-value<0.0001).

The proposed Wald test statistic rejects the null hypothesis at 5% significance level supporting the hypothesis of heterogeneity. The classical Wald test under ZINB model fails to reject the null hypothesis at 5% significance level. To evaluate the results further we conduct a stratification analysis by waves. As indicated by Figure 4.1, there is no strong evidence of inflation at zero in Wave 1, Wave 2 and Wave 3. As a result, the homogeneity test conducted under the Zero-Inflated Negative Binomial model has not been able to detect the heterogeneity under the assumption of constant mixing weight while the proposed robust Wald test has been able to detect the slight heterogeneity. The result from the proposed test is consistent with the findings from Hsu et al. (2014) who used the same dataset and applied a Wald test for heterogeneity assuming the mixing weight depend on covariates and supports for the existence of heterogeneity in the population. Even with the constant mixing weight in the zero-inflated model, the proposed robust Wald test has been able to identify the heterogeneity in the population.

Table 4.5: *Comparison of Wald test statistics and associated p-values of the homogeneity tests for longitudinal dental caries data.*

	Test Statistic	p-value
Robust Wald test	390	<0.0001
Classical Wald test under ZINB with $\log(\lambda_{ij})$ (assuming constant ω)	2.37	0.1238

Table 4.6: *Fitted constant mixing weight models for longitudinal dental caries data.*

	ZINB		Working Independence model	
Parameters	Estimates	S.E.	Estimates	S.E.
β_0	-0.9242^*	0.2083	-	-
β_1	4.2680^*	0.3174	-	-
β_2	2.0567^*	0.6182	-	-
β_3	2.1609^*	0.1798	-	-
β_4	2.6758^*	0.1962	-	-
β_5	-3.0576^*	0.2856	-	-
β_6	-4.0864^*	0.3129	-	-
β_7	-2.6886^*	0.9852	-	-
ω	0.0475	0.0308	0.2644^*	0.0134
κ	1.3989^*	0.1335	-	-
α^{**}	-	-	1.1709^*	0.0789
β^{**}	-	-	6.7135^*	0.4650

κ indicates the overdispersion parameter under the ZINB model
α^{**} and β^{**} are the parameters under the working independence model
* indicates p-value < 0.05

