



Electronic Delivery Cover Sheet

NOTICE WARNING CONCERNING COPYRIGHT RESTRICTIONS

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted materials.

Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be "used for any purpose other than private study, scholarship, or research." If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use," that user may be liable for copyright infringement.

This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve violation of copyright law.

This notice is posted in compliance with
Title 37 C. F. R., Chapter II, Part 201.14

Rapid #: -17768813

CROSS REF ID: 1582842

LENDER: FGA :: Main Library

BORROWER: VPI :: Newman Library

TYPE: Article CC:CCG

JOURNAL TITLE: Communications in statistics. Simulation and computation

USER JOURNAL TITLE: Communications in Statistics: Simulation and Computation

ARTICLE TITLE: Sample size calculations for hierarchical Poisson and zero-inflated Poisson regression models

ARTICLE AUTHOR: Channouf, N

VOLUME: 50

ISSUE: 4

MONTH:

YEAR: 2021

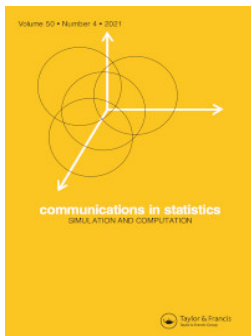
PAGES: 937-956

ISSN: 0361-0918

OCLC #:

Processed by RapidX: 7/13/2021 9:13:19 AM

This material may be protected by copyright law (Title 17 U.S. Code)



Sample size calculations for hierarchical Poisson and zero-inflated Poisson regression models

Nabil Channouf, Marc Fredette & Brenda MacGibbon

To cite this article: Nabil Channouf, Marc Fredette & Brenda MacGibbon (2021) Sample size calculations for hierarchical Poisson and zero-inflated Poisson regression models, Communications in Statistics - Simulation and Computation, 50:4, 937-956, DOI: [10.1080/03610918.2019.1577975](https://doi.org/10.1080/03610918.2019.1577975)

To link to this article: <https://doi.org/10.1080/03610918.2019.1577975>



Published online: 12 Mar 2019.



Submit your article to this journal [↗](#)



Article views: 103



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)



Sample size calculations for hierarchical Poisson and zero-inflated Poisson regression models

Nabil Channouf^a, Marc Fredette^b, and Brenda MacGibbon^c

^aDepartment of Operations Management and Business Statistics, College of Economics & Political Sciences, Sultan Qaboos University, Muscat, Oman; ^bDepartment of Management Sciences, HEC Montréal, Montréal, Canada; ^cDepartment of Mathematics, Université du Québec à Montréal, Montréal, Canada

ABSTRACT

In biomedical research there is a growing interest in the use of hierarchical Poisson regression models. Although sample size calculations for testing parameters in a Poisson regression model with prespecified power and size have been previously done, very little attention has been paid to this problem for the hierarchical model. We propose to use Monte Carlo simulations to calculate the sample size necessary to perform the Wald tests when the number of clusters is fixed in advance, but the cluster size is variable. The effect of the number of clusters and the covariance structure of the fixed effects is also studied. The method and the simulation study are also extended to the case of the hierarchical zero-inflated Poisson regression model in order to obtain analogous results there. The method is also illustrated on an interesting real dataset.

ARTICLE HISTORY

Received 24 August 2018
Accepted 29 January 2019

KEYWORDS

Wald test; Score function; Information matrix; Hierarchical generalized linear models; Intraclass correlation; Monte Carlo simulations

1. Introduction

In the applied literature hierarchical Poisson regression models are used in interesting ways such as in the studies of clustered data of Christiansen and Morris (1997), Yau and Lee (2001) and Royle (2008). We also cite Hall (2000), Tu and Piegorsch (2003) and Roy et al. (2007). In biomedical research the clusters could be hospitals, and for the patients within hospitals there would usually be several explanatory covariates such as sex, age, etc. For such studies, calculation of the sample size needed for a test with pre-specified power and size is usually a necessary part of the design.

Sample size calculations for the standard nonhierarchical Poisson model are often based on the Wald test (Signorini 1991; Shieh 2001, 2005; Channouf, Fredette, and MacGibbon 2014), the score test (Self and Mauritsen 1988) or the likelihood ratio test (Self, Mauritsen, and Ohara 1992; Shieh 2000).

Although statistical inferential methods for the hierarchical generalized linear model have been proposed by Lee and Nelder (1996, 2001), and Ha and Lee (2003), with few exceptions the problem of sample size determination has largely been ignored. The exceptions include, Liu and Liang (1997), who extended the method of Self and Mauritsen (1988) to the multivariate case

with correlated data using the generalized estimating equation method and the quasi-score statistic (Liang and Zeger 1986), and Matsui (2005), who used asymptotic approximations for a two-sample non-parametric test to develop a new formula for sample size calculations for clinical trials with Poisson or over-dispersed Poisson data.

We have chosen here to use the Wald test for sample size calculations for the hierarchical Poisson regression model: the test is frequently used due to its accessibility and this field of research has recent major contributions in sample size calculations for this test for the standard nonhierarchical Poisson model (Shieh 2001, 2005; Channouf, Fredette, and MacGibbon 2014).

Another problem with count data is that the observed data may have an excess number of zeros and thus the Poisson model may not be an adequate model for the counts in such situations. The more complex zero-inflated Poisson (ZIP) model could be a better choice here. For the nonhierarchical model, Channouf, Fredette, and MacGibbon (2014) used the estimation approach of Lambert (1992), based on the complete-data likelihood function (Hall 2000), to do sample size calculations for the ZIP model, while Williamson et al. (2007) used the incomplete log-likelihood function (equation (3) in Lambert (1992)) to do power calculations for ZIP and ZINB regression models. Hall (2000) presented a detailed study of the zero-inflated Poisson and binomial regression models and their extension to correlated data. The idea of using the zero-inflated Poisson model with correlated data in biomedical applications was first studied by Yau and Lee (2001), but sample size calculations were not considered.

Influenced by the original work of Shieh (2001) for the standard Poisson regression model when the covariates were assumed to have a finite support, the major contribution of this paper is the development of the methodology necessary to do sample size calculations using Monte Carlo simulations for the Wald test for different hierarchical Poisson regression models, and the extension of this methodology to the case of the hierarchical Zero-inflated Poisson model. It should be noted that, by using numerical integration and quasi-Monte Carlo simulations (Bratley and Fox 1988), our methodology works for continuous covariates. Without loss of generality, we restrict our numerical examples to the case of equal cluster size. For parameter estimation we use the h -likelihood method proposed by Lee and Nelder (1996). We also study the effect of different correlation structures on the sample size. Our work is motivated by the fact that no existing method was proposed in the literature for sample size calculations for hierarchical Poisson regression models.

The remainder of the paper is organized as follows. In the next section, we introduce the hierarchical Poisson regression model and describe our use of the h -likelihood method and the approach to sample size calculations for this model. In Sec. 3, we present the extension of our approach to zero-inflated hierarchical Poisson regression models using Hall (2000) estimation techniques. Section 4 contains numerical examples with an accompanying discussion of these results. Section 5 contains an illustrative data example. In Sec. 6, we draw some conclusions and discuss further directions of research related to this work.

2. Hierarchical Poisson regression model

The hierarchical Poisson regression model belongs to the class of hierarchical generalized linear models in which the distribution of the random effects is not restricted to be normal,

but of general form. Often, this distribution is conjugate to that of the outcomes. For such a model having both random and fixed effects, we use a conditional approach as in Lee and Nelder (1996).

In general for the hierarchical Poisson model, we consider the case of clustered data, where each vector of responses \mathbf{Y}_i denotes a cluster, and where the outcomes are $Y_{i,j}$, for $i = 1, \dots, n; j = 1, \dots, n_i$, with $N = \sum_{i=1}^n n_i$. Given a vector of unobserved random effects $\mathbf{U}_i = (U_{0,i}, \dots, U_{r,i})$ associated with the cluster \mathbf{Y}_i , the outcomes $Y_{i,1}, \dots, Y_{i,n_i}$ are mutually independent and have a Poisson distribution with mean $E[Y_{i,j}|\mathbf{U}_i] = \lambda'_{i,j}$, where the linear predictor takes the form:

$$\log(\lambda'_{i,j}) = \boldsymbol{\beta}^T \mathbf{X}_{i,j} + \nu(\mathbf{U}_i)^T \mathbf{Z}_{i,j}, \quad i = 1, \dots, n; \quad j = 1, \dots, n_i,$$

where $\mathbf{X}_{i,j}^T = (1, X_{1,i,j}, \dots, X_{p,i,j})$ represents the explanatory covariates associated with the fixed $(p+1)$ effects $\boldsymbol{\beta}^T = (\beta_0, \beta_1, \dots, \beta_p)$, and $\mathbf{Z}_{i,j}^T = (1, Z_{1,i,j}, \dots, Z_{r,i,j})$ the covariates associated with the $(r+1)$ random effects $\mathbf{v}_i = \nu(\mathbf{U}_i)$ for a strictly monotonic link function ν . The random variables $\mathbf{U}_{0,i}, \dots, \mathbf{U}_{r,i}$ are assumed independent and the distribution of each is assumed to belong to the exponential family and is characterized by one dispersion parameter α_k .

The choice of the link function ν is important and affects the estimation of the parameters and the dispersion. Some special distributions for the responses \mathbf{Y}_i and the random variables $\mathbf{U}_{0,i}, \dots, \mathbf{U}_{r,i}$ involve conjugate models. For such models, in addition to the joint likelihood function, as defined by (Lee and Nelder 1996, p. 621), to estimate the random effect and the parameters of the model, an explicit form of the marginal likelihood can be obtained as well without any integration.

