

C

Year

1 S
3 I

1 V
2 F
3 R

T

S
V

O
a

O
e
5c

d

CS

vi

ic

rv

di

Y

7

Parameter Inference for a Zero-Inflated Poisson Distribution and Its Variants

Sanku Debbar, Prabodh Kumar, Sujay Dasgupta^{2*} and Basudip A. Pal

August 10, 2009

¹ University of Calicut, Ponnampalayam, Kerala, India, ² American, Edinburgh, ³ Fred Hutchinson Cancer Research Center, Seattle, ⁴ Theoretical Network Centric System, * Corresponding author

1 Introduction

Count data are frequently encountered in real life and are often modelled by a Poisson distribution. However, it is well known that the number of zeroes in a dataset is often more than what is allowed by a Poisson model. So, in order to accommodate excess zeroes, a zero-inflated Poisson distribution (ZIP) was introduced (Gilchrist, 1965). Since then, many variants of ZIP have been proposed (e.g., Cohen (1966), Singh (1966), Gupta (1977), Gupta (1986) and Heilborn (1986)). Some interesting applications can be found in Heilborn and Gilchrist (1990), Gupta (1992), Gupta (1999), Saei and McGilchrist (1997), Li et al. (1999). There have been attempts to generalize the ZIP distribution in the past, such as the one in Consul and Jain (1973), and generalized versions of ZIP (Angers and Biswas (2003) proposed a generalized ZIP model. The ZIP model has the feature that its mean and variance are equal, i.e., the variances exceed their means. This necessitates the use of a more dispersed Poisson (OP) model. See Cox (1981) and Scollnik (1981). Overdispersed Poisson models have been used for modelling diseases (e.g., Gupta and Gupta (1990)) and one can show that a ZIP model is a special case of the CMP distribution (Gupta (2005) reviewed another variant of the CMP distribution, the CMP distribution, and earlier authors and demonstrated its usefulness as a model for count data. Later, Gupta et al. (2006) provided a Bayesian approach to the CMP distribution using conjugate priors. The CMP distribution can be extended to a ZIP distribution.

Here, we present parameter inference results for a ZIP distribution and its variants. In the above, we discuss point estimation, hypothesis testing, and Bayesian inference. Much of the existing work cited above came as

developments and are somewhat scattered. There is no such unified presentation of infinite distributions. However, this is not meant to detract from the results presented here, which are new (except for some references cited). Detailed proofs and derivations are given. We believe this article will serve as a valuable applied research on this interesting family. The paper is as follows. In section 2, we provide the distributions. In section 3, we derive the conditions of sum of ZIP variables. Section 4, fitting for ZIP data. Section 5 provides some new way-Maxwell-Poisson (CMP) distribution, the zero-inflated version of CMP, generalized Poisson distribution and generalized Poisson inference for the ZIP model. Section 6, while Appendix 2 describes the methodology outlined in section 4.

Preliminaries

The basic idea behind a ZIP distribution is to assign a higher probability to zero and lower probability to the corresponding Poisson distribution with parameters ϕ and λ . The ZIP distribution with two parameters is formally defined by

$$P(X=0) = \phi$$

$$P(X=x) = (1-\phi) \frac{e^{-\lambda} \lambda^x}{x!} \quad x=1,2,\dots$$

where ϕ is a number between 0 and 1 and λ is a positive real number.

$$P(X=x) = \phi I(x=0) + (1-\phi) \frac{e^{-\lambda} \lambda^x}{x!} \quad x=0,1,2,\dots$$

any non-negative integer x , where $I(A)$ is the indicator function of the event A .

The above definition is the most standard. However, there are several other ways to define a ZIP distribution. We shall discuss some of them here.

The first alternative definition is:

$$P(X=0) = e^{-\lambda}$$

$$P(X=x) = \left(1 - \frac{1}{1 + \lambda^x}\right) \frac{e^{-\lambda} \lambda^x}{x!} \quad x=1,2,\dots$$

To our knowledge, this is the first definition for this class of distributions. Most of the definitions are quoted from the literature wherever possible. The layout of the paper is as follows.

The layout of the paper is as follows. In section 2, we provide the distributions. In section 3, we derive the conditions of sum of ZIP variables. Section 4, fitting for ZIP data. Section 5 provides some new way-Maxwell-Poisson (CMP) distribution, the zero-inflated version of CMP, generalized Poisson distribution and generalized Poisson inference for the ZIP model. Section 6, while Appendix 2 describes the methodology outlined in section 4.

and λ is to assign a higher probability to zero and lower probability to the corresponding Poisson distribution with parameters ϕ and λ .

so write this as:

(1)

function of the event

ing a ZIP distribution. We shall discuss some of them here.

So, once again, this ZIP distribution is defined by the following simple-minded intuition: increase the probability at 0 by a constant number γ . One has to keep in mind that γ should be less than $1 - e^{-\lambda}$. A second way of defining the ZIP distribution is as follows: let δ and λ be two positive constants such that $\delta < 1 - e^{-\lambda}$. Then the ZIP distribution is defined by the following probability mass function:

$$P(X = 0) = (1 + \delta) e^{-\lambda}$$

$$P(X = x) = \left(1 + \delta - \frac{\delta}{1 - e^{-\lambda}}\right) e^{-\lambda} \frac{\lambda^x}{x!} \quad x > 0$$

This time, the two parameters are δ and λ . One can show that δ is nothing but $\frac{\phi(1 - e^{-\lambda})}{e^{-\lambda}}$. This definition is guided by the following intuition: increase the probability at 0 by multiplying the original probability by a number bigger than 1. Of course, one has to keep in mind that δ should be less than $e^{\lambda} - 1$. A third alternative definition of the ZIP distribution is:

$$P(X = 0) = 1 - \phi$$

$$P(X = x) = e^{-\mu} \frac{\mu^x}{x!} \quad \text{for } x \geq 1,$$

with $0 < \phi < 1$. Here the two parameters are ϕ and μ . One can easily observe that the two parameters of the original ZIP distribution, (λ, ϕ) , can be recovered from (ϕ, μ) as: $\lambda = \mu\phi$, $\phi = 1 - e^{-\mu + \mu\phi}$.