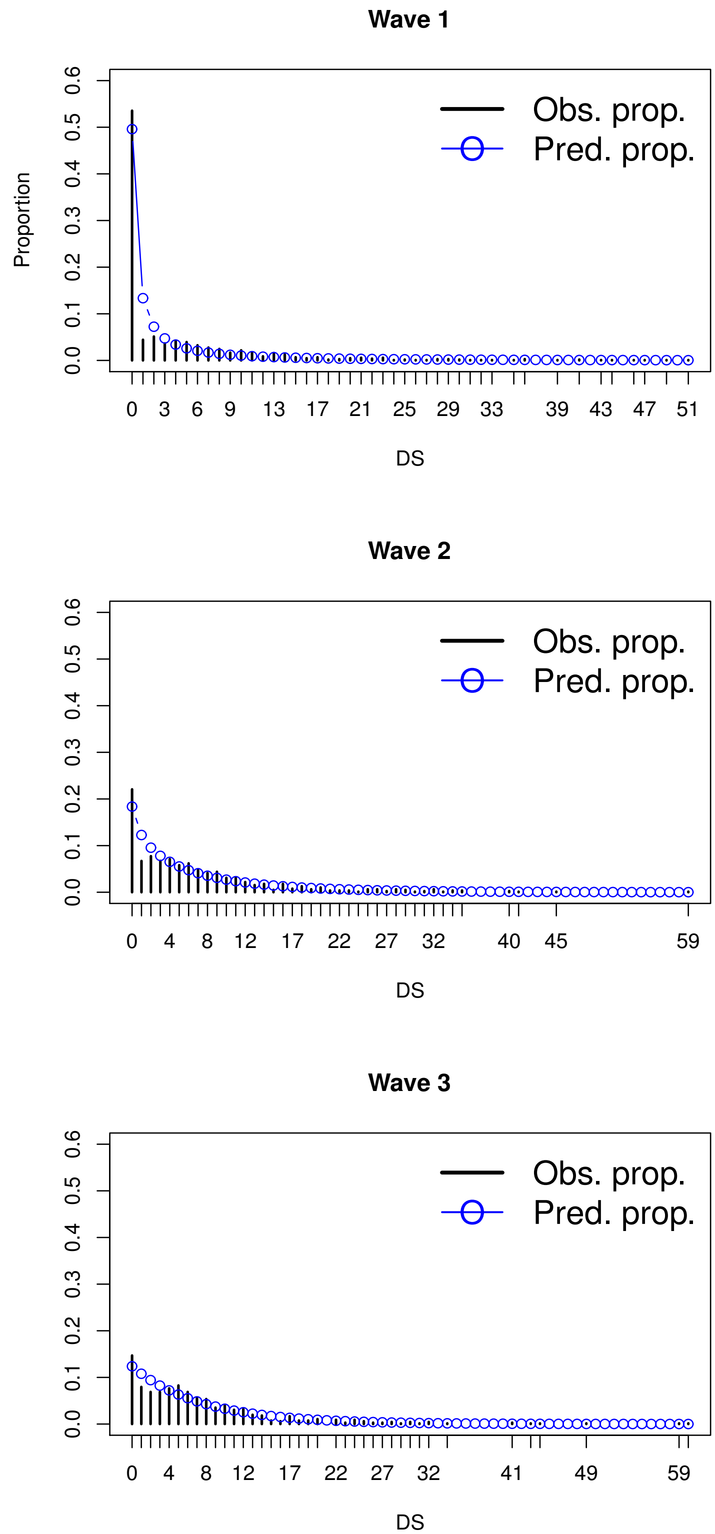

Figure 4.1: *Observed vs fitted proportions by waves under the NB model for Detroit dental caries data.*

Chapter 5

Discussion

5.1 Summary

This study has extended the literature by developing two homogeneity tests for evaluating heterogeneity under zero-inflated models. The classical testing procedures of homogeneity under this class of models require the correct specification of the baseline model, which may be subject to misspecification. In most of the existing testing procedures the conditional mean under the baseline model is assumed to be linearly related to covariates though the log link function. Such an assumption may not be reasonable in some situations where the linearity assumption may not correctly identify the true mean function. Depending on the nature of the response, the specified baseline distribution under the model could be also incorrect resulting unreliable statistical inferences. The proposed testing procedure address these issues of misspecification by developing a score test based on a more general parametric framework, which can incorporate various baseline distributions also without specifying the associated mean function. The proposed methodology is extended to incorporate correlated count data with excess zeros by developing a Wald test for testing heterogeneity under this class of models.

The proposed tests have some limitations as the mixing weight under the proposed methodology is assumed to be constant which may not be realistic for the heterogeneity that varies with the covariate profile. In other words, such an assumption may not be reasonable when the true mixing weights are related to covariates. Although the proposed methodology naturally embedded the testing for both zero inflation and deflation, it should be noted that the one-sided version of the test might be desirable for testing zero-inflation alone. If the alternative $\omega \neq 0$ is replaced by $\omega > 0$, the reference distribution of the score test statistic would not be a χ_1^2 distribution, instead, the limiting distribution would follow a mixture of a degenerate point mass at zero and a χ_1^2 distribution with equal mixing proportions.

In terms of the proposed Wald test statistic which relies on the working independence assumption in the estimation coupled with the sandwich estimator, the Generalized Estimating Equations (GEE) approach under a specified working correlation matrix would be an alternative approach to account for the correlation in data. Such approach, however, can be computationally demanding in practice.

From the practical standpoint, the proposed Wald test can be implemented in a commercial software such as SAS with minimal programming effort. The advantage of only having to estimate few parameters compared to the existing homogeneity tests would be useful in practice. The proposed Wald test would be a useful extension for the literature on homogeneity tests for correlated count data with excess zeros.

Given that the true underlying baseline distribution and the relation between the mean of the response and covariates is usually unknown in practice, the proposed general approach that does not specify the form of the relation *a priori* appears to be a more robust approach in practice. Extensions of the test when the true underlying distribution is Binomial or any other count distribution would be of interest. This extension and generalization of the methodology for the cases that the true baseline distribution is a member of the linear exponential family may be the subject of further research.