Our numerical examples include the usual conjugate hierarchical Poisson-Gamma model with different covariate distributions for the explanatory covariates. Analogous results could also be obtained for other conjugate hierarchical models in the generalized linear exponential family.

2.1. Hierarchical likelihood estimation

The hierarchical likelihood method, called *h*-likelihood, was introduced by Lee and Nelder (1996), and it is based on the joint distribution of \mathbf{Y} and \mathbf{U} . It should be noted that Kuk and Cheng (1999) gave a comparative study of different algorithms for maximum likelihood estimates in random effects models. They argued that the *h*-likelihood can be inappropriate, especially with binary data. Lee, Nelder, and Noh (2007) gave a detailed study of the *h*-likelihood and effectively responded to the criticism of Kuk and Cheng (1999).

For hierarchical generalized linear models like the hierarchical Poisson model, both fixed effects and random effects are of interest, so they are estimated simultaneously. In general, for models where the joint distribution is used, some difficulties arise and complex integrals need to be evaluated, especially for high dimensions. Different methods have been presented in order to circumvent these analytical difficulties by using approximations, such as penalized quasi-likelihood estimation or restricted maximum likelihood estimation, where Laplace transforms are used to approximate the integrals. For more details about likelihood-based inference in generalized linear mixed models, see

(Song 2007, chapter 7). The h -likelihood estimation method can be regarded as an extension of penalized quasi-likelihood estimation, since it does not involve any integration (Jiang 2007, chapter 3).

Assume we have a random sample from the joint distribution of $(\mathbf{Y}, \mathbf{X}, \mathbf{Z})$ with conditional p.d.f. $f(\mathbf{Y}, \mathbf{X}, \mathbf{Z}|\mathbf{U}) = f(\mathbf{Y}|\mathbf{X}, \mathbf{Z}, \mathbf{U})f(\mathbf{X}, \mathbf{Z})$, since \mathbf{X} and \mathbf{Z} do not depend on \mathbf{U} and their joint p.d.f $f(\mathbf{X}, \mathbf{Z})$ does not depend on any of the unknown parameters of the model $(\beta_0, \beta_1, \dots, \beta_p, \alpha_0, \dots, \alpha_r)$. The joint log-likelihood function for the hierarchical Poisson model, called the h -likelihood, associated with the data, the parameters β and the random effects \mathbf{U} , is given by:

$$h(\beta; \mathbf{y}, \mathbf{x}, \mathbf{z}, \mathbf{v}) = \log L(\beta; \mathbf{y}, \mathbf{x}, \mathbf{z}, \mathbf{v}) = l(\beta; \mathbf{y}, \mathbf{x}, \mathbf{z}|\mathbf{v}) + l(\alpha_0, \dots, \alpha_r; \mathbf{v}) \quad (1)$$

where $l(\beta; \mathbf{y}, \mathbf{x}, \mathbf{z}|\mathbf{v}) = \sum_{i=1}^n \sum_{j=1}^{n_i} \{\log f_Y(y_{i,j}|x_{i,j}, z_{i,j}, v_i) + \log f_{X,Z}(x_{i,j}, z_{i,j})\}$, with f_Y having the p.d.f. of a Poisson distribution \mathbf{Y} and $l(\alpha_0, \dots, \alpha_r; \mathbf{v})$ representing the log-likelihood of \mathbf{v} with parameters $\alpha_0, \dots, \alpha_r$. Note that we can use both the density of $\mathbf{Y}|\mathbf{v}(\mathbf{U})$ or $\mathbf{Y}|\mathbf{U}$, since \mathbf{v} is a strictly monotonic function of \mathbf{U} . We also note that in the estimation of the parameters in the hierarchical generalized linear model, these parameters are conditioned on the random effects, which are not observed. However, they can be estimated jointly with the fixed effects.

For this estimation problem of dimension $(p + n + 1)$, the score function for N observations is defined as follows

$$S_N(\beta; \mathbf{y}, \mathbf{x}, \mathbf{z}, \mathbf{v}) = \begin{cases} \frac{\partial h(\beta; \mathbf{y}, \mathbf{x}, \mathbf{z}, \mathbf{v})}{\partial \beta} = \frac{\partial l(\beta; \mathbf{y}, \mathbf{x}, \mathbf{z}|\mathbf{v})}{\partial \beta} \\ \frac{\partial h(\beta; \mathbf{y}, \mathbf{x}, \mathbf{z}, \mathbf{v})}{\partial \mathbf{v}} = \frac{\partial l(\beta; \mathbf{y}, \mathbf{x}, \mathbf{z}|\mathbf{v})}{\partial \mathbf{v}} + \frac{\partial l(\alpha_0, \dots, \alpha_r; \mathbf{v})}{\partial \mathbf{v}}. \end{cases} \quad (2)$$

The maximum h -likelihood estimates of β and \mathbf{v} are then obtained by solving the system of equations $S_N(\beta; \mathbf{y}, \mathbf{x}, \mathbf{z}, \mathbf{v}) = 0$. For the estimation of the dispersion parameters $\alpha_0, \dots, \alpha_r$, a second stage procedure is used on the estimators of the fixed and random effects (Lee and Nelder 1996).

2.2. Hierarchical Poisson-Gamma model with one random effect

Here we consider the special case of the hierarchical Poisson model with its conjugate gamma distribution, with the logarithm as the link function and with only one (intercept) random factor \mathbf{U}_0 for each cluster. This type of model is often called “the random-intercept model”.

For each group (cluster) $i = 1, \dots, n$, we have outcomes $Y_{i,1}, \dots, Y_{i,n_i}$, a vector of fixed effects $(\beta_0, \beta_1, \dots, \beta_p)^T$ associated with $\mathbf{X}_{i,j}^T = (1, X_{1,i,j}, \dots, X_{p,i,j})$, for $j = 1, \dots, n_i$, with $N = \sum_{i=1}^n n_i$, and one random effect $U_{0,i}$ having an i.i.d. gamma distribution with a common dispersion parameter α_0 and satisfying $E[U_{0,i}] = 1$. We have $Y_{i,j}|\mathbf{u}_{0,i} \sim \text{Poisson}(\lambda'_{i,j})$, where the conditional mean $E[Y_{i,j}|\mathbf{u}_{0,i}] = \lambda'_{i,j}$ satisfies:

$$\log(\lambda'_{i,j}) = \beta^T \mathbf{X}_{i,j} + v_{0,i}, \quad (3)$$

with $v_{0,i} = \log(u_{0,i})$. From (1), the h -likelihood function associated with the parameters $\beta^T = (\beta_0, \beta_1, \dots, \beta_p)$ and the unobserved random effects $v_{0,i}$, for a random sample

$(y_{i,j}, x_{1,i,j}, \dots, x_{p,i,j})$, is given by:

$$\begin{aligned} h(\boldsymbol{\beta}; \mathbf{y}, \mathbf{x}, \mathbf{v}_0) &= l(\boldsymbol{\beta}; \mathbf{y}, \mathbf{x} | \mathbf{v}_0) + l(\alpha_0; \mathbf{v}_0) \\ &= \sum_{i=1}^n \sum_{j=1}^{n_i} \left\{ y_{i,j} \left(\beta^T x_{i,j} + v_{0,i} \right) - e^{\beta^T x_{i,j} + v_{0,i}} - \log(y_{i,j}!) + \log f_X(x_{i,j}) \right\} \\ &\quad + \sum_{i=1}^n \left\{ \alpha_0 \log(\alpha_0) + \alpha_0 v_{0,i} - \alpha_0 e^{v_{0,i}} - \log \Gamma(\alpha_0) \right\}. \end{aligned} \quad (4)$$

Once α_0 is specified or fixed in the first step as in Lee and Nelder (1996), the estimates of the random effects can be easily obtained from the score function and the second equation of (2). We have:

$$\hat{u}_{0,i} = e^{\hat{v}_{0,i}} = \frac{y_{i,+} + \alpha_0}{\lambda_{i,+} + \alpha_0}, \quad (5)$$

where $y_{i,+} = \sum_{j=1}^{n_i} y_{i,j}$ and $\lambda_{i,+} = \sum_{j=1}^{n_i} \lambda_{i,j}$, with $\lambda_{i,j} = E[Y_{i,j}]$ the unconditional mean of $Y_{i,j}$. It should be noted that $E[U_{0,i}] = 1$ for any given α_0 , and when $\alpha_0 \rightarrow \infty$, we recover the Poisson regression model with fixed effects.

