



Published in final edited form as:

Stat Med. 2014 May 30; 33(12): 2103–2114. doi:10.1002/sim.6092.

A Score-type Test for Heterogeneity in Zero-inflated Models in a Stratified Population

Guanqun Cao^a, Wei-Wen Hsu^b, and David Todem^{c,*}

^aDepartment of Mathematics and Statistics, Auburn University, Auburn, AL 36849, USA

^bDepartment of Statistics, Kansas State University, Manhattan, KS 66506, USA

^cDepartment of Epidemiology and Biostatistics, Michigan State University, East Lansing, MI 48824, USA

Abstract

We propose a score-type statistic to evaluate heterogeneity in zero-inflated models for count data in a stratified population, where heterogeneity is defined as instances in which the zero counts are generated from two sources. Evaluating heterogeneity in this class of models has attracted considerable attention in the literature, but existing testing procedures have primarily relied on the constancy assumption under the alternative hypothesis. In this paper, we extend the literature by describing a score-type test to evaluate homogeneity against general alternatives that do not neglect the stratification information under the alternative hypothesis. The limiting null distribution of the proposed test statistic is a mixture of chi-squared distributions which can be well approximated by a simple parametric bootstrap procedure. Our numerical simulation studies show that the proposed test can greatly improve efficiency over tests of heterogeneity that ignore the stratification information. An empirical application to dental caries data in early childhood further shows the importance and practical utility of the methodology in using the stratification profile to detect heterogeneity in the population.

Keywords

Count data; Dental caries; Negative binomial model; One-sided alternatives; Score test; Poisson model; Zero inflation

1. Introduction

In many epidemiological and biomedical studies, the primary outcomes of interest take the form of counts which can be well represented by standard distributions such as the binomial, negative binomial and Poisson models. In practice, however, some count data exhibit a higher incidence of zeros which can not be accommodated by these standard distributions. For such data, it is a common practice to use zero-inflated models to characterize the population. These are two-component mixture models that combine a degenerate

distribution at zero with a parametric non-degenerate distribution \mathcal{F}_γ known up to a finite dimensional parameter $\gamma \in \Gamma$. Specifically, the mixture distribution is defined as follows,

$$P(y=y_i)=\begin{cases} \pi_i(\boldsymbol{\omega})+(1-\pi_i(\boldsymbol{\omega}))f(y_i;\boldsymbol{\gamma}), & y_i=0 \\ (1-\pi_i(\boldsymbol{\omega}))f(y_i;\boldsymbol{\gamma}), & y_i>0. \end{cases} \quad (1)$$

where y denotes the random count variable and y_i its observed version for the i th observation, $i = 1, \dots, n$, and $\pi_i(\boldsymbol{\omega})$, $0 \leq \pi_i(\boldsymbol{\omega}) \leq 1$, with $\boldsymbol{\omega}$ being an unknown parameter, represents the unknown mixing probability, $f(y_i; \boldsymbol{\gamma})$, with $\boldsymbol{\gamma}$ being a $\pi \times 1$ vector of unknown parameters, represents the probability mass function of y under the homogeneous distribution. In essence, the density $f(y_i; \boldsymbol{\gamma})$ is a positive smooth function that decays rapidly at infinity with some degree of uniformity. A good example of such functions is the exponential family density function, for which the associated mixture model has been extensively studied in statistical research [1, 2, 3, 4]. A non member of the exponential family such as the two-parameter negative binomial distribution also constitutes an example of this family [5, 6, 7, 8].

Under the mixture model in (1), zero counts are generated from two sources and the parameter $\pi_i(\boldsymbol{\omega})$ measures the extent of this heterogeneity in the population. In real life applications of this mixture model, one is interested in evaluating whether this heterogeneity is consistent with observed data. This question has been examined by many authors in a variety of settings [1, 9, 10, 11]. But existing methodologies often rely on restrictive formulations of the working mixing weight model. One general limitation is that the mixing weights are often assumed constant under the alternative hypothesis. Although important, such an approach may fail to reject homogeneity against alternatives that allow the mixing weight to depend on covariates. In recent papers, Jansakul and Hinde [10] and Todem *et al.* [11] have shown, using two-sided alternatives, that incorporating covariates into the mixing weights can greatly improve the test efficiency. By allowing potentially negative mixing weights under the alternative, the tests developed by these authors are only valid under the marginal representation of the mixture model which ignores its hierarchical representation. This is another limitation as zero-inflated models which maintain their hierarchical representation are usually fit in practice [12].

In this paper, we suggest an extension of existing homogeneity testing procedures to covariates with a focus on alternatives that are consistent with real applications of zero-inflated regression models. Specifically, we consider the situation where the mixing weight depends on a stratification variable with few strata under the alternative model. A complication, however, is that the implied hypotheses may not be typical and standard regularity conditions to conduct the test may not hold. There are hypothesized parameters under the null that lie on the boundary of the parameter space and one-sided composite hypotheses under the alternatives. We develop a score-type test of homogeneity that can accommodate these complications. Technically, the test statistic is similar in spirit to that of Silvapulle and Silvapulle [13] in detecting general alternatives and has the well known advantage of only requiring model estimation under the null hypothesis. Using numerical simulations and a real life example, the test statistic that can detect varying heterogeneity

under the alternative is found to be relatively more powerful than the test that assumes constant heterogeneity under alternative, when the underlying heterogeneity varies with the stratification profile.

The rest of this article is organized as follows. In Section 2, we develop a one-sided score test for homogeneity of zeros against alternatives from the mixture models with stratum dependent mixing probabilities. In Section 3, we conduct numerical studies to evaluate the finite sample properties of the proposed testing procedure and illustrate its practical utility using a real life example in early childhood dental caries research. Some remaining issues are discussed in Section 4.

2. A score test of homogeneity in a stratified population

Suppose we are interested in evaluating the hypothesis of zero mixing weights against the alternative that the sample $y_i, i = 1, \dots, n$, is obtained from model in (1). Under the null hypothesis $H_0: \pi_i(\omega) = 0$, the mixture model in (1) reduces to the homogeneous model with probability mass function $f(y_i; \gamma)$. A popular approach for assessing this hypothesis consists of assuming a constant mixing weight under alternatives to homogeneity [9]. Such an approach may suffer from reduced power if the underlying true mixing weight depends on covariates. In this paper, we construct a test statistic that allows the mixing probability $\pi_i(\omega)$ to depend upon a categorical covariate such as a stratification variable under the alternative. Suppose that the population under study consists of K known distinct strata and that each stratum k has its own mixing weight ω_k , with $0 < \omega_k < 1$. Let δ_{ik} be a non random binary variable which takes value 1 if subject i belongs to stratum $k, 1 \leq k \leq K$ and 0 otherwise, with $\sum_{k=1}^K \delta_{ik} = 1$. The mixing probability for subject i can be written as $\pi_i(\omega) = \prod_{k=1}^K \omega_k^{\delta_{ik}}$ with $\omega_k = \omega_k / (1 + \omega_k), \omega_k > 0$, and $\omega = (\omega_1, \omega_2, \dots, \omega_K)'$. In other words, ω_k represents the proportion of extra zeros in Stratum $k, k = 1, \dots, K$. To evaluate the homogeneity hypothesis, one is typically interested in the one-sided alternative,

$$H_0: \omega_k^* = 0 \text{ for all } k \quad \text{vs.} \quad H_1: \omega_k^* > 0 \text{ for some } k, \quad (2)$$

where ω_k^* represents the true value of $\omega_k, k = 1, 2, \dots, K$.