The mean and variance of ZIPD are given by:

$$E(X) = (1 - \phi)\lambda, \quad \text{Var}(X) = (1 - \phi)\lambda(1 + \phi\lambda)$$

The moment generating function $M_X(t)$ and the characteristic function $\Phi_X(t)$ are given by

$$M_X(t) = \phi + (1 - \phi)e^{-\lambda + \lambda e^t}$$

$$\Phi_X(t) = e^{i\lambda(e^{it} - 1)} \phi + (1 - \phi)e^{-\lambda + \lambda e^{it}}$$

Now it should be clear why a ZIP distribution is a special case of an overdispersed Poisson model. Next we look at a weighted ZIP distribution. It is given by

$$P_w(Y^w = y) = \frac{w(y)P(Y = y)}{E[w(Y)]},$$

where Y is a usual Poisson random variable. So, we want to consider the ZIP distribution as a weighted version of a Poisson distribution. On then,

$$\frac{w(y)}{E[w(Y)]} = \frac{1 - \phi}{1 - \phi + \phi e^{-\lambda}} e^{-\lambda} \frac{\lambda^y}{y!}$$

for $y > 0$ and

$$\frac{w(0)}{E[w(Y)]} = 1 + \delta,$$

where $\delta = \frac{\phi(1-e^{-\lambda})}{e^{-\lambda}}$. As a result, denoting $E[w(Y)]$ by A , we have $w(0) = A(1+\delta)$ and $w(y) = A(1-\phi)$ for $y > 0$. One observes that A can take any positive value. So, it is good enough to choose $w(0) = 1 + \delta$ and $w(y) = 1 - \phi$ for $y > 0$ where $\delta = \frac{(1-e^{-\lambda})}{e^{-\lambda}}$. Therefore, the ZIP distribution is nothing but a weighted overdispersed Poisson distribution with the weights $w(y)$ chosen as above for each $y \geq 0$.

3 Conditional and Unconditional Distributions of Sums of ZIP Variables

In order to discuss parametric inference for a ZIP distribution, one has to know the distribution of the convolution of i.i.d. ZIP variables. For this, we first observe that if $X_1, \dots, X_n$ are independent ZIP variables with X_i having parameters (ϕ_i, λ_i) for $i = 1, 2, \dots, n$ and if $Z = X_1 + \dots + X_n$, then using induction on n ,

$$P(Z = z) = \prod_{i=1}^n \phi_i I(z_i = 0) + \sum_{r=1}^n \sum_{(i_1, \dots, i_r)} \prod_{l=1}^r (1 - \phi_{i_l}) \prod_{j=1, j \neq i_1, \dots, i_r}^n \alpha_j e^{-\sum \lambda_{i_l}} \frac{(\sum \lambda_{i_l})^z}{z!}.$$

In case the X_i 's are i.i.d. with common parameters (ϕ, λ) , the convolution distribution reduces to

$$P(Z = z) = \phi^n I(z = 0) + \sum_{r=1}^n \binom{n}{r} (1 - \phi)^r \phi^{n-r} e^{-r\lambda} \frac{(r\lambda)^z}{z!}.$$

This can also be written as

$$P(Z = z) = \phi^n I(z = 0) + \sum_{r=1}^n P(Y_n = r) P(Z_r = z), \quad (2)$$

where Y_n is Binomial with parameters (n, ϕ) and Z_r is Poisson with parameter $r\lambda$. This can also be derived using a different method. One can start out by finding the conditional joint distribution of $X_1, X_2, \dots, X_n$ given $n_0 = j$ and the marginal distribution of n_0 , where n_0 is the number of zero values among the X_i 's. It can be shown that the conditional joint distribution of $X_1, \dots, X_n$ given $n_0 = j$ is

$$P(X_1 = x_1, \dots, X_n = x_n | n_0 = j) = \binom{n}{j}^{-1} \left(\frac{e^{-\lambda}}{1 - e^{-\lambda}} \right)^{n-j} \frac{\lambda^{\sum x_i}}{\prod_{i=1}^n x_i!}.$$

Next, for $j = 0$, one observes that the conditional joint distribution of $X_1, X_2, \dots, X_n$ given $n_0 = 0$ is the same as the unconditional joint distribution of n i.i.d. truncated Poisson random variables $X_1^*, X_2^*, \dots, X_n^*$ (truncated at zero), which is given by

$$P(X_1 = x_1, \dots, X_n = x_n | n_0 = 0) = \prod_{i=1}^n \frac{e^{-\lambda} \lambda^{x_i}}{(e^{-\lambda})^{x_i} x_i!}.$$

Note that this is free from ϕ . From this, we can derive the conditional pmf of $Z = \sum X_i$ given $n_0 = 0$ to be

$$P(\sum X_i = k | n_0 = 0) = \left(\frac{\xi_n(k)}{n^k} \right) (1 - e^{-\lambda}) P(Z_n = k),$$

where $Z_n \sim \text{Poisson}(n\lambda)$ and $\xi_n(k)$ is a known, parameter-free function. For $n = 2$, $\xi_2(k) = 2^k - 2$; for $n = 3$, $\xi_3(k) = 3^k - 3 \cdot 2^k + 3$, and so forth. In general there is a recursive formula for $\xi_n(k)$ as follows:

$$\xi_n(k) = \sum_{l=n-1}^{k-1} \binom{k}{l} \xi_{n-1}(l).$$

Next, it can be shown that the conditional pmf of $Z = \sum_{i=1}^n X_i$ given $n_0 = k$ is

$$P(\sum_{i=1}^n X_i = a | n_0 = k) = \frac{\xi_{n-k}(a)}{(n-k)^a} (1 - e^{-\lambda})^{-(n-k)} P(W = a),$$

where W has a usual Poisson distribution with parameter $(n-k)\lambda$. This is exactly the same as $P(\sum_{i=1}^{n-k} X_i = a | n_0 = 0)$. Finally, one easily observes that n_0 is Binomial with parameters n and $\phi = (1-\phi)e^{-\lambda}$ and multiplying this with the conditional pmf of $Z = \sum_{i=1}^n X_i$ given $n_0 = k$ yields the unconditional distribution of Z mentioned earlier.