2.3. Sample size determination

We want to find the minimum required sample size in order to achieve a given power with the Wald test for the hypothesis $H_0 : \beta_s = 0$ against $H_1 : \beta_s \neq 0$, for any given $s \in \{1, \dots, p\}$. To do this, we generalize the approach used by Signorini (1991) and Shieh (2001) for nonhierarchical models. In our case, we also prefer the conditional expectation, conditional on the covariates \mathbf{X} , using Monte Carlo simulation to generate samples from these covariates instead of the full expectation as used by Shieh (2001). Shieh (2001) chose to solve $\lim_{N \rightarrow \infty} E[S_N(\beta_0, \beta_1, \dots, \beta_{s-1}, 0, \beta_{s+1}, \dots, \beta_p)]/N$, where the full expectation is always taken with respect to the true value of β . In his case, this expectation can be calculated because it is always expressed as a finite sum since the support of the covariates \mathbf{X} is treated as finite. Even for continuous random variables such as the normal distribution, he evaluates it at a finite number of points. We, however, choose to avoid using such a finite approximation in the more complex case of the hierarchical Poisson regression model. In Eq. (2) we choose $N=K$ sufficiently large as a first step and calculate $S_N(\boldsymbol{\beta}) = S_K(\boldsymbol{\beta})$ directly by substituting the values of the covariates from our Monte Carlo simulations and then use K as the subscript for the score function. We first calculate the estimates of the parameters $(\beta_0^*, \beta_1^*, \dots, \beta_{s-1}^*, 0, \beta_{s+1}^*, \dots, \beta_p^*)$ and $\hat{\mathbf{v}}_0^T = (\hat{v}_{0,1}^*, \dots, \hat{v}_{0,n}^*)$ under H_0 , by solving the following system of equations:

$$E \left[S_K(\beta_0, \beta_1, \dots, \beta_{s-1}, 0, \beta_{s+1}, \dots, \beta_p, \mathbf{v}_0) | \mathbf{X} = \mathbf{x} \right] = 0, \quad (6)$$

where this conditional expectation is taken with respect to the true values of β_0, \dots, β_p . The score function is given by:

$$S_K(\boldsymbol{\beta}, \mathbf{v}_0) = \begin{cases} \frac{\partial h(\boldsymbol{\beta}; \mathbf{y}, \mathbf{x}, \mathbf{v}_0)}{\partial \beta_0} = \sum_{i=1}^n \sum_{j=1}^{n_i} (y_{i,j} - e^{\boldsymbol{\beta}^T \mathbf{x}_{i,j} + v_{0,i}}), \\ \frac{\partial h(\boldsymbol{\beta}; \mathbf{y}, \mathbf{x}, \mathbf{v}_0)}{\partial \beta_l} = \sum_{i=1}^n \sum_{j=1}^{n_i} x_{l,i,j} (y_{i,j} - e^{\boldsymbol{\beta}^T \mathbf{x}_{i,j} + v_{0,i}}); \quad l = 1, \dots, p, \\ \frac{\partial h(\boldsymbol{\beta}; \mathbf{y}, \mathbf{x}, \mathbf{v}_0)}{\partial v_{0,i}} = \alpha_0 - \alpha_0 e^{v_{0,i}} + \sum_{j=1}^{n_i} (y_{i,j} - e^{\boldsymbol{\beta}^T \mathbf{x}_{i,j} + v_{0,i}}); \quad i = 1, \dots, n. \end{cases}$$

In order to estimate the variance-covariance matrix of the maximum h -likelihood estimators of $\boldsymbol{\beta}$ and \mathbf{v}_0 , we use the Fisher information matrix by considering

$$V(\boldsymbol{\beta}, \mathbf{v}_0) = I(\boldsymbol{\beta}, \mathbf{v}_0)^{-1}, \quad (7)$$

where

$$I(\boldsymbol{\beta}, \mathbf{v}_0) = -\frac{1}{K} \begin{pmatrix} \mathbb{E} \left[\frac{\partial^2 h(\boldsymbol{\beta}; \mathbf{y}, \mathbf{x}, \mathbf{v}_0)}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T} \right] & \mathbb{E} \left[\frac{\partial^2 h(\boldsymbol{\beta}; \mathbf{y}, \mathbf{x}, \mathbf{v}_0)}{\partial \boldsymbol{\beta} \partial \mathbf{v}_0^T} \right] \\ \mathbb{E} \left[\frac{\partial^2 h(\boldsymbol{\beta}; \mathbf{y}, \mathbf{x}, \mathbf{v}_0)}{\partial \boldsymbol{\beta} \partial \mathbf{v}_0^T} \right]^T & \mathbb{E} \left[\frac{\partial^2 h(\boldsymbol{\beta}; \mathbf{y}, \mathbf{x}, \mathbf{v}_0)}{\partial \mathbf{v}_0 \partial \mathbf{v}_0^T} \right] \end{pmatrix}, \quad (8)$$

and where the asymptotic normality of the maximum h -likelihood estimators of $\boldsymbol{\beta}$ and \mathbf{v}_0 holds as the $n_i \rightarrow \infty$, at the same rate, whereas the number of clusters n remains constant (Lee and Nelder 1996). Xia, Zhang, and Ying (2006) also studied the asymptotic behavior of these estimators. They showed their strong consistency and calculated the convergence rate under the assumption of large number of clusters ($n \rightarrow \infty$).

Without loss of generality, let us assume that the test concerns the parameter β_1 , and we test $H_0 : \beta_1 = 0$, against the alternative $H_1 : \beta_1 \neq 0$, for a given power. Since asymptotic normality of the estimates of the $\boldsymbol{\beta}$ parameters holds, we can use Eqs. (7) and (8), as proposed by Signorini (1991) and Shieh (2001), to calculate the necessary variances in the formula for estimating the minimum sample size N_s , for $s = 1$.

$$N_s = \left\lceil \left(\frac{V(\beta_0^*, 0, \beta_2^*, \dots, \beta_p^*, \mathbf{v}_0^*)_{2,2}^{1/2} Z_{\alpha/2} + V(\beta_0, \beta_1, \dots, \beta_p, \mathbf{v}_0)_{2,2}^{1/2} Z_{1-\text{Power}}}{\beta_1} \right)^2 \right\rceil, \quad (9)$$

where Z_q is the $100(1-q)$ th percentile of the standard normal distribution, α is the size of the test, $V(\beta_0, \beta_1, \dots, \beta_p, \mathbf{v}_0)$ is computed with the prespecified values and $V(\beta_0^*, 0, \beta_2^*, \dots, \beta_p^*, \mathbf{v}_0^*)$ is computed with the values obtained in (6). It should be noted that here the prespecified value of the parameter β_1 represents the difference we want to detect by the test for a given power and size.

The procedure can be easily reversed to calculate the power of the test from Eq. (9) for given parameter values $(\beta_0, \beta_1, \dots, \beta_p)$ and a sample size N , with a specified significance level α .

The calculation of the sample size necessary for a significance test with specified power and size on any one parameter $\beta_s, s = 1, \dots, p$, is summarized in Algorithm 1.

Algorithm 1. Sample size calculation for the hierarchical Poisson-gamma model with one random effect (intercept).

The algorithm has the following input:

- The desired *Power* and the significance level α .
- The value of the difference Δ that is to be detected. This difference is set equal to e^{β_s} , where β_s is the parameter being tested.
- The distribution of the covariates \mathbf{X} in Eq. (3).
- The parameters of U_0 are chosen so that $\mathbb{E}[U_0] = 1$; $U_0 \sim \Gamma(\alpha_0, 1/\alpha_0)$.
- Some reasonable values for the other $\beta_k, k = 1, \dots, p$, also have to be specified. (These could come from previous studies, a pilot study or an educated guess). The value of β_0 is chosen to satisfy an overall fixed mean $\bar{\lambda} = E[e^{\beta^T \mathbf{X}}]$.
- An initial number K of simulated sets of covariates. Empirical studies suggest that the value of K has little effect on the results and that $K = 100$ suffices.
- The number of clusters n is specified.
- The desired number of Monte Carlo replications B .

For normal covariates, the quasi-Monte Carlo simulation is used for integration, whereas analytical formula can be obtained for other distributions (e.g. Bernoulli or multinomial).

The algorithm then returns the minimum sample size needed for the significance test. It has the following steps:

1. **Generate** a sample of K values from the distribution of \mathbf{X} and compute the rates $\lambda_i(X), i = 1, \dots, K$, under H_1 .
 2. **Compute** the values $\beta^* = (\beta_0^*, \beta_1^*, \dots, \beta_{s-1}^*, 0, \beta_{s+1}^*, \dots, \beta_p^*)$, and $u_{0,1}, \dots, u_{0,n}$, under H_0 , by solving the system of nonlinear equations in (6) with $N = K$, using the relation in (5);
 3. **Compute** the Fisher information matrix using (8) to obtain the variance functions $V(\beta^*, v_0^*)$ and $V(\beta, v_0)$ in (7) ($v_0 = \log(u_0)$), and then take its $(s+1)$ -th diagonal elements;
 4. **Repeat** steps 1 to 3 B times and each time return the sample size $N_{s,j}$ obtained from Eq. (9), for $j = 1, \dots, B$.
 5. **Take** the average value over the outputs of the B replications, $N_s = \frac{\sum_{j=1}^B N_{s,j}}{B}$.
-

3. Extension to the hierarchical zero-inflated Poisson regression model

For this type of model, we modify the methods of Hall (2000) who studied data having excess zeros using the generalized linear mixed model. We adapt his inference methodology for the regression parameters to the hierarchical zero-inflated Poisson model by using what he called the complete-data log-likelihood with one random intercept. However, when calculating the log-likelihood function, we consider the gamma distribution for the intercept instead of the normal one used by Hall (2000).