Let $\delta_i = \{\delta_{i1}, \dots, \delta_{iK}\}$ represent the stratification information for unit i and by x_i any potential covariates observed alongside y_i . We assume that observed data $\{y_i\}_{i=1}^n$ are random independent copies of y . Given data $\{y_i, \delta_i, x_i\}_{i=1}^n$ and setting $\theta = (\omega', \gamma)'$, we denote by $\ell_n(\theta)$ the log likelihood function associated with parameter vector θ and by $U_n(\theta) = \ell_n'(\theta) / \theta$ the corresponding first-order derivative with respect to parameter vector θ . Let $\theta^* = (\omega^{*'}, \gamma^{*'})'$ be the true value of θ with $E\{U_n(\theta^*)\} = 0$, to construct the score-based test statistic we assume the existence of a nonsingular matrix $G(\theta^*)$ and the following conditions:

- A1** All strata share a common subset of θ .
- A2** $n^{-1/2} U_n(\theta^*) \rightarrow^d N(0, G(\theta^*))$ as $n \rightarrow \infty$.

A3 For any $\alpha > 0$, $\sup_{\|\eta\| \leq \alpha} \|n^{-1/2}\{U_n(\theta^* + n^{-1/2}\eta) - U_n(\theta^*)\} + G(\theta^*)\eta\| = o_p(1)$, where $o_p(1)$ represents the convergence to 0 in probability as $n \rightarrow \infty$.

The discussion of these conditions and related technical details are provided in the Appendix. Let $\hat{\theta}_0 = (\mathbf{0}', \hat{\gamma}')'$, where $\hat{\gamma}$ is an estimator of γ^* obtained under the null hypothesis. Because of the one-sided constraints $\omega_k \geq 0, k = 1, \dots, K$, we consider the likelihood ratio test statistic for H_0 against H_1 based on the asymptotic normality of the ‘single’ observation $S_n(\theta_0) = n^{-1/2}U_n\omega(\hat{\theta}_0)$ to derive the general score statistic,

$$T_n = S_n(\hat{\theta}_0)' G^{\omega\omega}(\hat{\theta}_0) S_n(\hat{\theta}_0) - \inf_{\tilde{\omega} \geq 0} \left\{ (S_n(\hat{\theta}_0) - \tilde{\omega})' G^{\omega\omega}(\hat{\theta}_0) (S_n(\hat{\theta}_0) - \tilde{\omega}) \right\}, \quad (3)$$

where $\tilde{\omega} \geq 0$ with $\tilde{\omega} = (\tilde{\omega}_1, \tilde{\omega}_2, \dots, \tilde{\omega}_K)'$ is interpreted coordinatewise, that is $\tilde{\omega}_k \geq 0$, for all $k = 1, 2, \dots, K$. In practice, the quadratic function of $\tilde{\omega}$ can be easily minimized by using available optimization subroutines such as *optim* in the statistical software *R*. When $K = 1$, $\omega = \omega_1 \equiv \omega$ is a scalar and the test statistic simply reduces to $T_n =$

$S_n(\hat{\theta}_0)' G^{\omega\omega}(\hat{\theta}_0) S_n(\hat{\theta}_0) I\{S_n(\hat{\theta}_0) > 0\}$, where $I\{\cdot\}$ represents the indicator function. The null hypothesis is then rejected for large values of T_n . Statistic T_n has the limiting distribution

$\mathcal{T} = Z' G^{\omega\omega}(\theta_0^*) Z - \inf_{\tilde{\omega} \geq 0} \{ (Z - \tilde{\omega})' G^{\omega\omega}(\theta_0^*) (Z - \tilde{\omega}) \}$ with $Z \sim N(\mathbf{0}, G^{\omega\omega}(\theta_0^*)^{-1})$ for large values of n [14]. When $K = 1$, this limiting distribution is a mixture of chi-squared distributions χ_0^2 (a degenerate distribution with probability point mass at 0) and χ_1^2 , each with weight 0.5. For $K > 1$, the large sample distribution of the statistic T_n remains a mixture of chi-squared distributions but the mixing weights are quite complicated and are often approximated only through numerical methods [13].

When θ_0^* is known such as in simulation experiments for the study of the test empirical size, \mathcal{T} the limiting null distribution of T_n can be approximated by generating for example one million points from a centered multivariate normal distribution with covariance matrix $G^{\omega\omega}(\theta_0^*)^{-1}$. However, when θ_0^* is unknown such as in real data analysis and simulation studies for the study of the test empirical power, a simple parametric bootstrap procedure can be used to approximate the distribution of \mathcal{T} . The bootstrap resampling method was introduced by Efron [15] and has become a routine method for approximating distributions that are difficult to obtain analytically. Its validity follows automatically from arguments given in Efron and Tibshirani [16]. From a parametric point of view, the bootstrap methodology consists of three simple steps: i) an estimation step under which the parameters of the null model are estimated from observed data, ii) a Monte Carlo step to generate B pseudo-data sets from the fitted model and calculate the associated test statistics; and finally iii) construction of the bootstrap distribution for a sufficiently large value of B . We give below the details of this procedure to obtain the large sample distribution of the proposed score test T_n .

1. Estimation step: Given observed data $\{y_i, \delta_i, x_i\}_{i=1}^n$, compute the estimator $\hat{\gamma}$ of γ^* under the null model with probability mass function $f(y_i; \gamma)$.

2. Monte Carlo step: For each b , generate the Monte Carlo sample $\{y_i^{(b)}\}_{i=1}^n$ from the null model with $\boldsymbol{\gamma}$ fixed at $\hat{\boldsymbol{\gamma}}$ and assign the generated data point $y_i^{(b)}$ to δ_i and x_i . For each Monte Carlo sample $\{y_i^{(b)}, \delta_i, x_i\}, i = 1, \dots, n$, calculate the statistic

$$T_n^{(b)} = S_n(\hat{\boldsymbol{\theta}}_0^{(b)})' G^{\omega\omega}(\hat{\boldsymbol{\theta}}_0^{(b)}) S_n(\hat{\boldsymbol{\theta}}_0^{(b)}) - \inf_{\tilde{\omega} \geq 0} \left\{ \left(S_n(\hat{\boldsymbol{\theta}}_0^{(b)}) - \tilde{\omega} \right)' G^{\omega\omega}(\hat{\boldsymbol{\theta}}_0^{(b)}) \left(S_n(\hat{\boldsymbol{\theta}}_0^{(b)}) - \tilde{\omega} \right) \right\}.$$

where $\hat{\boldsymbol{\theta}}_0^{(b)} = (\mathbf{0}', \hat{\boldsymbol{\gamma}}^{(b)'})'$ with $\hat{\boldsymbol{\gamma}}^{(b)}$ being an estimate of $\hat{\boldsymbol{\gamma}}$ under the null hypothesis using the pseudo data $\{y_i^{(b)}, \delta_i, x_i\}, i = 1, \dots, n$.

3. Repeat Step 2 for $b = 1, 2, \dots, B$.

By the Glivenko-Cantelli Theorem, when B is large, an approximate p-value of the test is

$B^{-1} \sum_{b=1}^B I(T_n^{(b)} \geq t_n)$, the proportion of these artificial statistics which exceed t_n , the observed value of T_n . In our numerical applications, B is set to 1000.