We now move on to conditional distributions. We start with the conditional pmf of X_1 given $X_1 + X_2$. A little algebra will reveal that

$$P(X_1 = j | X_1 + X_2 = k) = \frac{a+b}{c+d}$$

where

$$c = 2\phi(1-\phi)e^{-\lambda} \frac{\lambda^k}{k!}, \quad d = (1-\phi)^2 e^{-2\lambda} \frac{(2\lambda)^k}{k!}$$

and either $a = 0, b = \binom{k}{j} \left(\frac{1}{2}\right)^k d$ or $a = \frac{c}{2}, b = \left(\frac{1}{2}\right)^k d$. In general, the conditional pmf of $\sum_{i=1}^m X_i$ given $\sum_{i=1}^n X_i$ (for $m \leq n$) involves both the parameters and can not be expressed by such a simple formula. This is unlike the ordinary Poisson distribution where the conditional pmf of $\sum_{i=1}^m X_i$ given

$\sum_{i=1}^n X_i = k$ (for $m \leq n$) is Binomial $(k, \frac{m}{n})$. But, in the case of ZIP variables, if we consider the conditional pmf of $\sum_{i=1}^m X_i$ given $\sum_{i=1}^n X_i = k$ (for $m \leq n$) and $n_0 = k^*$, then it turns out to be parameter-free, although not binomial. For $k^* = 0$, it is actually symmetric. For example,

$$P(X_1 = 1 | X_1 + X_2 = k, n_0 = 0) = 1;$$

$$P(X_1 = 1 | X_1 + X_2 = 3, n_0 = 0) = P(X_1 = 2 | X_1 + X_2 = 3, n_0 = 0) = \frac{1}{2};$$

$$P(X_1 = 1 | X_1 + X_2 + X_3 = 2, n_0 = 1) = \frac{1}{3}, P(X_1 = 1 | X_1 + X_2 + X_3 = 2, n_0 = 1) = \frac{2}{3};$$

etc. The general formula for $P(X_1 = l | X_1 + X_2 = k, n_0 = 0)$ is given by

$$P(X_1 = l | X_1 + X_2 = k, n_0 = 0) = \binom{k}{l} \xi_2^{-1}(k)$$

Similarly

$$P(X_1 = l | X_1 + X_2 + X_3 = k, n_0 = 0) = \sum_{m=1}^{k-l-1} \frac{1}{l!m!} \frac{k!}{(k-l-m)!} \xi_3^{-1}(k)$$

$$P(X_1 = l, X_2 = l | X_1 + X_2 + X_3 = k, n_0 = 0) = \sum_{m=1}^{l-1} \frac{1}{m!} \frac{k!}{(l-m)!(k-l)!} \xi_3^{-1}(k)$$

In general, for $m < n$,

$$P\left(\sum_{i=1}^m X_i = l \mid \sum_{i=1}^n X_i = k, n_0 = 0\right) = \sum_{l_1} \cdots \sum_{l_m} \sum_{j_1} \cdots \sum_{j_{n-m}} \frac{k!}{l_1! \cdots l_m! j_1! \cdots j_{n-m}!} \xi_n^{-1}(k)$$

where $\sum_{i=1}^m l_i = l$ and $\sum_{i=1}^{n-m} j_i = k - l$.

4 Parametric Inference for a ZIP Distribution

As mentioned earlier, the mean and the variance of a ZIP distribution are given by

$$E(X) = (1 - \phi)\lambda, \quad V(X) = (1 - \phi)\lambda(\phi + \phi\lambda)$$

So, if $X_1, X_2, \dots, X_n$ is a random sample from a ZIP distribution with parameters (ϕ, λ) , then $1 - \frac{\bar{X}}{\lambda}$ is an unbiased estimator for ϕ if λ is known, whereas $\frac{\bar{X}}{1 - \phi}$ is an unbiased estimator of λ if ϕ is known. But, in general, both ϕ and λ will be unknown and there is no sample-based unbiased estimator for the parameter-vector (ϕ, λ) . However, it is easy to obtain the method of moments (MOM) estimators which are

$$\hat{\lambda}_{MOM} = \frac{\sum X_i^2}{\sum X_i} - 1, \quad \hat{\phi}_{MOM} = 1 - \frac{n\bar{X}^2}{\sum X_i^2 - \sum X_i}$$

The sample likelihood function is given by

$$L(\phi, \lambda | x_1, \dots, x_n) = (\phi + (1 - \phi)e^{-\lambda})^{n_0} ((1 - \phi)e^{-\lambda})^{n - n_0} \frac{\lambda^{\sum x_i}}{\prod x_i!}$$

where n_0 is once again the number of zeros in the sample. From this, it is easy to see that the maximum likelihood estimators for the parameters can be obtained by solving

$$\hat{\lambda}_{MLE} = \frac{\bar{X}}{\hat{\phi}_{MLE}}, \quad \hat{\phi}_{MLE} = \frac{(n_0/n) - e^{-\hat{\lambda}_{MLE}}}{1 - e^{-\hat{\lambda}_{MLE}}}$$

For this pair of equations, no analytical solution exists and it must be solved numerically, which involves the Lambert's W function (see, for example, Corbridge et al. (1993)). In Appendix 2, we report the results from simulation studies conducted to explore the bias and the variance of the estimators mentioned above as functions of the two parameters. From the likelihood function, it should be clear that the 2-dimensional vector $(n_0, \sum X_i)$ is jointly sufficient for the two parameters (ϕ, λ) . In fact, it can be shown that they are minimally sufficient.