We again consider responses and covariates that come from clustered data. We assume that for the random effects $\mathbf{U}_i = (U_{0,i}, \dots, U_{r,i})$, and link functions $v_i = v(U_i)$, the outcomes $Y_{i,j}$, for $i = 1, \dots, n$, and $j = 1, \dots, n_i$, have the following conditional probability density function:

$$P(Y_{i,j} = y_{i,j} | u_i) = \begin{cases} \pi_{i,j} + (1 - \pi_{i,j})e^{-\lambda'_{i,j}} & \text{if } y_{i,j} = 0 \\ (1 - \pi_{i,j}) \frac{e^{-\lambda'_{i,j}} \lambda'^{y_{i,j}}_{i,j}}{y_{i,j}!} & \text{if } y_{i,j} > 0, \end{cases}$$

with $E[Y_{i,j} | u_i] = (1 - \pi_{i,j})\lambda'_{i,j}$ and $\text{Var}(Y_{i,j} | u_i) = (1 - \pi_{i,j})\lambda'_{i,j}(1 + \pi_{i,j}\lambda'_{i,j})$. In this model, the zeros for the outcomes can come from two states, the excess zeros with probability $\pi_{i,j}$, or Poisson with a probability $(1 - \pi_{i,j})e^{-\lambda'_{i,j}}$. In other words, $Y_{i,j}$ equals zero with probability $\pi_{i,j}$ or follows a Poisson distribution with probability $(1 - \pi_{i,j})$. Both sets of parameters $\lambda'_{i,j}$ and $\pi_{i,j}$ are modeled by link functions in order to have linear predictors as follows:

$$\text{logit}(\pi_{i,j}) = \gamma^T \mathbf{G}_{i,j} = \gamma_0 + \gamma_1 G_{1,i,j} + \dots + \gamma_q G_{q,i,j}, \quad (10)$$

with $\mathbf{G}_{i,j}^T = (1, G_{1,i,j}, \dots, G_{q,i,j})$, the $(q + 1)$ elements of the covariates, and $\gamma^T = (\gamma_0, \gamma_1, \dots, \gamma_q)$, a vector of $(q + 1)$ parameters. For the counts, we have

$$\begin{aligned} \log(\lambda'_{i,j}) &= \beta^T \mathbf{X}_{i,j} + v_i^T \mathbf{Z}_{i,j} \\ &= \beta_0 + \beta_1 X_{1,i,j} + \dots + \beta_p X_{p,i,j} + v_{0,i} + v_{1,i} Z_{1,i,j} + \dots + v_{r,i} Z_{r,i,j}, \end{aligned} \quad (11)$$

with $\mathbf{X}_{i,j}^T = (1, X_{1,i,j}, \dots, X_{p,i,j})$, the $(p + 1)$ elements of the covariates, $\beta^T = (\beta_0, \beta_1, \dots, \beta_p)$, a vector of $(p + 1)$ parameters, $\mathbf{Z}_{i,j}^T = (1, Z_{1,i,j}, \dots, Z_{r,i,j})$, the elements of the $(r + 1)$ covariates associated with the random effects $\mathbf{v}_i^T = (v_{0,i}, v_{1,i}, \dots, v_{r,i})$.

For a hierarchical ZIP model with one random gamma effect $\mathbf{U}_0 = e^{\mathbf{v}_0}$ (“the random-intercept model”), the joint log-likelihood combines the logarithm of the density of the outcomes conditional on \mathbf{v}_0 with parameters γ and β , and the logarithm of the density function of \mathbf{v}_0 with parameter α_0 . We have:

$$h(\gamma, \beta; \mathbf{y}, \mathbf{g}, \mathbf{x}, \mathbf{v}_0) = l(\gamma, \beta; \mathbf{y}, \mathbf{g}, \mathbf{x} | \mathbf{v}_0) + l(\alpha_0; \mathbf{v}_0),$$

where $l(\alpha_0; \mathbf{v}_0)$ is computed as for the hierarchical Poisson regression model given in (4), and $l(\gamma, \beta; \mathbf{y}, \mathbf{g}, \mathbf{x} | \mathbf{v}_0)$ is calculated following Hall (2000). Then for the h -likelihood, we obtain:

$$\begin{aligned} h(\gamma, \beta; \mathbf{y}, \mathbf{g}, \mathbf{x}, \mathbf{v}_0) &= \sum_{i=1}^n \sum_{j=1}^{n_i} \left\{ T_{i,j} \gamma^T \mathbf{g}_{i,j} - \log(1 + e^{\gamma^T \mathbf{g}_{i,j}}) \right\} \\ &+ \sum_{i=1}^n \sum_{j=1}^{n_i} (1 - T_{i,j}) \left\{ y_{i,j} \left(\beta^T \mathbf{x}_{i,j} + v_{0,i} \right) - e^{\beta^T \mathbf{x}_{i,j} + v_{0,i}} - \log(y_{i,j}!) \right\} \\ &+ \sum_{i=1}^n \sum_{j=1}^{n_i} \log f_{\mathbf{x}, \mathbf{Z}}(x_{i,j}, z_{i,j}) \\ &+ \sum_{i=1}^n \left\{ \alpha_0 \log(\alpha_0) + \alpha_0 v_{0,i} - \alpha_0 e^{v_{0,i}} - \log \Gamma(\alpha_0) \right\}, \end{aligned} \quad (12)$$

with

$$T_{ij} = \begin{cases} 1 & \text{if } Y_{ij} \text{ comes from excess zeros,} \\ 0 & \text{if } Y_{ij} \text{ comes from Poisson,} \end{cases}$$

for $i = 1, \dots, n$, and $j = 1, \dots, n_i$.

The sample size determination is then analogous to the methodology for the hierarchical Poisson regression model, since we use the same formula for the minimum sample size N_s , given by (9). However, the computations become more complex, since we have additional parameters for the logistic component, namely $\gamma^T = (\gamma_0, \gamma_1, \dots, \gamma_q)$.

To find the minimum required sample size for the Wald test, we have the following score function:

$$S_N(\gamma, \beta, \mathbf{v}_0) = \begin{cases} \frac{\partial l(\gamma, \beta; \mathbf{y}, \mathbf{g}, \mathbf{x}, \mathbf{v}_0)}{\partial \gamma_0} = \sum_{i=1}^n \sum_{j=1}^{n_i} \left\{ T_{ij} - \frac{e^{\gamma^T \mathbf{g}_{ij}}}{1 + e^{\gamma^T \mathbf{g}_{ij}}} \right\} \\ \frac{\partial l(\gamma, \beta; \mathbf{y}, \mathbf{g}, \mathbf{x}, \mathbf{v}_0)}{\partial \gamma_l} = \sum_{i=1}^n \sum_{j=1}^{n_i} g_{l,ij} \left\{ T_{ij} - \frac{e^{\gamma^T \mathbf{g}_{ij}}}{1 + e^{\gamma^T \mathbf{g}_{ij}}} \right\}; & l = 1, \dots, q. \\ \frac{\partial l(\gamma, \beta; \mathbf{y}, \mathbf{g}, \mathbf{x}, \mathbf{v}_0)}{\partial \beta_0} = \sum_{i=1}^n \sum_{j=1}^{n_i} (1 - T_{ij}) \{ y_{ij} - e^{\beta^T \mathbf{x}_{ij}} \}, \\ \frac{\partial l(\gamma, \beta; \mathbf{y}, \mathbf{g}, \mathbf{x}, \mathbf{v}_0)}{\partial \beta_k} = \sum_{i=1}^n \sum_{j=1}^{n_i} (1 - T_{ij}) \{ x_{k,ij} (y_{ij} - e^{\beta^T \mathbf{x}_{ij}}) \}; & k = 1, \dots, p. \\ \frac{\partial l(\gamma, \beta; \mathbf{y}, \mathbf{g}, \mathbf{x}, \mathbf{v}_0)}{\partial v_{0,i}} = \alpha_0 - \alpha_0 e^{v_{0,i}} + \sum_{j=1}^{n_i} (1 - T_{ij}) \{ y_{ij} - e^{\beta^T \mathbf{x}_{ij}} \}; & i = 1, \dots, n. \end{cases}$$

The maximum likelihood estimates of γ and β are given by the solution of the system $S_N(\gamma, \beta, \mathbf{v}_0) = 0$, and their variance-covariance matrix $V(\gamma, \beta, \mathbf{v}_0)$ is estimated by the inverse of the Fisher information matrix $I(\gamma, \beta, \mathbf{v}_0)$, given by:

$$I(\gamma, \beta, \mathbf{v}_0) = -\frac{1}{N} \begin{pmatrix} \mathbf{E} \left[\frac{\partial^2 l(\gamma, \beta; \mathbf{y}, \mathbf{g}, \mathbf{x}, \mathbf{v}_0)}{\partial(\gamma, \beta)^T \partial(\gamma, \beta)} \right] & \mathbf{E} \left[\frac{\partial^2 l(\gamma, \beta; \mathbf{y}, \mathbf{g}, \mathbf{x}, \mathbf{v}_0)}{\partial(\gamma, \beta)^T \partial \mathbf{v}_0^T} \right] \\ \mathbf{E} \left[\frac{\partial^2 l(\gamma, \beta; \mathbf{y}, \mathbf{g}, \mathbf{x}, \mathbf{v}_0)}{\partial(\gamma, \beta)^T \partial \mathbf{v}_0^T} \right]^T & \mathbf{E} \left[\frac{\partial^2 l(\gamma, \beta; \mathbf{y}, \mathbf{g}, \mathbf{x}, \mathbf{v}_0)}{\partial \mathbf{v}_0 \partial \mathbf{v}_0^T} \right] \end{pmatrix}. \quad (13)$$