3. Numerical Studies

We illustrate the use of the proposed score-type statistic to evaluate homogeneity against heterogeneity for the class of zero-inflated models for stratified discrete data. For the real data analysis, the zero-inflated negative binomial model will be used. The zero-inflated Poisson regression model will be utilized for the numerical experiment to reduce the computational burden. These two popular models verify all the smoothness and regularity conditions A2 and A3. Condition A1, however, will be discussed in regard to each working model specification. Derivations of the basic quantities required to construct the proposed test statistics for these models are relegated to the Appendix.

3.1. Dental Caries Data

Our motivating example comes from a longitudinal (three-wave) study designed to collect oral health information on inner-city low-income African American populations of Detroit, Michigan. Targeted participants were African American children under the age of 6 and their main caregivers of at least 14 years of age living in Detroit households with income below 250% of the 2000 federal poverty level. Study participants (index children and their caregivers) were randomly selected according to a two-stage probability sample design. At the first stage, census blocks were sampled with probabilities proportional to their size. Housing units were sampled at the second stage with probabilities inversely proportional to their size. From this sampling scheme, 1,021 families (children and their caregivers) were drawn and eventually examined. For selected participants, the oral examinations involved assessments of dental caries, periodontal disease, oral cancer and oral hygiene status. A set of questionnaires about demographics, social factors of oral health behaviors and practices including social support and depressive symptoms (measured by the Center of Epidemiologic Studies Scale CESD) among others, and biological characteristics was also administered. A food frequency questionnaire was administered to caregivers to collect

information (such as the amount of sugar intake) on their own dietary practices and those of the child.

In this paper, we focus on a mouth-level index which represents the overall experience of dental caries for each surveyed child. This index counts the number of decayed surfaces (ds) which are surfaces with signs of clinically detectable enamel lesions comprising both noncavitated and cavitated lesions. Although this type of mouth-level indices has well documented limitations regarding its ability to measure dental caries, it continues to be instrumental in evaluating and comparing the risks of dental caries across population groups. Most importantly, it remains popular in dental caries research because it enables one to conduct historical comparisons in population-based studies [17]. A more detailed description of the study can be found elsewhere [18, 19, 20].

Our numerical computations ignore aspects of the multi-stage sampling design and are based on cross-sectional data obtained in the first wave of examinations and interviews completed between 2002–2003. Covariates considered in our analysis include the child's age group (0 year Age 2 years; 2 years < Age 4 years; and Age > 4 years) and the child's sugar intake or consumption (SI) measured in grams per day, variables that have been shown to be associated with dental caries [21, 22]. The children's age group serves as the stratification variable in this paper (0 year Age 2 years for Stratum 1; 2 years < Age 4 years for Stratum 2; and Age > 4 years for Stratum 3). As a preliminary analysis, bar chart plots of relative frequencies for the overall sample and for each age-specific stratum are constructed and presented in Figure 1. From these plots, a sizeable proportion (about 50%) of children in the sample do not have dental caries. More importantly, zeros frequencies are more prominent in the youngest age group, with about 80% of these children having no tooth surfaces with signs of clinically detectable enamel lesions. This result is consistent with the fact that younger children do not have all their teeth erupted and therefore are less likely to have dental caries.

Following earlier analyses of dental caries indices in these African American children, a negative binomial model is used to accommodate overdispersion due to sizeable frequencies of children with large caries counts [11]. Specifically, we consider a negative binomial regression model with mean $\lambda_i = \exp \left\{ \sum_{k=1}^3 \delta_{ik} (\beta_{0k} + \beta_{1k} \times SI_i) \right\}$ and an overdispersion parameter $\kappa > 0$. All parameters of the mean λ_i are stratum-specific but the overdispersion parameter κ is common across all strata to ensure that condition A1 is satisfied. This condition will be evaluated against data in a subsequent analysis. For this homogeneous model, parameter vector $\boldsymbol{\gamma} = (\boldsymbol{\beta}', \kappa)'$ with $\boldsymbol{\beta} = (\beta_{01}, \beta_{11}, \dots, \beta_{03}, \beta_{13})'$, and variable δ_{ik} represents the binary indicator for Stratum k , $k = 1, 2, 3$. This basic starting model is then evaluated against the associated zero-inflated model with a constant mixing probability $\pi_i(\omega) = \omega / (1 + \omega) \equiv \omega_1 = \omega_2 = \omega_3$, $\omega \geq 0$. For the caries index considered, the homogeneity hypothesis $\omega^* = 0$ (no need for zero-inflation) fails to be rejected at 5% significance level against the zero-inflated model that neglects the mixing weight stratification (observed statistic = 0.158 with a p-value of 0.346 or 0.401, using respectively a mixture of χ_0^2 and χ_1^2 distributions, each with weight 0.5 and the proposed parametric bootstrap resampling scheme).

We then consider the proposed test that accommodates the stratification of mixing weights under the alternative hypothesis. Specifically, we assume a working zero-inflated negative binomial model with mixing probability $\pi_i(\boldsymbol{\omega}) = \prod_{k=1}^3 \bar{\omega}_k^{\delta_{ik}}$, where $\bar{\omega}_k = \omega_k / (1 + \omega_k)$, ω_k , 0 , represents the proportion of extra zeros in Stratum k , $k = 1, 2, 3$, and $\boldsymbol{\omega} = (\omega_1, \omega_2, \omega_3)'$. Under this working model, the homogeneity hypothesis is clearly rejected at 10^{-3} significance level (observed statistic $t_n = 101.304$, p-value < 0.001 obtained from a parametric bootstrap resampling scheme), therefore providing a strong evidence of heterogeneity in this population. Table 1 gives the parameter estimates and their standard errors, and summary statistics of the homogeneous (null) model and the heterogeneous models with a constant mixing weight and stratum-specific mixing weights. The AIC and BIC statistics provide another evidence that the zero-inflated negative binomial model coupled with stratum-specific mixing weights provides a better fit to these data compared to the homogeneous model and the zero-inflated negative binomial model with a constant mixing weight. Results from fitting the stratum-specific mixing weight model indicate that there is no inflation of zeros in Stratum 3 representing the oldest age group ($\bar{\omega}_3 = 0.0478$, $se(\bar{\omega}_3) = 0.0358$) in contrast to the two younger age groups ($\bar{\omega}_1 = 0.7667$, $se(\bar{\omega}_1) = 0.0243$, and $\bar{\omega}_2 = 0.2119$, $se(\bar{\omega}_2) = 0.0329$). Hence by averaging over the strata information, the test statistic based on the constant mixing weight approach loses power in detecting heterogeneity in these data. This example clearly illustrates the importance of formulating less restrictive alternative hypotheses when evaluating a null model. Given that the true underlying parametric model that provides the best representation of the data is often unknown, entertaining less restrictive formulations appears to be the most conservative strategy.

As a final analysis, the non-separability condition A1 was examined by evaluating whether the overdispersion parameter κ was constant across all strata in the zero-inflated negative binomial model with stratum-specific mixing weights. A two degrees of freedom Wald test performed to this effect yielded an observed value of 0.88 (p-value=0.6461), suggesting no evidence from data to reject the constancy of the overdispersion parameter κ over the strata. This result was also consistent with the likelihood ratio test (observed statistic=1.30 and p-value=0.4780) and other summary statistics such as the AIC and the BIC (see Table 2). We noticed however that even when the overdispersion parameter is allowed to depend on the stratification variable, the constant mixing weight test remains insensitive in detecting the varying heterogeneity of the mixing weight under the alternative, in contrast to the proposed test.