Now suppose that $X_1, \dots, X_n$ are i.i.d. following a $\text{ZIP}(\phi_1, \lambda_1)$ distribution, $Y_1, \dots, Y_n$ are i.i.d. following a $\text{ZIP}(\phi_2, \lambda_2)$ distribution and we want to test $H_0: (\phi_1, \lambda_1) = (\phi_2, \lambda_2)$ versus $H_1: \neq$. Let n_0 and m_0 be the counts of zero values in the two samples. Under H_0 , $n_0 + m_0$ is a Binomial($2n, \phi$) pmf where ϕ is the common value of ϕ_1 and ϕ_2 . If $n_0 + m_0$ is observed to be k and $\sum_{i=1}^n (X_i + Y_i)$ is observed to be k^* , we reject H_0 at the preassigned level α provided that the test-statistic $\sum_{i=1}^n X_i$ exceeds the $(1 - \alpha)^{th}$ percentile of, or falls below the α^{th} percentile of the conditional pmf of $\sum_{i=1}^n X_i$ given $\sum_{i=1}^n (X_i + Y_i) = k^*$ and $n_0 + m_0 = k$. Recall from the previous section that this conditional distribution is parameter-free. In order to achieve an exact significance level of α , it may be necessary to randomize this test. The statistics $\sum_{i=1}^n (X_i + Y_i)$ and $n_0 + m_0$ together induce a partitioning of the sample space and here we are conditioning on those partition cells. If the conditional test is performed at level α , the unconditional significance level is α as well, since $E(\text{Type I error} | H_0)$ is equal to

$$\sum_{k=0}^{2n} \sum_{k^*=0}^{\infty} P(\text{Type I error} | n_0, n_0 + m_0 = k, \sum_{i=1}^n (X_i + Y_i) = k^*) \cdot P(n_0 + m_0 = k, \sum_{i=1}^n (X_i + Y_i) = k^*)$$

This test is, in some sense, a generalization of the one introduced by Pruszycki and Wilenski (1993) for the equality of two Poisson means.

5 The Conway-Maxwell-Poisson Distribution

The CMP distribution was introduced by Conway and Maxwell (1962) and is a generalization of the Poisson distribution under certain conditions. Here the probability mass function is given by

$$p(x) = \frac{\lambda^x}{(x!)^\zeta} \sum_{k=0}^{\infty} \frac{\lambda^k}{(k!)^\zeta}$$

for $x = 0, 1, 2, \dots$ and λ and ζ are positive. For $\zeta = 1$, the probability mass function reduces to the Poisson distribution with parameter λ . For $\zeta > 1$, the probability at 0 is bigger than $e^{-\lambda}$, which may lead to believe that the CMP distribution with $\zeta > 1$ is an example of a usual ZIP distribution. This, however, is not the case because for values of $x > 0$, the variance is smaller than the mean, unlike the constant variance of a ZIP distribution. We still discuss it here as a generalization of the usual Poisson distribution, which will be discussed later. We begin our discussion with the distribution of the convolution of n independent CMP(λ, ζ) random variables, which is given by

$$P\left(\sum_{i=1}^n X_i = k\right) = \sum_{j=0}^k \frac{\lambda^j}{(j!)^\zeta} \sum_{i_1+\dots+i_n=j} \left(\frac{k!}{i_1! \dots i_n!} \right)$$

for $k = 0, 1, 2, \dots$ etc. This can be proved by induction on n . The mean and the variance can be calculated using the method-of-moments (Lukacs, 1950). It may be worth pointing out that for $\zeta > 1$, the CMP distribution is underdispersed, as opposed to the ZIP distribution, which is overdispersed.

Now let us look at the conditional distributions associated with the CMP. The simplest one, namely $P(X_1 = l | X_1 + X_2 = k)$ is given by

$$P(X_1 = l | X_1 + X_2 = k) = \frac{\binom{k}{l}^\zeta}{\sum_{i=0}^k \binom{k}{i}^\zeta}$$

Going one step further, we have

$$P(X_1 = l | X_1 + X_2 + \dots + X_n = k) = \frac{\sum_{n=1}^{k-l} \left(\frac{k!}{l!n!(k-l-n)!} \right)^\zeta}{\sum_{i_1+\dots+i_n=k-l} \left(\frac{k-l}{i_1! \dots i_n!} \right)^\zeta}$$

where $i_1 = \dots = i_m = k$. In general, we have,

$$P(\sum_{i=1}^n X_i = k) = \frac{\sum_{l_2} \sum_{l_3} \dots \sum_{l_n} \left(\frac{k!}{l_1! l_2! \dots l_n!} \right)^\zeta}{\sum_{x_1} \sum_{x_2} \dots \sum_{x_n} \left(\frac{k!}{x_1! \dots x_n!} \right)^\zeta},$$

where $\sum_{i=1}^m l_i = k - l$ and $\sum_{i=1}^n x_i = k$.

Also, $p_{m,n}(l, k) = P(\sum_{i=1}^m X_i = l | \sum_{i=1}^n X_i = k)$ for $m < n$, then,

$$p_{m,n}(l, k) = \frac{\left(\sum_{l_1} \dots \sum_{l_m} \left(\frac{l!}{l_1! \dots l_m!} \right)^\zeta \sum_{j_1} \dots \sum_{j_{n-m}} \left(\frac{(k-l)!}{j_1! \dots j_{n-m}!} \right)^\zeta \right)}{\sum_{x_1} \dots \sum_{x_n} \left(\frac{k!}{x_1! \dots x_n!} \right)^\zeta}$$

for $\sum_{i=1}^m j_i = k - l$, $\sum_{i=1}^n x_i = k$.