Without loss of generality, let us assume that the test concerns the parameter β_1 . We can now calculate the minimum required sample size with the Wald test for the hypothesis $H_0 : \beta_1 = 0$, against $H_1 : \beta_1 \neq 0$. Analogous to the Poisson regression model, the estimates of the parameters $\gamma^{*T} = (\gamma_0^*, \dots, \gamma_q^*)$ and $(\beta_0^*, 0, \beta_2^*, \dots, \beta_p^*)$ under H_0 , are calculated by solving the modified system of equations:

$$\mathbf{E} [S_N(\gamma_0, \dots, \gamma_q, \beta_0, 0, \beta_2, \dots, \beta_p; \mathbf{v}_0) | \mathbf{X} = \mathbf{x}] = 0, \quad (14)$$

where the conditional expectation is taken with respect to the true values of γ and β , and the T_{ij} are replaced by their conditional means obtained from the EM algorithm used for maximum likelihood estimation of parameters in ZIP models with random effects (Hall 2000). We have

$$\begin{aligned}
E[T_{i,j}|y_{i,j}, \gamma, \beta] &= P(\text{excess zeros}|y_{i,j}, \gamma, \beta) \\
&= \frac{P(Y_{i,j}|\text{excess zeros})P(\text{excess zeros})}{P(Y_{i,j}|\text{excess zeros})P(\text{excess zeros}) + P(Y_{i,j}|\text{Poisson})P(\text{Poisson})} \quad (15) \\
&= \begin{cases} \left(1 + e^{-G_{i,j}\gamma - e^{\beta^T X_{i,j} + v_{0,i}}}\right)^{-1} & \text{if } y_{i,j} = 0 \\ 0 & \text{if } y_{i,j} > 0. \end{cases}
\end{aligned}$$

As Hall (2000) has shown that the maximum likelihood estimators of β are asymptotically normal, the analog of formula (9) can be used to estimate the minimum sample size N_s , which is given by:

$$N_s = \left\lceil \left(\frac{V(\gamma^*, \beta_0^*, 0, \beta_2^*, \dots, \beta_p^*, \mathbf{v}_0^*)^{1/2} Z_{\alpha/2} + V(\gamma, \beta, \mathbf{v}_0)^{1/2} Z_{1-\text{Power}}}{\beta_1} \right)^2 \right\rceil, \quad (16)$$

where the variance $V(\gamma, \beta, \mathbf{v}_0)$ is computed with the pre-specified values of the parameters and $V(\gamma^*, \beta_0^*, 0, \beta_2^*, \dots, \beta_p^*, \mathbf{v}_0^*)$ is calculated using the values obtained in (14). All the steps for the calculations of the sample size for a significance test with specified power and size on any parameter $\beta_s, s = 1, \dots, p$, are described in Algorithm 2.

Algorithm 2 Sample size calculations for the hierarchical Poisson-gamma model and the hierarchical ZIP-gamma model with one random effect (intercept).

The algorithm has the following input:

- The desired *Power* and the significance level α to be used.
- The value of the difference Δ that is to be detected. (This difference is set equal to e^{β_s} , where β_s is the parameter being tested whether or not it equals 0.)
- The distribution of the covariates \mathbf{X} and \mathbf{G} in Eqs. (3) and (10).
- The parameters $1/\alpha_0$, equal to the variance of $U_0 \sim \Gamma(\alpha_0, 1/\alpha_0)$.
- Some reasonable values for the other $\beta_k, k = 1, \dots, p, (k \neq s)$ and $\gamma_l, l = 1, \dots, q$, also have to be specified. (These could come from previous studies, a pilot study or an educated guess). The values of β_0 and γ_0 are chosen to satisfy an overall fixed means $\bar{\lambda} = E[e^{\beta^T \mathbf{X}}]$ and $\bar{\pi} = E[e^{\gamma^T \mathbf{G}} / (1 + e^{\gamma^T \mathbf{G}})]$, respectively.
- An initial number K of simulated sets of covariates. Empirical studies suggest that the value of K has little effect on the results and that $K = 100$ suffices.
- The number of clusters n is specified.
- The desired number of Monte Carlo replications B .

For normal covariates, the quasi-Monte Carlo method is used for integration, whereas analytic formula can be obtained for other distributions (e.g. Bernoulli or multinomial).

The algorithm then returns the minimum sample size needed for the significance test. It has the following steps:

1. **Generate** a sample of K values from the distributions of \mathbf{X} and \mathbf{G} , and compute the rates $\lambda_i(X)$, and $\pi_i(G)$, for $i = 1, \dots, K$, under H_1 .
2. **Compute** the values $\gamma^* = (\gamma_0^*, \gamma_1^*, \dots, \gamma_q^*)$ and
3. $\beta^* = (\beta_0^*, \beta_1^*, \dots, \beta_{s-1}^*, 0, \beta_{s+1}^*, \dots, \beta_p^*)$, and $u_{0,1}, \dots, u_{0,n}$, under H_0 , by solving the system of nonlinear equations in (14) with $N = K$, and using the relation in (15);
4. **Compute** the Fisher information matrix using (13) to obtain the variance functions $V(\gamma^*, \beta^*, v_0^*)$ and $V(\gamma, \beta, v_0)$ ($v_0 = \log(u_0)$), and then take its $(q + s + 2)$ -th diagonal elements;
5. **Repeat** steps 1 to 3 B times and each time return the sample size $N_{s,j}$ obtained from Eq. (16), for $j = 1, \dots, B$.
6. **Take** the average value over the outputs of the B replications, $N_s = \frac{\sum_{j=1}^B N_{s,j}}{B}$, for the required sample size.

4. Numerical examples

All our simulation studies are illustrated using the conjugate Poisson-gamma model and clusters of equal size with different distributions for the covariates. The required sample sizes for our numerical examples are computed through Monte Carlo simulations to test about the significance of β_1 . For each replication the values of the covariates are generated from the specified distributions, given we use the conditional expectation of the score function (Eqs. (6) and (14)). We consider first an example with a single covariate and then we extend this to the case of two covariates. It should also be noted that here the prespecified value of β_1 also represents the difference we want to detect by the test for a given power and size. We consider the cases of $X \sim \text{Bernoulli}(p)$ with $p = 0.1, 0.5, 0.9$, and $X \sim N(\mu, 1)$ with $\mu = 0, 0.5$. The case of two covariates, (X_1, X_2) , corresponding to the values $(0,0), (0,1), (1,0)$, and $(1,1)$, is considered as a multinomial with probabilities $(p_1, p_2, p_3, p_4 = 1 - p_1 - p_2 - p_3)$. The three following cases from Shieh (2001) are considered: $(p_1, p_2, p_3, p_4) = (0.25, 0.25, 0.25, 0.25)$, $(p_1, p_2, p_3, p_4) = (0.4, 0.1, 0.1, 0.4)$, and the extreme case $(p_1, p_2, p_3, p_4) = (0.76, 0.19, 0.01, 0.04)$. The “true” parameters are chosen as follows (Shieh 2001; Channouf, Fredette, and MacGibbon 2014): β_0 is chosen to satisfy $\bar{\lambda} = E[e^{\beta^T \mathbf{X}}] = 0.05$, $\beta_1 = \beta_2 = \log(2)$, and we arbitrarily take $\alpha_0 = 2$. We also consider equal size clusters, i.e. $n_i = N/n$, for $i = 1, \dots, n$, and fixed number of clusters n . The cluster size n_i is variable and can be increased to meet the calculated overall sample size.

For the hierarchical ZIP model, we use one additional covariate for the logit part of the model and the same types of covariate distributions for the counts as were used with the hierarchical Poisson model. We also add the case of two covariates having the bivariate normal distribution $N(0, \Sigma_\rho)$, with $\rho = -0.5, 0, 0.5$, where $\Sigma_\rho = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ is the covariance matrix. The “true” parameters of the logit model are taken as follows: γ_0 is chosen to satisfy $\bar{\pi} = E[\frac{e^{\gamma^T \mathbf{G}}}{1 + e^{\gamma^T \mathbf{G}}}] = 0.05$ and $\gamma_1 = \log 2$.

The estimates of the sample size N_s for testing $H_0 : \beta_1 = 0$, against $H_1 : \beta_1 \neq 0$, are based on independent Monte Carlo simulations with 5000 replications, using the software R. The significance level is taken to be $\alpha = 0.05$, and the powers considered are

Table 1. Results for the hierarchical Poisson regression model for testing $H_0 : \beta_1 = 0$, against $H_1 : \beta_1 \neq 0$, with significance level 5%, where β_1 corresponds to the effect of the odds ratio e^{β_1} that has to be detected. One covariate X has the Bernoulli distribution.

	Number of clusters	N_s	
		Power = 0.9	Power = 0.8
$Y u_0 \sim \text{Pois}(\lambda)$	1	7106	5578
$\lambda = e^{\beta_0 + \beta_1 X + \log(u_0)}$	2	5086	4015
$X \sim \text{Bernoulli}(0.1)$	4	4466	3496
$U_0 \sim \Gamma(2, 1/2)$	10	4163	3292
	25	4101	3210
	50	4069	3191
	100	4060	3176
Poisson	–	4046	3164
$Y u_0 \sim \text{Pois}(\lambda)$	1	3059	2397
$\lambda = e^{\beta_0 + \beta_1 X + \log(u_0)}$	2	2279	1702
$X \sim \text{Bernoulli}(0.5)$	4	2019	1503
$U_0 \sim \Gamma(2, 1/2)$	10	1900	1407
	25	1857	1374
	50	1847	1364
	100	1838	1362
Poisson	–	1823	1355
$Y u_0 \sim \text{Pois}(\lambda)$	1	10477	8036
$\lambda = e^{\beta_0 + \beta_1 X + \log(u_0)}$	2	7689	5538
$X \sim \text{Bernoulli}(0.9)$	4	6837	4855
$U_0 \sim \Gamma(2, 1/2)$	10	6482	4568
	25	6342	4478
	50	6300	4450
	100	6284	4429
Poisson	–	6271	4418

80% and 90%. We also calculated the intraclass correlation (ICC) as defined by Commenges and Jacqmin (1994) for the various hierarchical Poisson models defined in Tables 1 to 6. They defined the ICC as the ratio of the “variation within classes” to the “total variation”. However, each of these models considered here satisfies the properties of the general intraclass correlation random effects model as defined by Commenges and Jacqmin (1994). Thus, the

$$\text{ICC} = \text{corr}(Y_{i,j}, Y_{i,k}),$$

which only depends on i . It should be noted that our chosen $\alpha_0 = 2$ yields small intra-class correlations (ICC) varying between 1% and 2.5%.