Bibliography

Bera, A. K. and Yoon, M. J. (1993). Specification testing with locally misspecified alternatives. *Econometric theory 9*(7), 649–658.

Böhning, D., Dietz, E. , Schlattmann, P. , Mendonca, L. , and Kirchner, U. (1999). The zero-inflated poisson model and the decayed, missing and filled teeth index in dental epidemiology. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 162*(12), 195–209.

Deng, D. and Paul, S. R. (2000). Score tests for zero inflation in generalized linear models. *Canadian Journal of Statistics 28*(3), 563–570.

DiRienzo, A. and Lagakos, S. (2001). Effects of model misspecification on tests of no randomized treatment effect arising from cox's proportional hazards model. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 63*(4), 745–757.

Farewell, V. T. and Sprott, D. (1988). The use of a mixture model in the analysis of count data. *Biometrics* (3), 1191–1194.

Freedman, D. A. (2006). On the so-called huber sandwich estimator and robust standard errors. *The American Statistician 60*(4), 299–302.

Godfrey, L. G. (1991). *Misspecification tests in econometrics: the Lagrange multiplier principle and other approaches.* Number 4. Cambridge University Press.

Hall, D. B. (2000). Zero-inflated poisson and binomial regression with random effects: a case study. *Biometrics 56*(13), 1030–1039.

Hsu, W.-W., Todem, D. , Kim, K. , and Sohn, W. (2014). A wald test for zero inflation and deflation for correlated count data from dental caries research. *Statistical Modelling 14*(6), 471–488.

Hu, M.-C., Pavlicova, M. , and Nunes, E. V. (2011). Zero-inflated and hurdle models of count data with extra zeros: examples from an hiv-risk reduction intervention trial. *The American journal of drug and alcohol abuse 37*(5), 367–375.

Huber, P. J. (1967). The behavior of maximum likelihood estimates under nonstandard conditions. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, Volume 1, pp. 221–233. Berkeley, CA.

Hur, K., Hedeker, D. , Henderson, W. , Khuri, S. , and Daley, J. (2002). Modeling clustered count data with excess zeros in health care outcomes research. *Health Services and Outcomes Research Methodology 3*(1), 5–20.

Jansakul, N. and Hinde, J. (2002). Score tests for zero-inflated poisson models. *Computational statistics & data analysis 40*(15), 75–96.

Jansakul, N. and Hinde, J. P. (2008). Score tests for extra-zero models in zero-inflated negative binomial models. *Communications in statistics-simulation and computation 38*(16), 92–108.

Kassahun, W., Neyens, T. , Faes, C. , Molenberghs, G. , and Verbeke, G. (2014). A zero-inflated overdispersed hierarchical poisson model. *Statistical Modelling 14*(5), 439–456.

Lambert, D. (1992). Zero-inflated poisson regression, with an application to defects in manufacturing. *Technometrics 34*(6), 1–14.

Lee, A. H., Wang, K. , Scott, J. A. , Yau, K. K. , and McLachlan, G. J. (2006). Multi-level zero-inflated poisson regression modelling of correlated count data with excess zeros. *Statistical methods in medical research 15*(1), 47–61.

Liang, K.-Y. and Self, S. G. (1996). On the asymptotic behaviour of the pseudolikelihood ratio test statistic. *Journal of the Royal Statistical Society. Series B (Methodological)* (9), 785–796.

Min, Y. and Agresti, A. (2005). Random effect models for repeated measures of zero-inflated count data. *Statistical Modelling 5*(1), 1–19.

Moulton, L. H., Curriero, F. C. , and Barroso, P. F. (2002). Mixture models for quantitative hiv rna data. *Statistical Methods in Medical Research 11*(4), 317–325.

Ridout, M., Demétrio, C. G. , and Hinde, J. (1998). Models for count data with many zeros. In *Proceedings of the XIXth international biometric conference*, Volume 19, pp. 179–192.

Rosenkranz, R. R., Behrens, T. K. , and Dzewaltowski, D. A. (2010). A group-randomized controlled trial for health promotion in girl scouts: healthier troops in a snap (scouting nutrition & activity program). *BMC public health 10*(1), 81.