For a sample of size n from a CMP population, the likelihood function is given by

$$L(\zeta, \lambda | x_1, \dots, x_n) = \frac{\lambda^{\sum x_i}}{(\prod_{i=1}^n x_i!)^\zeta \Psi^n(\zeta, \lambda)},$$

where $\Psi(\zeta, \lambda) = \sum_{k=0}^{\infty} \frac{\lambda^k}{(k!)^\zeta}$. Then the vector $(\prod_{i=1}^n X_i!, \sum_{i=1}^n X_i)$ is jointly minimally sufficient for (ζ, λ) . Now let us write $Y = \prod_{i=1}^n Y_i = \prod_{i=1}^n X_i!$. Then Y_i takes values $1, 1!, 2!, \dots$ etc., i.e., $1, 1, 2, 6, 24, \dots$ etc. As a result, if y is the product of n negative integer x , then

$$P(Y = y) = P(X_i! = x!) = P(X_i = x) = \frac{\lambda^x}{(x!)^\zeta \sum_{k=0}^{\infty} \frac{\lambda^k}{(k!)^\zeta}}$$

So, we can write $P(Y = y)$ where y is the product of n uniquely determined factorials of $x_1, \dots, x_n$ (it is not necessary that all $x_1, \dots, x_n$ are different) as $y_i = x_i!$ for $i = 1, 2, \dots, n$. Suppose that the only distinct numbers $x_1, \dots, x_n$ are $z_1, \dots, z_k$. Let n_j be the number of times z_j is present for $j = 1, 2, \dots, k$, so that $\sum n_j = n$. Then we have

$$P(Y = y) = \frac{n!}{n_1! \dots n_k!} \prod_{j=1}^k \left(\frac{\lambda^{z_j}}{(z_j!)^\zeta \sum_{l=0}^{\infty} \frac{\lambda^l}{(l!)^\zeta}} \right)^{n_j}$$

which is the same as

$$P(Y = y) = \frac{n!}{n_1! n_2! \dots n_k!} \frac{\lambda^{\sum x_i}}{(\prod_{i=1}^n x_i!)^\zeta \Psi^n(\zeta, \lambda)}$$

6 Zero-inflated Conway-Maxwell-Poisson distribution

The zero-inflated CMP distribution is defined as

$$P(X = 0) = \phi + (1 - \phi) \frac{1}{\Psi(\zeta, \lambda)}$$

$$\text{and } P(X = k) = (1 - \phi) \frac{\frac{\lambda^k}{(k!)^\zeta}}{\Psi(\zeta, \lambda)}$$

for $0 < \phi < 1$ and $k = 1, 2, \dots$, where $\Psi(\zeta, \lambda)$ is as in the previous section. Then the bivariate distribution for two i.i.d. ZICMP random variables is given by

$$P(X_1 + X_2 = 0) = \left(\phi + (1 - \phi) \frac{1}{\Psi(\zeta, \lambda)} \right)^2$$

$$\text{and } P(X_1 + X_2 = k) = 2\phi(1 - \phi) \frac{\frac{\lambda^k}{k!}}{\Psi(\zeta, \lambda)} + (1 - \phi)^2 \frac{\frac{\lambda^k}{k!}}{(\Psi(\zeta, \lambda))^2} \sum_{l=0}^k \binom{k}{l}^\zeta.$$

In general, if we denote $G_{m,n}(k, \zeta) = \sum_{i_{m+1}} \dots \sum_{i_n} \left(\frac{k!}{i_{m+1}! \dots i_n!} \right)^\zeta$, then we have

$$P\left(\sum_{i=1}^n X_i = k\right) = \sum_{m=0}^n \binom{n}{m} \frac{\lambda^k}{k!} \left(\phi + (1 - \phi) \frac{1}{\Psi(\zeta, \lambda)} \right)^m \left(\frac{1 - \phi}{\Psi(\zeta, \lambda)} \right)^{n-m} G_{m,n}(k, \zeta).$$

If we condition on the event $n_0 = 0$ where n_0 is once again the number of zero-values, the conditional pmf of $\sum_{i=1}^n X_i$ has a somewhat simpler form. Of course, n_0 itself is binomial with probability of success $p = \phi + (1 - \phi) \frac{1}{\Psi(\zeta, \lambda)}$.

The simplest case is

$$P(X_1 + X_2 = k | n_0 = 0) = \frac{\lambda^k}{(k!)^\zeta} \left(\frac{1}{\Psi(\zeta, \lambda) - 1} \right)^2 \sum_{l=1}^{k-1} \binom{k}{l}^\zeta.$$

In general,

$$P\left(\sum_{i=1}^n X_i = k | n_0 = 0\right) = \frac{\lambda^k}{(k!)^\zeta} \left(\frac{1}{\Psi(\zeta, \lambda) - 1} \right)^n \sum_{i_1} \dots \sum_{i_n} \left(\frac{k!}{i_1! \dots i_n!} \right)^\zeta,$$

where each i_i is positive and $i_1 + \dots + i_n = k$. Also, as in Section 3, it can be shown that $P(\sum_{i=1}^n X_i = k | n_0 = j)$ is the same as $P(\sum_{i=1}^{n-j} X_i = k | n_0 = 0)$. As for the conditional joint pmf of $X_1, \dots, X_n$ given $n_0 = 0$, it is

$$P(X_1 = x_1, \dots, X_n = x_n | n_0 = 0) = \frac{\lambda^{\sum x_i}}{(\prod x_i!)^\zeta} \left(\frac{1}{\Psi(\zeta, \lambda) - 1} \right)^n$$

and, as was the case in Section 3, it turns out the joint pmf of n i.i.d. zero-truncated CMP random variables, if we derive the conditional pmf of a usual ZIP distribution, we derive the conditional pmf which is

$$P(X_1 = l_1 | \sum_{i=1}^n X_i = k, n_0 = 0) = \frac{\sum_{l_2} \dots \sum_{l_n} \frac{k!}{l_1! \dots l_n!}}{\sum_x \frac{k!}{x!}}$$

with $\sum_{i=2}^n l_i = k - l_1$ and $\sum_{i=1}^n x_i = k$.