The second column of each table contains the initial number of clusters, where the total number of observations generated from the original distributions of X (and G) is fixed at $K=100$ for all the cases. This is done since simulations showed almost no change if we used $K = 1,000$ or $K = 100$.

In each table, the first case corresponding to a unique cluster is an extreme case where all the outcomes are correlated and we expect to have an extremely large sample size here. As the number of clusters increases, the sample sizes obtained always converge toward the “independent” case of a nonhierarchical Poisson regression model and can thus be compared with the results of Shieh (2001) and Channouf, Fredette, and MacGibbon (2014). It is important to note that, to the best of our knowledge, no other existing studies with numerical results are presented in the literature for hierarchical Poisson regression models with more than two clusters, so we are thus unable to present a comparative study with the results of others.

Table 2. Results for the hierarchical Poisson regression model for testing $H_0 : \beta_1 = 0$, against $H_1 : \beta_1 \neq 0$, with significance level 5%, where β_1 corresponds to the effect of the odds ratio e^{β_1} that has to be detected. One covariate X has the standard normal distribution.

	Number of clusters	N_s	
		Power = 0.9	Power = 0.8
$Y u_0 \sim \text{Pois}(\lambda)$	1	734	587
$\lambda = e^{\beta_0 + \beta_1 X + \log(u_0)}$	2	541	415
$X \sim N(0,1)$	4	483	362
$U_0 \sim \Gamma(2, 1/2)$	10	454	341
	25	444	334
	50	443	331
	100	441	330
Poisson	–	437	326
$Y u_0 \sim \text{Pois}(\lambda)$	1	762	570
$\lambda = e^{\beta_0 + \beta_1 X + \log(u_0)}$	2	544	413
$X \sim N(0.5,1)$	4	477	361
$U_0 \sim \Gamma(2, 1/2)$	10	454	341
	25	445	332
	50	442	331
	100	441	330
Poisson	–	439	328

In Table 1, we present examples with one covariate X having a binomial distribution with $p = 0.1, 0.5$, and 0.9 . In all cases the sample size is fairly large when we have one cluster, especially when $p = 0.9$, and then the total sample size decreases when the number of clusters increases. This indicates the importance of using the hierarchical structure when the number of clusters is small or moderate even when the ICC is very close to 0. Changing power has an important impact on the cluster size. In general, between 27% and 42% more individuals are needed when the power increases from 80% to 90%, and the largest difference is observed when $p = 0.9$. When the covariate X follows a normal distribution (Table 2), we observe the same tendency when the number of clusters increases. Increasing power from 80% to 90%, the sample size can increase by an amount between 25% and 34%, and the shift in mean by 0.5 has no big effect on the sample size. For the numerical example with two covariates having the multinomial distribution, the results are quite similar to the model with a single covariate having a Bernoulli distribution in the well balanced case with $(p_1, p_2, p_3, p_4) = (0.25, 0.25, 0.25, 0.25)$. For the extremely skewed case, the sample size increases substantially for both levels of power. When we increase the power from 80% to 90%, the sample size increases from 32% to 35% for the balanced case, and from 29% to 36% for the unbalanced case. The results with the multinomial case are listed in Table 3. As the number of clusters increases, the total sample sizes approach those of Shieh (2001) and Channouf, Fredette, and MacGibbon (2014) for the non hierarchical model.

For the hierarchical Zero-inflated Poisson regression model, we consider the same distributions cited previously and we also add the bivariate normal distribution. The results are summarized in Tables 4–7. In general, when the power increases from 80% to 90%, the sample size increases by an amount ranging between 25% and 42%, and it is more pronounced with one covariate X following Bernoulli (0.9). Compared to the hierarchical Poisson model, calculation with the hierarchical ZIP model provides sample sizes which are between 4% and 14% larger. As the number of clusters increases, the sample sizes approach those of Channouf, Fredette, and MacGibbon (2014) obtained

Table 3. Results for the hierarchical Poisson regression model for testing $H_0 : \beta_1 = 0$, against $H_1 : \beta_1 \neq 0$, with significance level 5%, where β_1 corresponds to the effect of the odds ratio e^{β_1} that has to be detected. Two covariates X_1 and X_2 have the multinomial distribution.

	Number of clusters	N_s	
		Power = 0.9	Power = 0.8
$Y u_0 \sim \text{Pois}(\lambda)$	1	3154	2389
$\lambda = e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \log(u_0)}$	2	2251	1715
$(X_1, X_2) \sim \text{Multinomial}(p)$	4	2034	1530
$\rho = (0.25, 0.25, 0.25, 0.25)$	10	1893	1400
$u_0 \sim \Gamma(2, 1/2)$	25	1863	1376
	50	1847	1367
	100	1843	1362
Poisson	–	1835	1356
$Y u_0 \sim \text{Pois}(\lambda)$	1	4895	3784
$\lambda = e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \log(u_0)}$	2	3694	2715
$(X_1, X_2) \sim \text{Multinomial}(p)$	4	3301	2434
$\rho = (0.4, 0.1, 0.1, 0.4)$	10	3080	2268
$u_0 \sim \Gamma(2, 1/2)$	25	3018	2214
	50	2998	2197
	100	2989	2191
Poisson	–	2927	2184
$Y u_0 \sim \text{Pois}(\lambda)$	1	12104	11487
$\lambda = e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \log(u_0)}$	2	7184	5873
$(X_1, X_2) \sim \text{Multinomial}(p)$	4	6314	4911
$\rho = (0.76, 0.19, 0.01, 0.04)$	10	5903	4609
$u_0 \sim \Gamma(2, 1/2)$	25	5770	4514
	50	5757	4485
	100	5741	4457
Poisson	–	5661	4390

with the nonhierarchical ZIP regression model, with all the types of distributions of the covariates studied. The only exception is about the extremely skewed case when X follows a multinomial distribution with $(p_1, p_2, p_3, p_4) = (0.76, 0.19, 0.01, 0.04)$. The sample size decreases when we go from Poisson to ZIP model with a single cluster. The results for the normal case with two covariates exhibit larger sample sizes with positive correlation when compared to the results with the nonhierarchical ZIP regression model. When $\rho = -0.5$, the sample size is larger by approximately a percentage of 10%, and for $\rho = 0$, it is only 5% larger. When $\rho = 0.5$, the sample size is larger than results with the nonhierarchical ZIP regression model by approximately 75%. This dramatic increase in sample size must be taken into account when introducing random effects into a standard Poisson or ZIP models.

For both types of Poisson and ZIP models, and with all types of covariate distributions considered, the results indicated, as expected, that the required total sample size always decreases when the number of clusters increases. This means fewer correlated observations and therefore fewer individuals are needed.

5. Illustrative example

In order to illustrate our methodology with a real example, we reconsider here an application first presented in Signorini (1991), and then studied by Channouf, Fredette, and MacGibbon (2014) using the ZIP regression model. During a study of water pollution around Sydney Australia, the number of illnesses and infections contracted during a swimming season was examined by the Sydney Water Board. The objective was to

Table 4. Results for the hierarchical zero-inflated Poisson regression model for testing $H_0 : \beta_1 = 0$, against $H_1 : \beta_1 \neq 0$, with significance level 5%, where β_1 corresponds to the effect of the odds ratio e^{β_1} that has to be detected. One covariate G with Bernoulli distribution and one covariate X has the Bernoulli distribution.

	Number of clusters	N_s	
		Power = 0.9	Power = 0.8
$Y u_0 \sim \text{ZIP}(\lambda, \pi)$	1	7762	6225
$\lambda = e^{\beta_0 + \beta_1 X + \log(u_0)}$	2	5563	4423
$\pi = \frac{e^{\alpha_0 + \gamma_1 G}}{1 + e^{\alpha_0 + \gamma_1 G}}$	4	4893	3875
$X \sim \text{Bernoulli}(0.1)$	10	4586	3599
$G \sim \text{Bernoulli}(0.5)$	25	4474	3525
$U_0 \sim \Gamma(2, 1/2)$	50	4452	3496
	100	4423	3471
ZIP	–	4261	3331
Poisson	–	4040	3164
$Y u_0 \sim \text{ZIP}(\lambda, \pi)$	1	3513	2670
$\lambda = e^{\beta_0 + \beta_1 X + \log(u_0)}$	2	2499	1841
$\pi = \frac{e^{\alpha_0 + \gamma_1 G}}{1 + e^{\alpha_0 + \gamma_1 G}}$	4	2203	1643
$X \sim \text{Bernoulli}(0.5)$	10	2067	1542
$G \sim \text{Bernoulli}(0.5)$	25	2019	1501
$U_0 \sim \Gamma(2, 1/2)$	50	2004	1490
	100	1998	1483
ZIP	–	1933	1428
Poisson	–	1823	1355
$Y u_0 \sim \text{ZIP}(\lambda, \pi)$	1	11439	8325
$\lambda = e^{\beta_0 + \beta_1 X + \log(u_0)}$	2	8591	6041
$\pi = \frac{e^{\alpha_0 + \gamma_1 G}}{1 + e^{\alpha_0 + \gamma_1 G}}$	4	7470	5301
$X \sim \text{Bernoulli}(0.9)$	10	7048	5003
$G \sim \text{Bernoulli}(0.5)$	25	6867	4881
$U_0 \sim \Gamma(2, 1/2)$	50	6833	4847
	100	6804	4824
ZIP	–	6599	4650
Poisson	–	6271	4418

Table 5. Results for the hierarchical zero-inflated Poisson regression model for testing $H_0 : \beta_1 = 0$, against $H_1 : \beta_1 \neq 0$, with significance level 5%, where β_1 corresponds to the effect of the odds ratio e^{β_1} that has to be detected. One covariate G with Bernoulli distribution and one covariate X has the normal distribution.