3.2. Simulations

In this section, we conduct a numerical study to evaluate the finite sample performances of the proposed score test statistic T_n which are in turn compared to those of the constant mixing weight score test statistic. Similarly to the real data example, we consider a stratified population made of three strata. But to keep the simulation scheme simple, we restrict the true non-degenerate model \mathcal{F}_r to a Poisson process with covariate-dependent mean

$\lambda_i^* = \exp\left\{\sum_{k=1}^3 \gamma_k^* \delta_{ik} + \gamma_4^* v_i\right\}$, where covariate v_i is generated from a uniform random variable on the interval $[0, 1]$ and the membership indicators δ_{i1} , δ_{i2} and δ_{i3} are generated

from a multinomial distribution with membership probabilities $p_1 = 0.3, p_2 = 0.3, p_3 = 0.4$, respectively. The stratum-specific intercept term, represented by the vector $(\gamma_1^*, \gamma_2^*, \gamma_3^*)$, takes values in the set $\{(-0.05, -0.10, -0.15), (0.05, 0.10, 0.15)\}$ and the slope γ_4^* associated with v_i is set to 0.5. In this simulation study, these intercept terms are used to capture the separation between the mean of the true homogeneous model \mathcal{F}_{γ^*} and 0. Indeed, large values of $\gamma_k^*, k = 1, 2, 3$, shift the mean of the true non-degenerate model further away from 0, therefore making the two components of the mixture model easily detectable. For data generated under the associated zero-inflated model, the true mixing weight is

$$\pi_i^*(\omega^*) = \prod_{k=1}^K \varpi_k^{*\delta_{ik}}, \text{ with } \varpi_k^* = \omega_k^*/(1 + \omega_k^*) \text{ being the true mixing weight in stratum } k, k = 1, 2, 3 \text{ and } \omega^* = (\omega_1^*, \omega_2^*, \omega_3^*)'.$$

Throughout the simulations, we assume the working mixing weight model

$$\pi_i(\omega) = \prod_{k=1}^K \varpi_k^{\delta_{ik}} \text{ with } \varpi_k = \omega_k/(1 + \omega_k) \text{ under the alternative hypothesis. Conversely under the constant mixing weight alternative, the working mixing weight model is } \pi_i(\omega) = \omega/(1 + \omega). \text{ The maximum likelihood estimate } \hat{\gamma}^* \text{ of } \gamma^* = (\gamma_1^*, \gamma_2^*, \gamma_3^*, \gamma_4^*) \text{ under the null are obtained from the homogeneous Poisson model with the working mean model}$$

$$\lambda_i = \exp\left\{\sum_{k=1}^3 \gamma_k \delta_{ik} + \gamma_4 v_i\right\}. \text{ Note here that the parameter } \gamma_4 \text{ is constant across all strata in this working model to ensure that condition A1 is satisfied. Finally, all simulations were replicated 1,000 times and for sample sizes } n \text{ from 50 to 400.}$$

To investigate the empirical type I error rate of the proposed tests, we generate data from the homogeneous Poisson model ($\omega_k^* = 0, k = 1, 2, 3$) with the true mean λ_i^* specified above. In Figure 2, as expected, the constant mixing weight and the proposed test statistics maintain their sizes at 5% nominal level regardless of the degree of separation between the homogeneous model and the degenerate distribution. We further investigate the empirical power of the tests when the true mixing weight is constant across the three strata by setting $\omega^* = \omega_c^*$, where ω_c^* takes values from the set $\{(0.1, 0.1, 0.1), (0.2, 0.2, 0.2)\}$. Overall, larger sample sizes and increasing values of $\omega_k^*, k = 1, 2, 3$, improve the power of detecting the alternatives under consideration for the two tests. Moreover, the power also increases with increasing values of the intercept γ_k^* . The constant mixing weight test appears to perform slightly better than the stratum-adjusted test when the true mixing weight is constant. But this loss of power of the stratum-adjusted test appears to be minimal.

We finally conduct a study to evaluate the performances of T_n when the constancy assumption under heterogeneity is violated. Figure 2 also presents the results of this study for $\omega^* = \omega_{ss}^*$, where ω_{ss}^* takes values from the set $\{(0, 0.1, 0.25), (0, 0.1, 0.5)\}$. As expected, larger sample sizes and increasing values of one component of ω^* while maintaining the remaining components fixed, improve the power of detecting the alternatives under consideration for all tests. Moreover, the power also increases when the two components of the mixture distribution are well separated, occurring when $\gamma_k^*, k = 1, 2, 3$, are large. The stratum-adjusted test clearly outperforms the constant mixing weight test. And the maximum gain of efficiency of the stratum-adjusted test is achieved when one stratum-specific mixing

probability is large while the remaining ones are negligible. Thus by averaging over the stratification variable, the stratum-unadjusted test may have weaker power to capture heterogeneity in the population. Findings from the dental caries data analysis are very consistent to this simulation conjecture. Indeed, the estimates from the alternative model show heterogeneous mixing weights across the three age groups, with the youngest children exhibiting a significantly higher mixing weight. This heterogeneity of mixing weights across the three age groups adversely affects power of the test statistic based on the constant mixing weight in detecting zero-inflation. Our simulations and real data analysis are reassuring and lend credence to the need to formulate more elaborate zero-inflated parametric models under alternative in evaluating zero-inflation.

4. Discussion

Whenever the inference for the mixing weight in zero-inflated regression models is of interest, existing testing methodologies have relied essentially on restrictive assumptions under heterogeneity. Typically, score tests that restrict the mixing weights to be constant under alternatives to homogeneity are often advocated. Our real life example and simulation studies have shown that testing procedures that rely on such restrictions may adversely affect statistical power in detecting heterogeneity. Using the general theory from Silvapulle and Silvapulle [13], we have extended the literature by developing a testing procedure to evaluate homogeneity in two-component mixture models, which explicitly takes into account the presence of stratification under heterogeneity. This procedure is shown to have satisfactory statistical power even when the true mixing weight is constant under heterogeneity.

One may argue that the test of zero inflation that assumes constancy of the mixing probability under the alternative can simply be performed within each stratum. This approach is in principle feasible but not necessarily optimal. Indeed, if there are parameters that are shared by some strata, a joint estimation requiring information across all these strata is unavoidable. In the dental caries example, the working regression model for the mean of the non-degenerate distribution is stratum specific but the three strata share the same overdispersion parameter κ . Likewise in our numerical studies, although the intercept of the working mean model of the non-degenerate distribution is stratum specific, the slope of the covariate v_i is stratum-independent and therefore requires information across all strata for its estimation.

To perform the test in practice, a parametric bootstrap is proposed to approximate the limiting distribution of the test statistic T_n under the null hypothesis. This parametric bootstrap requires the null model to be fitted for each simulated Monte Carlo sample, which can be computationally very cumbersome in some practical settings. Alternatively, a more efficient resampling procedure that perturbs the influence function using normal random variates to approximate the true distribution of the test statistics T_n can be used. This technique requires the null model to be fitted only once and has been extensively used in the literature when the desired limiting distribution is difficult or even impossible to derive analytically [23, 24, 25, 26].

The proposed test can be extended to refine the model specification. For example when the null hypothesis is rejected, it may be of interest to identify subgroups of the population that have nonzero mixing probabilities. We did not discuss how to identify these components, but our results can be used for this purpose. Inference regarding significance of each stratum mixing weight can be conducted taking into account that multiple tests are performed, which necessitates an adjustment of the type I error. Our testing procedure also assumes that the mixing weight under heterogeneity depends only on the stratification variable, which is essentially a categorical variable. In practice however, a two-component mixing weight model with a more general specification of the mixing weight is often considered. A good and popular example is the situation where the mixing weight under heterogeneity is also related to some continuous covariates in addition to the stratification variable. For such working mixing weight models, it is unclear how to construct a test of homogeneity against heterogeneity. These issues and other generalizations of the methodology are out of the scope of this work and are the subject of further research.