Also, if we denote $p_{m,n}(l, k) = P(\sum_{i=1}^m X_i = l, \sum_{i=m+1}^n X_i = k - l, n_0 = 0)$ for $m < n$, then,

$$p_{m,n}(l, k) = \frac{\binom{k}{l} \sum_{l_1} \dots \sum_{l_m} \frac{l!}{l_1! \dots l_m!}}{\sum_{x_1} \dots \sum_{x_m} \frac{k!}{x_1! \dots x_m!}}$$

for $\sum_{i=1}^m l_i = l$, $\sum_{i=m+1}^n j_i = k - l$, $\sum_{i=1}^n x_i = k$ same as that of usual COM-Poisson.

7 Generalized Zero-Inflated

Angers and Biswas (2003) introduced the following generalized zero-inflated Poisson (henceforth GZIP) distribution

$$P(X = 0) = \phi + (1 - \phi)e^{-\lambda}$$

$$\text{and } P(X = k) = (1 - \phi) \frac{\lambda^k}{k!} e^{-\lambda}$$

for $k = 1, 2, \dots$ with the parameters satisfying $0 \leq \phi < 1$, $0 \leq \alpha < \lambda^{-1}$ and $\lambda > 0$. So, a new parameter α is introduced. It is easy to see that choosing $\alpha = 0$ brings us back to the ZIP distribution. See the appendix for a proof of the fact that the mean and the variance are given by

$$E(X) = \frac{\lambda(1 - \phi)}{1 - \phi + \phi\lambda}$$

and

$$V(X) = \frac{\phi(1 - \phi)\lambda^2}{(1 - \phi + \phi\lambda)^2}$$

The sample likelihood function is given by

$$L(\alpha, \phi, \lambda | x_1, \dots, x_n) = (\phi + (1 - \phi)e^{-\lambda})^{n_0} \prod_{i=1}^n \frac{(1 - \phi)e^{-\lambda} \lambda^{x_i}}{x_i!}$$

the unconditional pmf, as in the case of ZIP given $\sum_{i=1}^n X_i = k$, $n_0 = 0$ for

$$\frac{\sum_{l_1} \dots \sum_{l_n} \frac{k!}{l_1! \dots l_n!}}{\sum_x \frac{k!}{x!}}$$

if

$$\left(\frac{(k-l)!}{j_1! \dots j_{n-m}!} \right)$$

the formulas are exactly

istribution

generalized zero-

$(-e^{-\lambda})^{-1} < \phi < 1$ the picture. It is the ZIP distribution. $\phi = 1$ and that the

$$\prod_{i=1}^n \frac{(1 + \alpha e^{-\lambda})^{x_i - 1}}{x_i!}$$

The 3×1 vector (X_1, X_2, X_3) is jointly sufficient for the parameters (α, ϕ, λ) . Under this pmf of $X_1 + X_2$ will be

$$P(X_1 + X_2 = k) = 2\phi(1 - \phi)P(U_1 = k) + (1 - \phi)^2 P(U_2 = k) = \left[\frac{e^{\alpha\lambda k} (1 + \frac{k}{2}\alpha)}{1 + k\alpha} \right]^{k-1}$$

where $U_j \sim GP(\alpha, \phi)$ for $j = 1, 2$ and $GP(a, b)$ is the GP given by

$$P(U = k) = e^{-b}; P(U = k) = \frac{(1 + ak)^{k-1} e^{-b}}{k!}$$

for positive values of k .

An alternate expression for the pmf of $X_1 + X_2$ is

$$P(X_1 + X_2 = k) = 2\phi(1 - \phi)P(V_1 = k) + (1 - \phi)^2 P(V_1 + V_2 = k)$$

where $V_j \sim GP(\alpha, \lambda)$ for $j = 1, 2$. We can go two steps further and obtain the pmfs of $X_1 + X_2 + X_3$ and $X_1 + X_2 + X_3 + X_4$ as follows

$$P(X_1 + X_2 + X_3 = k) = 3P(X_1 = 0)P(X_2 + X_3 = k) \\ = 3 \sum_{k_1, k_2, k_3 \neq 0} \prod_{i=1}^3 \frac{(1 + k_i \alpha)^{k_i-1}}{k_i!} e^{-\lambda(3 + \alpha k)} \lambda^k,$$

where the sum is over all positive integers k_1, k_2, k_3 adding up to k . This simplifies to

$$P(X_1 + X_2 + X_3 = k) = 6\phi(1 - \phi)[\phi + (1 - \phi)e^{-\lambda}]P(V_1 = k) + (1 - \phi)^2 P(V_1 + V_2 = k) \\ + (1 - \phi)^3 P(V_1 + V_2 + V_3 = k).$$