	Number of clusters	N_s	
		Power = 0.9	Power = 0.8
$Y u_0 \sim \text{ZIP}(\lambda, \pi)$	1	826	641
$\lambda = e^{\beta_0 + \beta_1 X + \log(u_0)}$	2	596	457
$\pi = \frac{e^{\alpha_0 + \gamma_1 G}}{1 + e^{\alpha_0 + \gamma_1 G}}$	4	530	396
$X \sim N(0, 1)$	10	496	375
$G \sim \text{Bernoulli}(0.5)$	25	484	365
$U_0 \sim \Gamma(2, 1/2)$	50	481	361
	100	479	359
ZIP	–	453	339
Poisson	–	437	326
$Y u_0 \sim \text{ZIP}(\lambda, \pi)$	1	817	621
$\lambda = e^{\beta_0 + \beta_1 X + \log(u_0)}$	2	605	454
$\pi = \frac{e^{\alpha_0 + \gamma_1 G}}{1 + e^{\alpha_0 + \gamma_1 G}}$	4	530	400
$X \sim N(0.5, 1)$	10	497	374
$G \sim \text{Bernoulli}(0.5)$	25	484	365
$U_0 \sim \Gamma(2, 1/2)$	50	482	361
	100	479	360
ZIP	–	462	345
Poisson	–	439	328

Table 6. Results for the hierarchical zero-inflated Poisson regression model for testing $H_0 : \beta_1 = 0$, against $H_1 : \beta_1 \neq 0$, with significance level 5%, where β_1 corresponds to the effect of the odds ratio e^{β_1} that has to be detected. One covariate G with Bernoulli distribution and two covariates X_1 and X_2 have the multinomial distribution.

	Number of clusters	N_s	
		Power = 0.9	Power = 0.8
$Y u_0 \sim \text{ZIP}(\lambda, \pi)$	1	3345	2710
$\lambda = e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \log(u_0)}$	2	2489	1907
$\pi = \frac{e^{\gamma_0 + \gamma_1 G}}{1 + e^{\gamma_0 + \gamma_1 G}}$	4	2191	1651
$(X_1, X_2) \sim \text{Multinomial}(p)$	10	2059	1549
$p = (0.25, 0.25, 0.25, 0.25)$	25	2017	1500
$G \sim \text{Bernoulli}(0.5)$	50	2008	1492
$U_0 \sim \Gamma(2, 1/2)$	100	1998	1485
ZIP	–	1932	1428
Poisson	–	1835	1356
$Y u_0 \sim \text{ZIP}(\lambda, \pi)$	1	5469	4143
$\lambda = e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \log(u_0)}$	2	4038	3009
$\pi = \frac{e^{\gamma_0 + \gamma_1 G}}{1 + e^{\gamma_0 + \gamma_1 G}}$	4	3573	2631
$(X_1, X_2) \sim \text{Multinomial}(p)$	10	3336	2488
$p = (0.4, 0.1, 0.1, 0.4)$	25	3279	2416
$G \sim \text{Bernoulli}(0.5)$	50	3255	2399
$U_0 \sim \Gamma(2, 1/2)$	100	3249	2389
ZIP	–	3138	2297
Poisson	–	2927	2184
$Y u_0 \sim \text{ZIP}(\lambda, \pi)$	1	10943	8649
$\lambda = e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \log(u_0)}$	2	7770	6193
$\pi = \frac{e^{\gamma_0 + \gamma_1 G}}{1 + e^{\gamma_0 + \gamma_1 G}}$	4	6893	5438
$(X_1, X_2) \sim \text{Multinomial}(p)$	10	6474	5069
$p = (0.76, 0.19, 0.01, 0.04)$	25	6289	4953
$G \sim \text{Bernoulli}(0.5)$	50	6278	4886
$U_0 \sim \Gamma(2, 1/2)$	100	6250	4881
ZIP	–	5972	4633
Poisson	–	5661	4390

determine if there was a significant difference between non-ocean or infrequent swimmers ($X_1 = 0$) and ocean swimmers ($X_1 = 1$). Using a Poisson regression to model the number of illnesses and infections per swimmer, it was originally assumed that the infection rate per non-ocean or infrequent swimmer was 0.85 ($e^{\beta_0} = 0.85$) and that the sample size large enough to detect an increase of 30% for ocean swimmers ($\Delta = e^{\beta_1} = 1.3$) should be calculated for a given nominal power. It was also assumed that an equal number from both groups of swimmers would be sampled ($X_1 \sim \text{Bernoulli}(0.5)$), and that the proportion of each group of swimmers in each city is approximately 50%.

We used the method described here for hierarchical ZIP models for more flexibility in order to account for a possible excess of zeros among swimmers and we can assume that the swimmers were selected from 5 different cities (clusters). In this case, swimmers within the same city should be correlated and we examine the effect of the ICC on the sample size for different values of the dispersion parameter α_0 . At first, we consider that all swimmers have the same probability π of being measured from a zero-only random variable ($\pi = 0\%$ represents the hierarchical Poisson regression model). In a second step, we suppose that the practitioners would want to consider that both groups of swimmers

Table 7. Results for the hierarchical zero-inflated Poisson regression model for testing $H_0 : \beta_1 = 0$, against $H_1 : \beta_1 \neq 0$, with significance level 5%, where β_1 corresponds to the effect of the odds ratio e^{β_1} that has to be detected. One covariate G with Bernoulli distribution and two covariates X_1 and X_2 have the bivariate normal distribution.

	Number of clusters	N_s	
		Power = 0.9	Power = 0.8
$Y u_0 \sim \text{ZIP}(\lambda, \pi)$	1	1417	1075
$\lambda = e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \log(u_0)}$	2	1013	776
$\pi = \frac{e^{\gamma_0 + \gamma_1 G}}{1 + e^{\gamma_0 + \gamma_1 G}}$	4	907	683
$(X_1, X_2) \sim N(0, \Sigma_\rho)$	10	851	636
$\rho = -0.5$	25	830	626
$G \sim \text{Bernoulli}(0.5)$	50	824	619
$U_0 \sim \Gamma(2, 1/2)$	100	820	616
ZIP	–	746	558
Poisson	–	445	333
$Y u_0 \sim \text{ZIP}(\lambda, \pi)$	1	1297	1016
$\lambda = e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \log(u_0)}$	2	963	741
$\pi = \frac{e^{\gamma_0 + \gamma_1 G}}{1 + e^{\gamma_0 + \gamma_1 G}}$	4	851	647
$(X_1, X_2) \sim N(0, \Sigma_\rho)$	10	809	612
$\rho = 0$	25	792	595
$G \sim \text{Bernoulli}(0.5)$	50	785	591
$U_0 \sim \Gamma(2, 1/2)$	100	784	590
ZIP	–	747	561
Poisson	–	440	329
$Y u_0 \sim \text{ZIP}(\lambda, \pi)$	1	2103	1605
$\lambda = e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \log(u_0)}$	2	1587	1209
$\pi = \frac{e^{\gamma_0 + \gamma_1 G}}{1 + e^{\gamma_0 + \gamma_1 G}}$	4	1429	1086
$(X_1, X_2) \sim N(0, \Sigma_\rho)$	10	1344	1024
$\rho = 0.5$	25	1315	996
$G \sim \text{Bernoulli}(0.5)$	50	1313	988
$U_0 \sim \Gamma(2, 1/2)$	100	1308	987
ZIP	–	751	561
Poisson	–	423	315

Table 8. Results for testing $H_0: \beta_1 = 0$, against $H_1: \beta_1 \neq 0$, with significance level 5%. Hierarchical ZIP regression model, with $X \sim \text{Bernoulli}(0.5)$ ($e^{\beta_0} = 0.85$ and $e^{\beta_1} = 1.3$), $G \sim \text{Bernoulli}(0.5)$ and 5 clusters. Two cases: one case with a single probability of an excess zero ($\pi^0 = \pi^1 = \pi$), and a second case with two different probabilities of an excess zero for the two groups (π^0 and π^1).

	Power	$\pi^0 = \pi^1 = \pi$			$\pi^0 = 0.1$		
		$\pi = 0.05$	$\pi = 0.1$	$\pi = 0.25$	$\pi^1 = 0.05$	$\pi^1 = 0.15$	$\pi^1 = 0.2$
$\alpha_0 = 2$	0.8	573	639	906	603	678	713
	0.9	765	854	1214	804	903	948
	0.95	949	1060	1496	996	1116	1178
	ICC	0.26	0.26	0.28	0.26	0.26	0.27
$\alpha_0 = 5$	0.8	537	595	843	566	632	663
	0.9	719	798	1124	757	843	887
	0.95	888	991	1380	935	1038	1090
	ICC	0.15	0.15	0.16	0.15	0.15	0.15
$\alpha_0 = 10$	0.8	526	582	815	552	615	645
	0.9	705	788	1081	740	818	858
	0.95	869	959	1327	914	1009	1062
	ICC	0.08	0.09	0.09	0.09	0.09	0.09
$\alpha_0 = 100$	0.8	514	565	776	538	592	622
	0.9	686	750	1016	717	786	823
	0.95	847	923	1239	884	966	1010
	ICC	0.01	0.01	0.01	0.01	0.01	0.01

could have a different probability of being measured from a zero-only random variable. We have:

$$\pi_i = \begin{cases} \pi^0 & \text{if } G_{1,i} = 0, \\ \pi^1 & \text{if } G_{1,i} = 1. \end{cases}$$

We also have $G_{1,i} = X_{1,i}$ and the parameters γ_0 and γ_1 are chosen such that $\pi^0 = .10$ and $\pi^1 = .05, .15$ or $.20$.