Acknowledgments

Contract/grant sponsor:

The authors are grateful to Amid Ismail and Woosung Sohn for their permission to use the dental caries data. They also thank the two anonymous referees for carefully reading the manuscript and making helpful suggestions for improving the presentation. This work was supported by Dr Todem's NCI/NIH K-award, 1K01 CA131259 and its supplement from the 2009 ARRA funding mechanism.

References

1. Deng D, Paul S. Score tests for zero inflation in generalized linear models. *The Canadian Journal of Statistics / La Revue Canadienne de Statistique*. 2000; 28(3):563–570.10.2307/3315965
2. Deng D, Paul S. Score tests for zero-inflation and over-dispersion in generalized linear models. *Statistica Sinica*. 2005; 15:257–276.10.1002/sim.2616
3. Hall D. Zero-inflated Poisson and binomial regression with random effects: A case study. *Biometrics*. 2000; 56(4):1030–1039.10.1111/j.0006-341X.2000.01030.x [PubMed: 11129458]
4. Lee A, Wang K, Scott J, Yau K, McLachlan G. Multi-level zero-inflated poisson regression modelling of correlated count data with excess zeros. *Statistical Methods in Medical Research*. 2006; 15(1):47–61.10.1191/0962280206sm429oa [PubMed: 16477948]
5. Ridout M, Hinde J, Demétrio CGB. A score test for testing a zero-inflated Poisson regression model against zero-inflated negative binomial alternatives. *Biometrics*. 2001; 57(1):219–223.10.1111/j.0006-341X.2001.00219.x [PubMed: 11252601]
6. Yau K, Wang K, Lee A. Zero-inflated negative binomial mixed regression modeling of over-dispersed count data with extra zeros. *Biometrical Journal*. 2003; 45(4):437–452.10.1002/bimj.200390024
7. Mwalili S, Lesaffre E, Declerck D. The zero-inflated negative binomial regression model with correction for misclassification: an example in caries research. *Statistical Methods in Medical Research*. 2008; 17(2):123–139.10.1177/0962280206071840 [PubMed: 17698937]
8. Agresti, A. *Analysis of ordinal categorical data*. Wiley; Hoboken, New Jersey: 2010.
9. van den Broek J. A score test for zero inflation in a Poisson distribution. *Biometrics*. 1995; 51:738–743. [PubMed: 7662855]
10. Jansakul N, Hinde J. Score tests for extra-zero models in zero-inflated negative binomial models. *Communications in Statistics-Simulation and Computation*. 2009; 38(1):92–108.10.1080/03610910802421632

11. Todem D, Hsu WW, Kim K. On the efficiency of score tests for homogeneity in two-component parametric models for discrete data. *Biometrics*. 2012; 68 :975–982.10.1111/j.1541-0420.2011.01737.x [PubMed: 22348298]
12. Lambert D. Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics*. 1992; 34:1–14.10.2307/1269547
13. Silvapulle M, Silvapulle P. A score test against one-sided alternatives. *Journal of the American Statistical Association*. 1995; 90:342–349.10.1080/01621459.1995.10476518
14. Shapiro A. Towards a unified theory of inequality constrained testing in multivariate analysis. *International Statistical Review*. 1988; 56:46–62.10.1234/12345678
15. Efron B. Bootstrap methods: another look at the jackknife. *The annals of Statistics*. 1979; 7(1):1–26.10.1214/aos/1176344552
16. Efron, B.; Tibshirani, R. *An Introduction to the Bootstrap*. Chapman & Hall Ltd; 1993.
17. Lewsey J, Thomson W. The utility of the zero-inflated poisson and zero-inflated negative binomial models: a case study of cross-sectional and longitudinal dmf data examining the effect of socio-economic status. *Community Dentistry and Oral Epidemiology*. 2004; 32(3):183–189. [PubMed: 15151688]
18. Tellez M, Sohn W, Burt B, Ismail A. Assessment of the relationship between neighborhood characteristics and dental caries severity among low-income african-americans: A multilevel approach. *Journal of Public Health Dentistry*. 2006; 66:30–36.10.1111/j.1752-7325.2006.tb02548.x [PubMed: 16570748]
19. Sohn W, Ismail A, Amaya A, Lepkowski J. Determinants of dental care visits among low-income african-american children. *The Journal of the American Dental Association*. 2007; 138(3):309–318. [PubMed: 17332036]
20. Ismail AI, Lim S, Sohn W. A transition scoring system of caries increment with adjustment of reversals in longitudinal study: evaluation using primary tooth surface data. *Community Dentistry and Oral Epidemiology*. 2011; 39(1):61–68.10.1111/j.1600-0528.2010.00565.x [PubMed: 20690932]
21. Woodward M, Walker A. Sugar consumption and dental caries: evidence from 90 countries. *British Dental Journal*. 1994; 176(8):297–302.10.1038/sj.bdj.4808437 [PubMed: 8186040]
22. Warren J, Weber-Gasparoni K, Marshall T, Drake D, Dehkordi-Vakil F, Dawson D, Tharp K. A longitudinal study of dental caries risk among very young low ses children. *Community dentistry and oral epidemiology*. 2009; 37(2):116–122.10.1111/j.1600-0528.2008.00447.x [PubMed: 19046332]
23. Parzen M, Wei L, Ying Z. A resampling method based on pivotal estimating functions. *Journal of the American Statistical Association*. 1994; 81:341–350.10.1093/biomet/81.2.341
24. Lin DY, Fleming TR, Wei LJ. Confidence bands for survival curves under the proportional hazards model. *Biometrika*. 1994; 81:73–81. doi: 10.1.1.31.2736.
25. Hansen BE. Inference when a nuisance parameter is not identified under the null hypothesis (STMA V37 4551). *Econometrica*. 1996; 64:413–430.
26. Zhu H, Zhang H. Generalized score test of homogeneity for mixed effects models. *The Annals of Statistics*. 2006; 34(3):1545–1569.10.1214/009053606000000380

Appendix

Appendix 1. The technical details of regularity conditions

Condition A1 is the non-separability criterion that requires all strata to share a common subset of θ . It prevents the test of zero-inflation that assumes constancy of the mixing probability under the alternative to be simply performed within each stratum. Although our methodology remains valid even when the distributions of data across strata do not have some parameters in common, its efficiency however can be compromised. Conditions A2

and A3 essentially ensure that the asymptotic normality, the smoothness and regularity requirements of the score function are met.

We partition U_n as $U_n = (U'_{n\omega}, U'_{n\gamma})'$ to conform with the partition (ω', γ') of θ and we denote by $G_{\omega\omega}(\theta)$, $G_{\omega\gamma}(\theta)$, and $G_{\gamma\gamma}(\theta)$ the corresponding blocks in $G(\theta)$. The estimator $\hat{\gamma}$ is consistent for γ^* only when the null hypothesis is true. However, when the null is not true, $\hat{\gamma}$ converges to γ^* , which is not necessarily γ^* . In other words, $\hat{\theta}_0 = (\theta', \hat{\gamma}')$ converges to $\theta_0^* = (\theta', \gamma^{*'})'$ with θ^* -probability 1, where θ_0^* satisfies the null hypothesis. Conditions A2 and A3 ensure that $n^{-1/2}U_{n\omega}(\hat{\theta}_0)$ under H_0 converges in distribution to a centered normal model with asymptotic variance-covariance matrix $G^{\omega\omega}(\theta_0^*)^{-1} = G_{\omega\omega}(\theta_0^*) = G_{\omega\gamma}(\theta_0^*)G_{\gamma\gamma}^{-1}(\theta_0^*)G_{\gamma\omega}(\theta_0^*)$.