Sohn, W., Ismail, A. , Amaya, A. , and Lepkowski, J. (2007). Determinants of dental care visits among low-income african-american children. *The Journal of the American Dental Association 138*(3), 309–318.

Tellez, M., Sohn, W. , Burt, B. A. , and Ismail, A. I. (2006). Assessment of the relationship between neighborhood characteristics and dental caries severity among low-income african-americans: a multilevel approach. *Journal of public health dentistry 66*(1), 30–36.

Todem, D., Hsu, W.-W. , and Kim, K. (2012). On the efficiency of score tests for homogeneity in two-component parametric models for discrete data. *Biometrics 68*(18), 975–982.

Van den Broek, J. (1995). A score test for zero inflation in a poisson distribution. *Biometrics* (8), 738–743.

Wang, K., Yau, K. K. , and Lee, A. H. (2002). A zero-inflated poisson mixed model to analyze diagnosis related groups with majority of same-day hospital stays. *Computer Methods and Programs in Biomedicine 68*(3), 195–203.

Xiang, L., Lee, A. H. , Yau, K. K. , and McLachlan, G. J. (2006). A score test for zero-inflation in correlated count data. *Statistics in medicine 25*(10), 1660–1671.

Yau, K. K. and Lee, A. H. (2001). Zero-inflated poisson regression with random effects to evaluate an occupational injury prevention programme. *Statistics in medicine 20*(19), 2907–2920.

Yau, K. K., Wang, K. , and Lee, A. H. (2003). Zero-inflated negative binomial mixed regression modeling of over-dispersed count data with extra zeros. *Biometrical Journal 45*(4), 437–452.

Yip, K. C. and Yau, K. K. (2005). On modeling claim frequency data in general insurance with extra zeros. *Insurance Mathematics and Economics 36*(2), 153–163.

Appendix A

Score function and second derivatives of log likelihood function for Robust score test

Based on the log-likelihood function $L = l(\boldsymbol{y}, \alpha, \beta, \omega)$,

$$\begin{aligned}
\frac{\partial L}{\partial \omega} &= \sum_{i=1}^{n} \left\{ \mathrm{I}_{(y_i=0)} \frac{1}{[\omega + (1-\omega)(1+\beta)^{-\alpha}]} [1 - (1+\beta)^{-\alpha}] - \mathrm{I}_{(y_i>0)} \frac{1}{1-\omega} \right\} \\
S_w &= \left(\frac{\partial L}{\partial \omega} \right)_{\omega=0} \\
&= \sum_{i=1}^{n} \left\{ \mathrm{I}_{(y_i=0)} \frac{[1 - (1+\beta)^{-\alpha}]}{(1+\beta)^{-\alpha}} - \mathrm{I}_{(y_i>0)} 1 \right\} \\
&= \sum_{i=1}^{n} \left\{ \frac{\mathrm{I}_{(y_i=0)}.1 - \mathrm{I}_{(y_i=0)}(1+\beta)^{-\alpha} - \mathrm{I}_{(y_i>0)}(1+\beta)^{-\alpha}}{(1+\beta)^{-\alpha}} \right\} \\
&= \sum_{i=1}^{n} \left\{ \frac{\mathrm{I}_{(y_i=0)} - (1+\beta)^{-\alpha}}{(1+\beta)^{-\alpha}} \right\} \\
&= \sum_{i=1}^{n} \left\{ \frac{\mathrm{I}_{(y_i=0)}}{(1+\beta)^{-\alpha}} - 1 \right\} \\
&= \sum_{i=1}^{n} \left\{ \mathrm{I}_{(y_i=0)} (1+\beta)^{\alpha} - 1 \right\}
\end{aligned}$$