Table 8 shows the required sample sizes for hierarchical ZIP regression for nominal powers 80%, 90% and 95%. with the two cases for π . For the single probability ZIP model as π increases from 0.05 to 0.25, the necessary sample size increases by 56%-58% for $\alpha_0 = 2$. For $\alpha_0 = 100$, this increase is of the order of 46%-51%. In the more complex model given by Eq. (17) with $\pi^0 = 0.1$ and π^1 varying between 0.05 and 0.2, the sample size increase varies between 14% and 17% for all α_0 considered. The necessary sample sizes increase as π increases in the simple model and as π^1 increases in the more complex model. Increasing the power while keeping the other parameters fixed increases the necessary sample size. It should be noted that the necessary sample size for $\pi = \pi^1 = 0.05$ are larger for the complex model.

6. Conclusion

We have succeeded in developing Monte Carlo simulation methods for calculating the sample size necessary for the Wald test of one parameter for hierarchical Poisson models using the estimation procedure proposed by Lee and Nelder (1996, 2001) for the parameters, where the power and the significance levels, as well as the values of the parameters of interest are specified under the null and the alternative hypothesis. We show that these methods can be extended also to more complex models, such as Zero-inflated hierarchical Poisson models. Our simulations indicate that the sample size will be sensitive to the type of covariate distributions, as well as the size of the intraclass correlations.

Our results indicate that caution should be taken when adding random effects to Poisson regression models, because the required sample size for the same size and power can be substantially larger when the number of clusters is small or moderate.

Further work should include the case of unequal sizes for the clusters which will result in more complex calculations. For these to be efficient, we will use quasi-Monte Carlo methods (Bratley and Fox 1988). Further research should also study possibly inflated type 1 error due to small number of clusters and if correction is needed, as well as the impact of the design effect (DE) on the sample size and if an adjustment is needed in addition to the effect of the intra class correlation (ICC). Deep investigations should also be done about the mathematical proprieties and the asymptotic behavior of the estimators when the cluster size is relatively small. This methodology could also be applied to other hierarchical generalized linear models. It would be interesting for example to consider the case of overdispersed count data where the negative binomial distribution is used, which is very common in medical studies.

Acknowledgments

This work was supported by the NSERC Discovery Grants of the second and third authors. All three authors would like to thank the referees for their helpful comments which improved the original manuscript.

References

- Bratley, P., and B. L. Fox. 1988. Algorithm 659: Implementing Sobol's quasirandom sequence generator. *ACM Transactions on Mathematical Software* 14 (1):88–100. doi:[10.1145/42288.214372](https://doi.org/10.1145/42288.214372).
- Channouf, N., M. Fredette, and B. MacGibbon. 2014. Power and sample size calculations for Poisson and zero-inflated Poisson regression models. *Computational Statistics & Data Analysis* 72 :241–51. doi:[10.1016/j.csda.2013.09.029](https://doi.org/10.1016/j.csda.2013.09.029).
- Christiansen, C. L., and C. N. Morris. 1997. Hierarchical Poisson regression modeling. *Journal of the American Statistical Association* 92 (438):618–32. doi:[10.2307/2965709](https://doi.org/10.2307/2965709).
- Commenges, D., and H. Jacqmin. 1994. The intraclass correlation coefficient: distribution-free definition and test. *Biometrics* 50 (2):517–26. doi:[10.2307/2533395](https://doi.org/10.2307/2533395).
- Ha, I. D., and Y. Lee. 2003. Estimating frailty models via Poisson hierarchical generalized linear models. *Journal of Computational and Graphical Statistics* 12 (3):663–81. doi:[10.1198/1061860032256](https://doi.org/10.1198/1061860032256).
- Hall, D. B. 2000. Zero-inflated Poisson and binomial regression with random effects: a case study. *Biometrics* 56 (4):1030–9. doi:[10.1111/j.0006-341X.2000.01030.x](https://doi.org/10.1111/j.0006-341X.2000.01030.x).
- Jiang, J. 2007. *Linear and generalized linear mixed models and their applications*. Springer Series in Statistics. New York: Springer.
- Kuk, A. Y. C., and Y. W. Cheng. 1999. Pointwise and functional approximations in Monte Carlo maximum likelihood estimation. *Statistics and Computing* 9 (2):91–9.
- Lambert, D. 1992. Zero-inflated poisson regression, with an application to defects in manufacturing. *Technometrics* 34 (1):1–14. doi:[10.2307/1269547](https://doi.org/10.2307/1269547).
- Lee, Y., and J. A. Nelder. 1996. Hierarchical generalized linear models. *Journal of the Royal Statistical Society. Series B (Methodological)* 58 (4):619–78. doi:[10.1111/j.2517-6161.1996.tb02105.x](https://doi.org/10.1111/j.2517-6161.1996.tb02105.x).
- Lee, Y., and J. A. Nelder. 2001. Hierarchical generalized linear models: A synthesis of generalised linear models, random-effect models and structured dispersions. *Biometrika* 88 (4):987–1006. doi:[10.1093/biomet/88.4.987](https://doi.org/10.1093/biomet/88.4.987).
- Lee, Y., J. A. Nelder, and M. Noh. 2007. H-likelihood: problems and solutions. *Statistics and Computing* 17 (1):49–55.
- Liang, K. Y., and S. Zeger. 1986. Longitudinal data analysis using generalized linear models. *Biometrika* 73 :13–22. doi:[10.2307/2336267](https://doi.org/10.2307/2336267).
- Liu, G., and K.-Y. Liang. 1997. Sample size calculations for studies with correlated observations. *Biometrics* 53 (3):937–47. doi:[10.2307/2533554](https://doi.org/10.2307/2533554).
- Matsui, S. 2005. Sample size calculations for comparative clinical trials with over-dispersed Poisson process data. *Statistics in Medicine* 24 (9):1339–56.
- Roy, A., D. K. Bhaumik, S. Aryal, and R. D. Gibbons. 2007. Sample size determination for hierarchical longitudinal designs with differential attrition rates. *Biometrics* 63 (3):699–707. doi:[10.1111/j.1541-0420.2007.00769.x](https://doi.org/10.1111/j.1541-0420.2007.00769.x).
- Royle, J. A. 2008. Hierarchical modeling of cluster size in wildlife surveys. *Journal of Agricultural, Biological, and Environmental Statistics* 13 (1):23–36. doi:[10.1198/108571108X273188](https://doi.org/10.1198/108571108X273188).
- Self, S., and R. H. Mauritsen. 1988. Power/sample size calculations for generalized linear models. *Biometrics* 44 (1):79–86. doi:[10.2307/2531897](https://doi.org/10.2307/2531897).
- Self, S., R. H. Mauritsen, and J. Ohara. 1992. Power calculations for likelihood ratio tests in generalized linear models. *Biometrics* 48 (1):31–9. doi:[10.2307/2532736](https://doi.org/10.2307/2532736).

- Shieh, G. 2000. On power and sample size calculations for likelihood ratio tests in generalized linear models. *Biometrics* 56 (4):1192–6. doi:[10.1111/j.0006-341X.2000.01192.x](https://doi.org/10.1111/j.0006-341X.2000.01192.x).
- Shieh, G. 2001. Sample size calculations for logistic and Poisson regression models. *Biometrika* 88 (4):1193–9. doi:[10.1093/biomet/88.4.1193](https://doi.org/10.1093/biomet/88.4.1193).
- Shieh, G. 2005. On power and sample size calculations for Wald tests in generalized linear models. *Journal of Statistical Planning and Inference* 128 (1):43–59. doi:[10.1016/j.jspi.2003.09.017](https://doi.org/10.1016/j.jspi.2003.09.017).
- Signorini, D. F. 1991. Sample size for poisson regression. *Biometrika* 78 (2):446–50. doi:[10.2307/2337274](https://doi.org/10.2307/2337274).
- Song, P. X.-K. 2007. *Correlated data analysis*. Springer Series in Statistics. New York: Springer.
- Tu, W., and W. W. Piegorsch. 2003. Empirical Bayes analysis for a hierarchical Poisson generalized linear model. *Journal of Statistical Planning and Inference* 111 (1-2):235–48. doi:[10.1016/S0378-3758\(02\)00305-1](https://doi.org/10.1016/S0378-3758(02)00305-1).
- Williamson, J. M., H.-M. Lin, R. H. Lyles, and A. W. Hightower. 2007. Power calculations for ZIP and ZINB models. *Journal of Data Science* 5 :519–34.
- Xia, N., Z.-Z. Zhang, and Z.-L. Ying. 2006. Convergence rate of the L-N estimator in Poisson-gamma models. *Acta Mathematicae Applicatae Sinica, English Series* 22 (4):639–54. doi:[10.1007/s10255-006-0338-z](https://doi.org/10.1007/s10255-006-0338-z).
- Yau, K. K. W., and A. H. Lee. 2001. Zero-inflated Poisson regression with random effects to evaluate an occupational injury prevention programme. *Statistics in Medicine* 20 (19):2907–20. doi:[10.1002/sim.860](https://doi.org/10.1002/sim.860).