Next,

$$P(X_1 + X_2 + X_3 + X_4 = k) = 4P(X_1 = 0)P(X_2 + X_3 + X_4 = k) \\ = 4 \sum_{k_1, k_2, k_3, k_4 \neq 0} \prod_{i=1}^4 \frac{(1 + k_i \alpha)^{k_i-1}}{k_i!} e^{-\lambda(4 + \alpha k)} \lambda^k,$$

where the sum is over all positive integers k_1, k_2, k_3, k_4 adding up to k . This simplifies to

$$P(X_1 + X_2 + X_3 + X_4 = k) = 24\phi(1 - \phi)[\phi + (1 - \phi)e^{-\lambda}]^2 P(V_1 = k) \\ + 12(1 - \phi)^2 [\phi + (1 - \phi)e^{-\lambda}]P(V_1 + V_2 = k) \\ + 4\phi(1 - \phi)^3 P(V_1 + V_2 + V_3 = k) + (1 - \phi)^4 P(V_1 + V_2 + V_3 + V_4 = k).$$

In general, by the binomial theorem, we have

$$P(X_1 + \dots + X_n = k) = \sum_{j=0}^k \binom{k}{j} (1-\phi)^j \phi^{k-j} e^{-\lambda} j^{n-j-1} P(V_1 + \dots + V_j = k)$$

Suppose we have $V_1, \dots, V_n$ are independent and identically distributed random variables with $V_1 \sim \text{Poisson}(\lambda)$ and $V_2, \dots, V_n \sim \text{Poisson}(\lambda + \alpha)$. Then, we have

we have this for $n+1$ as follows:

$$\begin{aligned} & P(X_1 + \dots + X_{n+1} = k) \\ &= \sum_{j=0}^k \binom{k}{j} (1-\phi)^j \phi^{k-j} e^{-\lambda} j^{n-j-1} P(V_1 + \dots + V_j = k) \\ &+ (1-\phi)^n P(V_1 = k) \\ &= \sum_{j=0}^{n-1} \binom{k}{j} (1-\phi)^j \phi^{k-j} e^{-\lambda} j^{n-j-1} P(V_1 + \dots + V_j = k) \\ &+ (1-\phi)^n P(V_1 = k) \\ &= \sum_{j=0}^{n-1} \binom{k}{j} (1-\phi)^j \phi^{k-j} e^{-\lambda} j^{n-j-1} P(V_1 + \dots + V_j = k) \\ &+ (1-\phi)^n P(V_1 = k) \\ &= \sum_{j=0}^{n-1} \binom{k}{j} (1-\phi)^j \phi^{k-j} e^{-\lambda} j^{n-j-1} P(V_1 + \dots + V_j = k) \\ &+ (1-\phi)^n P(V_1 = k) \end{aligned}$$

8 Bayesian inference for the ZIP model

Bayesian analysis of the ZIP model has been discussed in the literature (e.g., (2006)). In this section, we consider a Bayesian hierarchical model for the ZIP model. We assume that the parameters λ and α are unknown and we adopt a conjugate prior distribution for them. Specifically, we assume that λ and α are independent and follow Gamma distributions. The joint posterior distribution of λ and α given the data $\mathbf{y}$ is then given by

that leads naturally to a method of two-sample (or multi-sample) comparison. Then, to represent a conjugate Bayesian analysis for the sum of (conditionally) i.i.d. random variables, we have independent random variables from J ZIP populations may, representing counts of individual visits, correspond in a particular way to J different treatments administered independently at K time points. Let Y_{jk} be the k th replicate count observed under the j th treatment ($j = 1, \dots, J$ and $k = 1, \dots, K$). We assume $Y_{jk} \sim ZIP(p, \lambda_{jk})$. In other words:

$$\text{pr}(Y_{jk} = y) = \binom{y}{p} p^p (1 - p)^{y-p} \text{pr}(Y_{jk}^* = y) \quad (3)$$

for some $0 < \phi < 1$, where $Y_{jk}^* \sim \text{Pois}(\lambda_{jk})$. Then we model $\log(\lambda_{jk})$ as

$$\log(\lambda_{jk}) = \beta_j + \epsilon_{jk}, \quad (4)$$

where ϵ_{jk} is the random residual component following $\text{Normal}(0, \sigma_\epsilon^2)$. The use of a random component in the link function specification is consistent with the belief that there may be unexplained sources of variation in the data, perhaps due to unmeasured covariate or unmodeled random effects. This is particularly appropriate for Poisson data sets with over-dispersion. The use of residual effects in GLMs is discussed in Surratt et al. (2000) and is a special case of the class of generalized linear mixed models (Zeger and Karim, 1991; Easlow and Clayton, 1993). Equation (3) boils down to

$$\log(\lambda_{jk}) \sim \text{Normal}(\beta_j, \sigma_\epsilon^2) \quad (5)$$

where β_j is the effect of the j th treatment. We use conjugate priors for this hierarchical model and center the parameters for efficient MCMC sampling (Gelfand et al., 1990). Let $\mathcal{NIG}$ be the Normal-Inverse Gamma family of conjugate distributions in which the mean follows a normal distribution conditionally on the variance and the variance marginally follows an Inverse-Gamma distribution with the hyper-prior parameters μ, σ^2, u, v having the appropriate subscripts. In other words,

$$\begin{aligned} \theta, \sigma^2 &\sim \mathcal{NIG}(\theta_0, \sigma_0^2, u, v) \text{ implies that} \\ \theta | \sigma^2 &\sim \mathcal{N}(\theta_0, \sigma^2) \text{ and} \\ \sigma^2 &\sim \text{IG}(\sigma_0^2, v). \end{aligned}$$

With this notation in mind, this is how we specify our priors:

$$\begin{aligned} \beta_j, \sigma_\beta^2 &\sim \mathcal{NIG}(\mu_\beta, \sigma_\beta^2, u_\beta, \pi, \beta, \pi) \\ \mu, \sigma_\mu^2 &\sim \mathcal{NIG}(\mu_0, \sigma_\mu^2, u_\mu, \pi, \mu, \pi) \end{aligned}$$