$$
\begin{aligned}
\frac{\partial^2 L}{\partial \omega^2} &= \sum_{i=1}^{n}\left\{\mathrm{I}_{(y_i=0)}\frac{-[1-(1+\beta)^{-\alpha}]^2}{[\omega+(1-\omega)(1+\beta)^{-\alpha}]^2} - \mathrm{I}_{(y_i>0)}\frac{1}{(1-\omega)^2}\right\} \\
-E\left(\frac{\partial^2 L}{\partial \omega^2}\right) &= \sum_{i=1}^{n}\left\{\frac{[1-(1+\beta)^{-\alpha}]^2}{[\omega+(1-\omega)(1+\beta)^{-\alpha}]} + \frac{[1-(1+\beta)^{-\alpha}]}{(1-\omega)}\right\} \\
I_{\omega\omega} &= -E\left(\frac{\partial^2 L}{\partial \omega^2}\right)_{\omega=0} \\
&= n\left\{\frac{[1-(1+\beta)^{-\alpha}]^2}{(1+\beta)^{-\alpha}} + [1-(1+\beta)^{-\alpha}]\right\} \\
&= n\left\{[1-(1+\beta)^{-\alpha}]\left[\frac{[1-(1+\beta)^{-\alpha}]}{(1+\beta)^{-\alpha}} + 1\right]\right\} \\
&= n\left\{\frac{[1-(1+\beta)^{-\alpha}]}{(1+\beta)^{-\alpha}}\right\} \\
&= n\left\{\frac{1}{(1+\beta)^{-\alpha}} - 1\right\} \\
&= n\big\{(1+\beta)^{\alpha} - 1\big\}
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 L}{\partial \omega \partial \alpha} &= \sum_{i=1}^{n}\left\{\mathrm{I}_{(y_i=0)}\frac{(1+\beta)^{-\alpha}\log(1+\beta)}{[\omega+(1-\omega)(1+\beta)^{-\alpha}]} + \frac{[1-(1+\beta)^{-\alpha}](1-\omega)(1+\beta)^{-\alpha}\log(1+\beta)}{[\omega+(1-\omega)(1+\beta)^{-\alpha}]^2}\right\} \\
-E\left[\frac{\partial^2 L}{\partial \omega \partial \alpha}\right] &= \sum_{i=1}^{n}\left\{-(1+\beta)^{-\alpha}\log(1+\beta) - \frac{[1-(1+\beta)^{-\alpha}](1-\omega)(1+\beta)^{-\alpha}\log(1+\beta)}{[\omega+(1-\omega)(1+\beta)^{-\alpha}]}\right\} \\
I_{\omega\alpha} &= -E\left[\frac{\partial^2 L}{\partial \omega \partial \alpha}\right]_{\omega=0} \\
&= n\{-(1+\beta)^{-\alpha}\log(1+\beta) - [1-(1+\beta)^{-\alpha}]\log(1+\beta)\} \\
&= n\{-(1+\beta)^{-\alpha}\log(1+\beta) - \log(1+\beta) + (1+\beta)^{-\alpha}]\log(1+\beta)\} \\
&= -n \;\; \log(1+\beta)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 L}{\partial\omega\partial\beta} &= \sum_{i=1}^{n}\{\mathrm{I}_{(y_i=0)}\frac{\alpha(1+\beta)^{-\alpha-1}}{[\omega+(1-\omega)(1+\beta)^{-\alpha}]} + \frac{[1-(1+\beta)^{-\alpha}](1-\omega)\alpha(1+\beta)^{-\alpha-1})}{[\omega+(1-\omega)(1+\beta)^{-\alpha}]^2}\} \\
-E\left(\frac{\partial^2 L}{\partial\omega\partial\beta}\right) &= \sum_{i=1}^{n}\{-\alpha(1+\beta)^{-\alpha-1} - \frac{[1-(1+\beta)^{-\alpha}](1-\omega)\alpha(1+\beta)^{-\alpha-1})}{[\omega+(1-\omega)(1+\beta)^{-\alpha}]}\} \\
I_{\omega\beta} &= -E\left(\frac{\partial^2 L}{\partial\omega\partial\beta}\right)_{\omega=0} \\
&= n\{-\alpha(1+\beta)^{-\alpha-1} - \frac{[1-(1+\beta)^{-\alpha}]\alpha(1+\beta)^{-\alpha-1})}{(1+\beta)^{-\alpha}}\} \\
&= n\{-\alpha(1+\beta)^{-\alpha-1} - \alpha(1+\beta)^{-1} + \alpha(1+\beta)^{-\alpha-1})\} \\
&= \frac{-n\alpha}{1+\beta}
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 L}{\partial\alpha\partial\beta} &= \sum_{i=1}^{n}\Bigg\{\mathrm{I}_{(y_i=0)}\frac{(\omega-1)(1+\beta)^{-\alpha-1}[1-\alpha\log(1+\beta)]}{[\omega+(1-\omega)(1+\beta)^{-\alpha}]} \\