Appendix 2. Detailed calculations for the zero-inflated negative model

Consider the zero-inflated negative binomial (ZINB) model for which the probability mass function of the homogeneous distribution is,

$$f(y_i; \gamma) = \frac{\Gamma(\kappa^{-1} + y_i)}{\Gamma(\kappa^{-1})y_i!} \left(\frac{\kappa\lambda_i}{1 + \kappa\lambda_i} \right)^{y_i} \left(\frac{1}{1 + \kappa\lambda_i} \right)^{\kappa^{-1}}, \quad y_i = 0, 1, \dots,$$

where λ_i and $\kappa > 0$ represent the mean and the overdispersion parameter, respectively. We consider the working mean model $\lambda_i = \exp\{\mathbf{x}'_i\beta\}$ where β is a regression parameter vector of dimension $(p - 1) \times 1$ associated with covariate $\mathbf{x}_i = (x_{1i}, \dots, x_{p-1i})$, $i = 1, 2, \dots, n$. For this model, $\theta = (\omega', \gamma')$ with $\omega = (\omega_1, \omega_2, \dots, \omega_K)$ and $\gamma = (\beta', \kappa)'$. Let $\hat{\beta}$ and $\hat{\kappa}$ be the maximum likelihood estimates of β^* and κ^* , the values of β and κ under the null hypothesis that satisfy the condition $E\{U_n(\theta_0^*)\} = 0$, where $\theta_0^* = (\theta', \gamma^{*'})'$ with $\gamma^* = (\beta^{*'}, \kappa^{*'})'$.

The score vector for the ZINB model evaluated at $\theta = \hat{\theta}_0$ is $U_{n\omega}(\hat{\theta}_0) = \frac{\partial \ell_n(\theta)}{\partial \omega} \Big|_{\theta = \hat{\theta}_0}$ with

$$\frac{\partial \ell_n(\theta)}{\partial \omega_k} \Big|_{\theta = \hat{\theta}_0} = \sum_{i=1}^n \left\{ I(y_i = 0) \left(1 + \hat{\kappa} \hat{\lambda}_i \right)^{1/\hat{\kappa}} - 1 \right\} \delta_{ik},$$

where $\hat{\lambda}_i = \exp\{\mathbf{x}'_i\hat{\beta}\}$.

Matrices $G_{\omega\omega}(\theta_0^*)$, $G_{\omega\gamma}(\theta_0^*)$ and $G_{\gamma\gamma}(\theta_0^*)$ are given as follows:

$$G_{\omega\omega}(\theta_0^*) = \text{diag} \left\{ -E \left(\frac{\partial^2 \ell_n(\theta)}{\partial \omega_k^2} \Big|_{\theta = \hat{\theta}_0} \right) \right\}_{k=1}^K$$

with

$$E \left(\frac{\partial^2 \ell_n(\boldsymbol{\theta})}{\partial \omega_k^2} \middle| \boldsymbol{\theta} = \boldsymbol{\theta}_0^* \right) = \sum_{i=1}^n \left(1 - (1 + \kappa^* \lambda_i^*)^{1/\kappa^*} \right) \delta_{ik}, \quad \lambda_i^* = \exp\{\mathbf{x}_i' \boldsymbol{\beta}^*\},$$

$$G_{\omega\gamma}(\boldsymbol{\theta}_0^*) = \left(\left\{ -E \left(\frac{\partial^2 \ell_n(\boldsymbol{\theta})}{\partial \omega_k \partial \beta_j} \middle| \boldsymbol{\theta} = \boldsymbol{\theta}_0^* \right) \right\}_{(j,k)=(1,1)}^{(p-1,K)}, \quad \left\{ -E \left(\frac{\partial^2 \ell_n(\boldsymbol{\theta})}{\partial \omega_k \partial \kappa} \middle| \boldsymbol{\theta} = \boldsymbol{\theta}_0^* \right) \right\}_{k=1}^K \right) \text{ with}$$

$$E \left(\frac{\partial^2 \ell_n(\boldsymbol{\theta})}{\partial \omega_k \partial \beta_j} \middle| \boldsymbol{\theta} = \boldsymbol{\theta}_0^* \right) = \sum_{i=1}^n x_{ji} \lambda_i^* (1 + \kappa^* \lambda_i^*)^{-1} \delta_{ik}, \quad k=1, 2, \dots, K; \quad j=1, 2, \dots, p-1$$

and

$$E \left(\frac{\partial^2 \ell_n(\boldsymbol{\theta})}{\partial \omega_k \partial \kappa} \middle| \boldsymbol{\theta} = \boldsymbol{\theta}_0^* \right) = \sum_{i=1}^n \left\{ -\frac{1}{\kappa^{*2}} \log(1 + \kappa^* \lambda_i^*) + \frac{\lambda_i^*}{\kappa^* (1 + \kappa^* \lambda_i^*)} \right\} \delta_{ik}, \quad k=1, 2, \dots, K,$$

and

$$G_{\gamma\gamma}(\boldsymbol{\theta}_0^*) = \left(\begin{array}{cc} \left\{ -E \left(\frac{\partial^2 \ell_n(\boldsymbol{\theta}_0^*)}{\partial \beta_j \partial \beta_l} \middle| \boldsymbol{\theta} = \boldsymbol{\theta}_0^* \right) \right\}_{(j,l)=(1,1)}^{(p-1,p-1)} & \mathbf{0} \\ \mathbf{0}' & -E \left(\frac{\partial^2 \ell_n(\boldsymbol{\theta})}{\partial \kappa^2} \middle| \boldsymbol{\theta} = \boldsymbol{\theta}_0^* \right) \end{array} \right) \text{ with}$$

$$E \left(\frac{\partial^2 \ell_n(\boldsymbol{\theta})}{\partial \beta_j \partial \beta_l} \middle| \boldsymbol{\theta} = \boldsymbol{\theta}_0^* \right) = \sum_{i=1}^n \frac{\lambda_i^*}{\kappa^* (1 + \kappa^* \lambda_i^*)} x_{ji} x_{li}, \quad j, l=1, 2, \dots, p-1$$

and

$$E \left(\frac{\partial^2 \ell_n(\boldsymbol{\theta})}{\partial \kappa^2} \middle| \boldsymbol{\theta} = \boldsymbol{\theta}_0^* \right) = \sum_{i=1}^n \left\{ -\frac{2}{\kappa^{*3}} \log(1 + \kappa^* \lambda_i^*) + \frac{\lambda_i^*}{\kappa^{*2} (1 + \kappa^* \lambda_i^*)} \right. \\ \left. + E \left[\frac{2}{\kappa^{*3}} (\text{ddg}(y_i + \kappa^{*-1}) - \text{ddg}(\kappa^{*-1})) + \frac{1}{\kappa^{*4}} (\text{dtg}(y_i + \kappa^{*-1}) - \text{dtg}(\kappa^{*-1})) \right] \right\},$$

where $\text{ddg}(\cdot)$ and $\text{dtg}(\cdot)$ are di-gamma and tri-gamma functions respectively.