&\quad - \frac{(1-\omega)^2\alpha(1+\beta)^{-\alpha}\log(1+\beta)(1+\beta)^{-\alpha-1}}{[\omega+(1-\omega)(1+\beta)^{-\alpha}]^2} - \mathrm{I}_{(y_i>0)}\frac{1}{1+\beta}\Bigg\} \\
-E\left(\frac{\partial^2 L}{\partial\alpha\partial\beta}\right) &= \sum_{i=1}^{n}\Bigg\{\mathrm{I}_{(y_i=0)} - (\omega-1)(1+\beta)^{-\alpha-1}[1-\alpha\log(1+\beta)] \\
&\quad + \frac{(1-\omega)^2\alpha(1+\beta)^{-\alpha}\log(1+\beta)(1+\beta)^{-\alpha-1}}{[\omega+(1-\omega)(1+\beta)^{-\alpha}]} + (1-\omega)\frac{[1-(1+\beta)^{-\alpha}]}{1+\beta}\Bigg\} \\
I_{\alpha\beta} &= -E\left(\frac{\partial^2 L}{\partial\alpha\partial\beta}\right)_{\omega=0} \\
&= n\Bigg\{(1+\beta)^{-\alpha-1}[1-\alpha\log(1+\beta)] + \alpha\log(1+\beta)(1+\beta)^{-\alpha-1} \\
&\quad + [1-(1+\beta)^{-\alpha}](1+\beta)^{-1}\Bigg\} \\
&= n\big\{(1+\beta)^{-1}\big\} \\
&= \frac{n}{1+\beta}
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 L}{\partial \beta^2} &= \sum_{i=1}^{n}\Bigg\{\mathrm{I}_{(y_i=0)}\frac{(1-\omega)\alpha(\alpha+1)(1+\beta)^{-\alpha-2}}{[\omega+(1-\omega)(1+\beta)^{-\alpha}]} - \frac{(1-\omega)^2\alpha^2[(1+\beta)^{-\alpha-1}]^2}{[\omega+(1-\omega)(1+\beta)^{-\alpha}]^2} \\
&\quad + \mathrm{I}_{(y_i>0)}\left[\frac{-y_i}{\beta^2}+\frac{(y_i+\alpha)}{(1+\beta)^2}\right] \\
-E\left(\frac{\partial^2 L}{\partial \beta^2}\right) &= \sum_{i=1}^{n}\Bigg\{-(1-\omega)\alpha(\alpha+1)(1+\beta)^{-\alpha-2} + \frac{(1-\omega)^2\alpha^2[(1+\beta)^{-\alpha-1}]^2}{[\omega+(1-\omega)(1+\beta)^{-\alpha}]} \\
&\quad - (1-\omega)[1-(1+\beta)^{-\alpha}]\left[\frac{-y_i}{\beta^2}+\frac{(y_i+\alpha)}{(1+\beta)^2}\right]\Bigg\} \\
I_{\beta\beta} &= -E\left(\frac{\partial^2 L}{\partial \beta^2}\right)_{\omega=0} \\
&= \sum_{i=1}^{n}\Bigg\{-\alpha(\alpha+1)(1+\beta)^{-\alpha-2}+\alpha^2(1+\beta)^{-\alpha-2}-[1-(1+\beta)^{-\alpha}]E\left[\frac{-y_i}{\beta^2}+\frac{(y_i+\alpha)}{(1+\beta)^2}\right]\Bigg\} \\
&= \sum_{i=1}^{n}\Bigg\{-\alpha(1+\beta)^{-\alpha-2}-[1-(1+\beta)^{-\alpha}]E\left[\frac{-y_i}{\beta^2}+\frac{(y_i+\alpha)}{(1+\beta)^2}\right]\Bigg\}
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 L}{\partial \alpha^2} &= \sum_{i=1}^{n}\Bigg\{\frac{-(\omega-1)\log(1+\beta)^2(1+\beta)^{-\alpha}}{[\omega+(1-\omega)(1+\beta)^{-\alpha}]} - \frac{(1-\omega)^2[(1+\beta)^{-\alpha}\log(1+\beta)]^2}{[\omega+(1-\omega)(1+\beta)^{-\alpha}]^2} \\
&\quad + \mathrm{I}_{(y_i>0)}[trigamma(y_i+\alpha)-trigamma(\alpha)]\Bigg\} \\
-E\left(\frac{\partial^2 L}{\partial \alpha^2}\right) &= \sum_{i=1}^{n}\Bigg\{(\omega-1)\log(1+\beta)^2(1+\beta)^{-\alpha} + \frac{(1-\omega)^2[(1+\beta)^{-\alpha}\log(1+\beta)]^2}{[\omega+(1-\omega)(1+\beta)^{-\alpha}]} \\
&\quad - (1-\omega)[1-(1+\beta)^{-\alpha}][trigamma(y_i+\alpha)-trigamma(\alpha)]\Bigg\} \\
I_{\alpha\alpha} &= -E\left(\frac{\partial^2 L}{\partial \alpha^2}\right)_{\omega=0} \\
&= \sum_{i=1}^{n}[(1+\beta)^{-\alpha}-1][E(d^2lgamma(y_i+\alpha))-E(d^2lgamma(\alpha))],
\end{aligned}
$$

where $trigamma(y_i{+}\alpha) = d^2lgamma(y_i{+}\alpha) = \dfrac{\partial^2 \log\Gamma(y+\alpha)}{\partial\alpha^2}$ and $d^2lgamma(\alpha) = \dfrac{\partial^2 \log\Gamma(\alpha)}{\partial\alpha^2}$

Appendix B

Sample R Code: Robust score test

```
##########Robust score test###################
library(maxLik)

n=dim(y)[1]
g=matrix(1,n,1)
df1=dim(g)[2]

a_initial=1
b_initial=1
alpha=0.05

Likhood=function(param){
a_par=param[1]
b_par=param[2]
return(sum(lgamma(y+a_par)-lgamma(a_par)-lgamma(y+1)
+y*(log(b_par)-log(1+b_par))-(a_par*log(1+b_par))))
}
```

```
A=matrix(rbind(c(1,0),c(0,1)),2,2)
B=matrix(rbind(0,0),2,1)
mle_results=maxLik(Likhood,start=c(a_initial,b_initial),
constraints=list(ineqA=A,ineqB=B))
alpha_res=mle_results$estimate[1]
beta_res=mle_results$estimate[2]

sw=sum((((y==0)*1)*((1+beta_res)^alpha_res))-1)
Iww=n*(((1+beta_res)^alpha_res)-1)
Iwa=-n*log(1+beta_res)
Iwb=(-n*alpha_res)/(1+beta_res)
Iwab=rbind(Iwa,Iwb)
vc=vcov(mle_results)
v=Iww-(t(Iwab)%*%vc%*%Iwab)

#score test statistic
s_robust=sw^2/v
s_robust=c(s_robust)
p_value=pchisq(s_robust,df1,lower.tail = FALSE)

s_robust
p_value
########################################
```

B.1 Sample SAS Code: Robust Wald test

```
%MACRO modelFit(y);
```

```
ods output Contrasts=testStat;
proc nlmixed data=d2 empirical noad qpoints=1 MAXITER=1000;
parms alpha=0.1 beta=0.1 wi=0;
bounds alpha >0;
bounds beta >0;

/*Robust test log-likelihood function*/
if y=0 then
 ll=log(wi+(1-wi)*((1+beta)**(-alpha)))+b0 ;
 else ll= log(1-wi)+lgamma(y+alpha)-lgamma(alpha)-lgamma(y+1)+y*log(beta)
 -(y+alpha)*log(1+beta)+b0;
model y~general(ll);
random b0~normal(0,1) subject=SubjectID;
contrast "F-test" wi;
run;
ods output close;
%MEND modelFit;

data d2;
set dental_data;
SubjectID=subjectid;
%modelFit(y);
run;
```