Appendix 3. Detailed calculations for the zero-inflated Poisson model

Consider the zero-inflated Poisson (ZIP) model for which the probability mass function of the homogeneous distribution is,

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

where λ_i represents the mean parameter. We consider the working mean model

$\lambda_i = \exp\{\mathbf{x}'_i \boldsymbol{\beta}\}$ where $\boldsymbol{\beta}$ is a regression parameter vector of dimension $p \times 1$ associated with covariate $\mathbf{x}_i = (x_{1i}, \dots, x_{pi})$, $i = 1, 2, \dots, n$. For this model, $\boldsymbol{\gamma} = \boldsymbol{\beta}$. Let $\hat{\boldsymbol{\beta}}$ be the maximum likelihood estimate of $\boldsymbol{\beta}^*$, the true value of $\boldsymbol{\beta}$ under the null hypothesis.

The score vector for the ZIP model evaluated at $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ $U_{n\omega}(\hat{\boldsymbol{\theta}}) = \frac{\partial \ell_n(\boldsymbol{\theta})}{\partial \boldsymbol{\omega}} \Big|_{\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}}$ with

$$\frac{\partial \ell_n(\boldsymbol{\theta})}{\partial \omega_k} \Big|_{\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}} = \sum_{i=1}^n \left\{ I\{y_i=0\} e^{\hat{\lambda}_i} - 1 \right\} \delta_{ik},$$

where $\hat{\lambda}_i = \exp\{\mathbf{x}'_i \hat{\boldsymbol{\beta}}\}$.

Matrices $G_{\omega\omega}(\boldsymbol{\theta}_0^*)$, $G_{\omega\gamma}(\boldsymbol{\theta}_0^*)$ and $G_{\gamma\gamma}(\boldsymbol{\theta}_0^*)$ are given as follows:

$$G_{\omega\omega}(\boldsymbol{\theta}_0^*) = \text{diag} \left\{ -E \left(\frac{\partial^2 \ell_n(\boldsymbol{\theta})}{\partial \omega_k^2} \Big|_{\boldsymbol{\theta} = \boldsymbol{\theta}_0^*} \right) \right\}_{k=1}^K$$

with

$$E \left(\frac{\partial^2 \ell_n(\boldsymbol{\theta})}{\partial \omega_k^2} \Big|_{\boldsymbol{\theta} = \boldsymbol{\theta}_0^*} \right) = \sum_{i=1}^n \left(1 - e^{-\lambda_i^*} \right) \delta_{ik}, \quad \lambda_i^* = \exp\{\mathbf{x}'_i \boldsymbol{\beta}^*\}$$

$$G_{\omega\gamma}(\boldsymbol{\theta}_0^*) = \left\{ -E \left(\frac{\partial^2 \ell_n(\boldsymbol{\theta})}{\partial \omega_k \partial \beta_j} \Big|_{\boldsymbol{\theta} = \boldsymbol{\theta}_0^*} \right) \right\}_{(j,k)=(1,1)}^{(p,K)}$$

with

$$E \left(\frac{\partial^2 \ell_n(\boldsymbol{\theta})}{\partial \omega_k \partial \beta_j} \Big|_{\boldsymbol{\theta} = \boldsymbol{\theta}_0^*} \right) = \sum_{i=1}^n x_{ji} \lambda_i^* \delta_{ik}, \quad k=1, \dots, K; \quad j=1, \dots, p$$

and

$$G_{\gamma\gamma}(\boldsymbol{\theta}_0^*) = \left\{ -E \left(\frac{\partial^2 \ell_n(\boldsymbol{\theta})}{\partial \beta_j \partial \beta_l} \Big|_{\boldsymbol{\theta} = \boldsymbol{\theta}_0^*} \right) \right\}_{(j,l)=(1,1)}^{(p,p)}$$

with

$$E \left(\frac{\partial^2 \ell_n(\boldsymbol{\theta})}{\partial \beta_j \partial \beta_l} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0^*} \right) = \sum_{i=1}^n \lambda_i^* x_{ji} x_{li}, \quad j, l=1, 2, \dots, p.$$

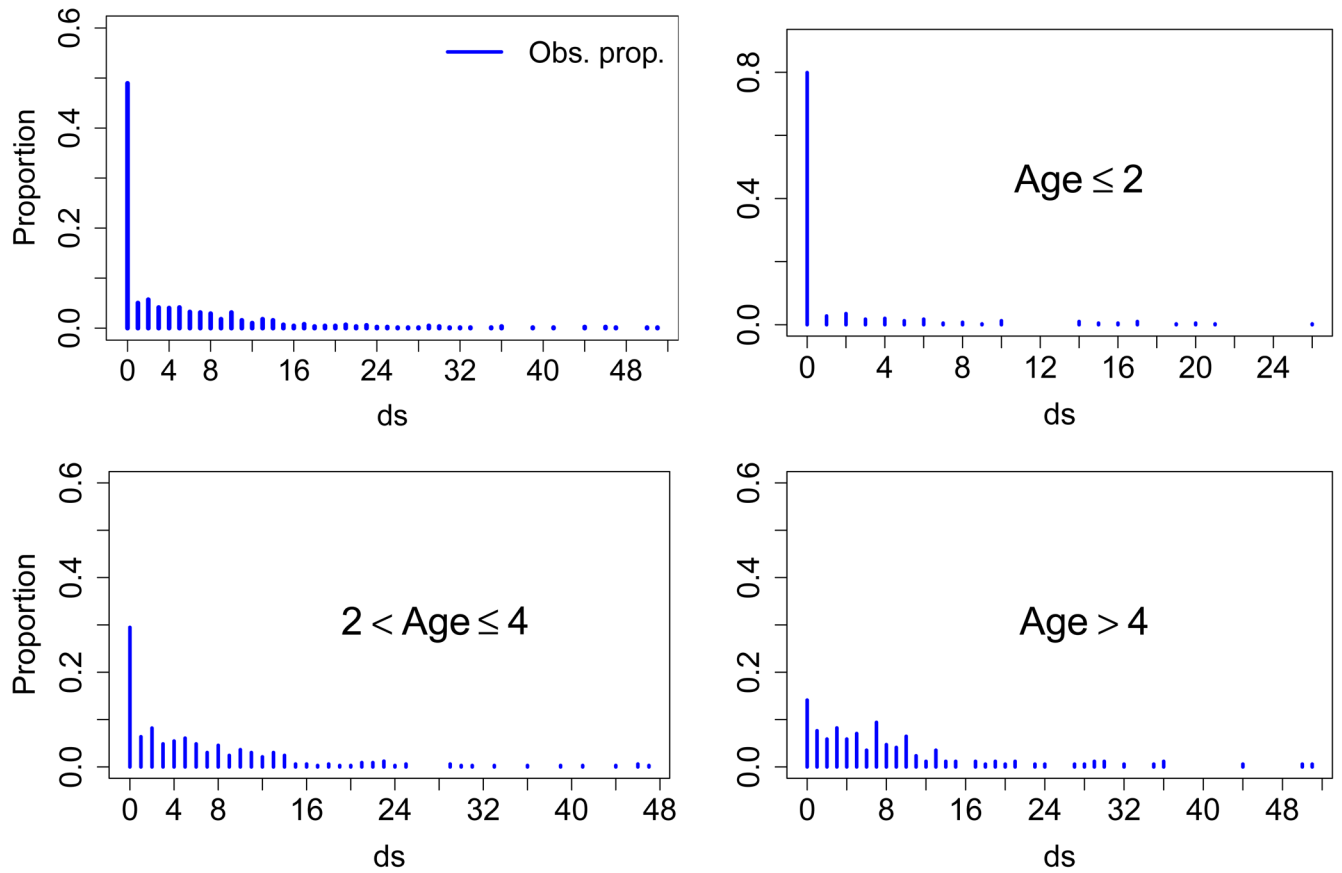


Figure 1.

Observed proportions of the number of decayed surfaces (ds) for the overall sample and for each age group.

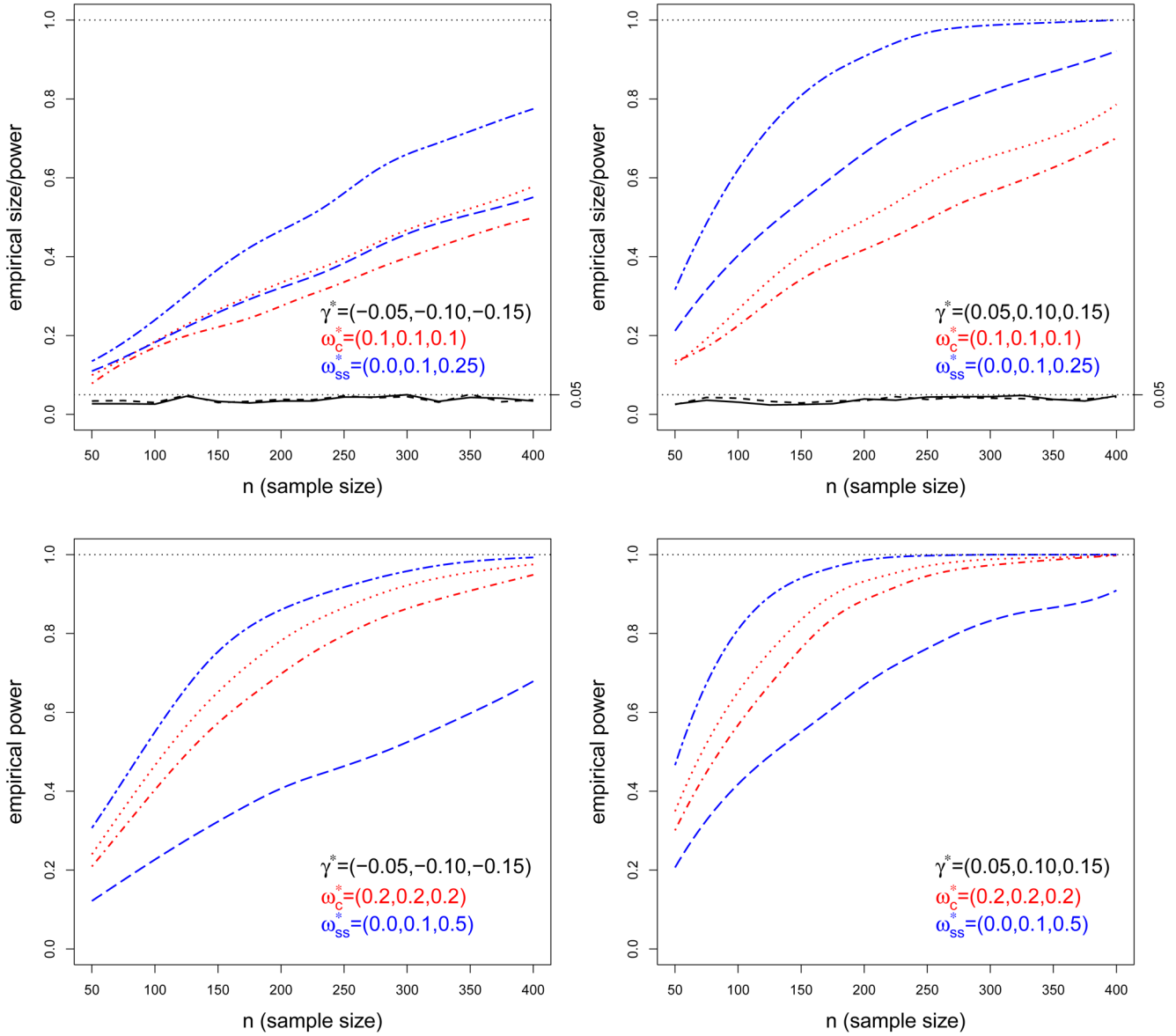


Figure 2.

Empirical size and power of constant mixing weight and stratum-specific mixing weight tests at 5% significance level, homogeneous model is Poisson. The black lines represent the data generated under the null model ($\omega_c^* = \omega_{ss}^* = 0$). The red and blue lines represent the data generated under the constant (ω_c^*) and the stratum-specific (ω_{ss}^*) mixing weight model, respectively:

True mixing weight is zero (size): ——— stratum-specific mixing weight test; - - - constant mixing weight test
 True mixing weight is constant (power): - - - - stratum-specific mixing weight test; ····· constant mixing weight test
 True mixing weight is stratum-specific (power): - - - - stratum-specific mixing weight test; ——— constant mixing weight test.

Table 1
The fitted homogeneous (null) model and heterogeneous models with constant mixing weight and stratum-specific mixing weight to dental caries index ds

Vector θ	No mixing weight (null) model			Constant mixing weight model			Stratum-specific mixing weight model		
	Estimates	Std. Errors		Estimates	Std. Errors		Estimates	Std. Errors	
β_{01}	0.2675	0.1795		0.3846	0.2605		1.6342	0.2147	
β_{11}	0.0009	0.0026		0.0010	0.0026		0.0025	0.0032	
β_{02}	1.6561	0.2015		1.7376	0.2317		1.9541	0.1450	
β_{12}	0.0011	0.0010		0.0010	0.0010		0.0008	0.0007	
β_{03}	2.2924	0.2599		2.3211	0.2503		2.3369	0.1651	
β_{13}	-0.0010	0.0012		-0.0009	0.0012		-0.0010	0.0008	
κ	2.8603	0.1905		2.4527	0.6249		0.9601	0.1107	
ω_1	-	-		0.0811	0.1218		0.7667	0.0243	
ω_2	-	-		0.0811	0.1218		0.2119	0.0329	
ω_3	-	-		0.0811	0.1218		0.0478	0.0358	
-2LogLik.		4030.5			4030.2			3791.4	
AIC		4044.5			4046.2			3811.4	
BIC		4078.1			4084.6			3859.4	

-2LogLik. = $-2\ell_n(\hat{\theta})$, AIC = $-2\ell_n(\hat{\theta}) + 2q$, BIC = $-2\ell_n(\hat{\theta}) + q \log(n)$, where $\hat{\theta}$ is the maximum likelihood estimate of θ , $q = \#$ model parameters.

Table 2

Summary statistics comparing models that assume constant and stratum-specific overdispersion parameters

Summary Statistics	Constant overdispersion parameter κ	Stratum-dependent overdispersion parameter κ
AIC	3811.4	3816.7
BIC	3859.4	3874.3
-2LogLik.^{\S}	3791.4	3792.7

$^{\S} -2\text{LogLik.} = -2 \times \text{Log-Likelihood value at } \hat{\theta}, \text{ where } \hat{\theta} \text{ is the maximum likelihood estimate of } \theta^*.$ $(-2\text{LogLik.}) = 1.3$ is the difference in $-2 \times \text{Log-Likelihood values between the constant and the stratum-dependent overdispersion parameter models. The p-value associated with } (-2\text{LogLik.})$ is 0.4780 which is computed from a χ^2_2 distribution.