However, the specification of the hyper-inflation parameter makes the sampling from the conditional posterior distribution extremely difficult. Agarwal et al (2002) and Ghosh et al (2006) cleverly handle the problem by introducing a latent variable. In the present context, denoting the latent variable corresponding

to Y_{jk} by Z_{jk} , the complete likelihood of the data is $L(y, z | \phi, \lambda) =$

$$\prod_j \prod_k \phi^{z_{jk}} \left\{ (1 - \phi) \frac{e^{-\lambda_{jk}} \lambda_{jk}^{y_{jk}}}{y_{jk}!} \right\}^{1-z_{jk}} \quad (6)$$

or, equivalently, $L(y, z | \phi, \lambda) =$

$$\phi^{n_0} (1 - \phi)^{n - n_0} \prod_{y_{jk} > 0} \frac{e^{-\lambda_{jk}} \lambda_{jk}^{y_{jk}}}{y_{jk}!} \prod_{y_{jk} = 0} (e^{-\lambda_{jk}})^{1 - z_{jk}}, \quad (7)$$

where $n_0 = \sum_j \sum_k z_{jk}$ and $n = JK$.

We assume a $\text{Beta}(a, b)$ prior on ϕ and elicit conjugate priors for all the variance parameters. In summary, our hierarchical model is given by:

$$\begin{aligned} Y_{jk} &\sim \text{ZIP}(\phi, \lambda_{jk}) \\ \phi &\sim \text{Beta}(a, b) \\ \log(\lambda_{jk}) &\sim \text{Normal}(\mu_j, \sigma_\epsilon^2) \\ \sigma_\epsilon^2 &\sim \text{IG}(u_\epsilon, \pi) \\ \beta_j, \sigma_\beta^2 &\sim \text{NIG}(\mu_\beta, \nu_\beta, \pi) \\ \mu, \sigma_\mu^2 &\sim \text{NIG}(\mu_\mu, \nu_\mu, \pi) \end{aligned}$$

Sampling from the posterior distribution can be performed using a block Gibbs sampler. All the conditional distributions except for those of the λ -values and ϕ have conjugate forms. Using latent variables has the advantage that sampling from the conditional distribution of the zero-inflation parameter reduces to sampling from its conjugate distribution. However, a Metropolis-Hastings step is needed for drawing the λ -values and a log-Normal proposal distribution will work. We would prefer using relatively flat priors for all the variance parameters.

Next we move onto a conjugate Bayesian analysis of $Z = \sum_{i=1}^n X_i$ where $X_1, \dots, X_n$ are conditionally i.i.d. ZIP (λ) variables given the parameters. One can assume that

(i) λ has an *a priori* $\Gamma(\mu, \alpha)$ density given by

$$f(\lambda) = \frac{\mu^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\mu\lambda}$$

for $\lambda > 0$ and a positive integer α ;

(ii) ϕ has an *a priori* $\text{Beta}(a, b)$ density for some $a > 0$ and $b > 0$ (one can choose $a = b = 1$ yielding a $U(0, 1)$ prior);

(iii) λ and ϕ are independent, in which case the joint prior density denoted by $h(\lambda, \phi)$ will be the product of the above two.

Then, if $\{X_1, X_2, \dots, X_n\}$ are independent from the p.d.f. in (1), which implies that $Z = \sum_{i=1}^n X_i$, with the usual posterior calculation, its distribution is $Z = z | \lambda, \mu$ in (2), one does not need the conditional or marginal distribution of Z is

$$P(Z = z) = \frac{1}{n+1} \sum_{r=0}^z P(N_r = z-r) + \frac{1}{n+1} \sum_{r=1}^n P(N_r = z-r)$$

where N_r is negative binomial with parameters $(\alpha, \frac{\mu}{\mu+1})$. The posterior joint distribution of λ and ϕ is

$$h^*(\lambda, \phi | z) = \frac{[\phi^n I(z=0) + \sum_{r=1}^n P(Y_n = r | Z_r = z)] \frac{\alpha^{-1} e^{-\mu\lambda}}{(n+1) \sum_{r=0}^z P(N_r = z-r)}}{\frac{1}{n+1} \sum_{r=0}^z P(N_r = z-r) + \sum_{r=1}^n P(N_r = z-r)}$$

END

Appendix 1A: Probability distribution
Consider the paper by Consul and Jain (1971) where they introduced a generalized Poisson distribution as follows

$$P(X = x | \lambda_1, \lambda_2) = \frac{(1 + \lambda_1)^x}{x!} e^{-(\lambda_1 + x\lambda_2)}$$

for $x = 0, 1, 2, \dots$ so that

$$P(X = x | \lambda_1, \lambda_2) = 0 \text{ for } x < 0$$

for $x \geq m$ if $\lambda_1 + m\lambda_2 \leq 0$. But in this case, is good enough to consider $\lambda \geq 0$ so that $P(X = x | \lambda_1, \lambda_2)$ is not zero. On the other hand, observe that setting $\lambda_2 = 0$, we can get back the usual Poisson distribution.

Consul and Jain referred to Consul (1902) in proving that $P(X = x | \lambda_1, \lambda_2) = 0$ for $x < 0$. They used one-dimensional Laplace's formula to prove that

$$\phi(z) = \phi(0) + \sum_{x=1}^{\infty} \frac{1}{x!} \left(\frac{d^x}{dz^x} \phi(z) \right)_{z=0} z^x \quad (8)$$

where $\phi(z) = e^{\lambda_1 z}$ and $f(z) = (1 + \lambda_1 + \lambda_2 z)^{-1}$. It is shown that

$$\left(\frac{d^{x-1}}{dz^{x-1}} f(z) \right)_{z=0} = (-1)^{x-1} (1 + \lambda_2 z)^{x-1} \Big|_{z=0} = (-1)^{x-1}$$

so that (8) becomes

$$e^{\lambda_1 z} = \sum_{x=0}^{\infty} \frac{1}{x!} \left(\frac{d^x}{dz^x} f(z) \right)_{z=0} z^x = \sum_{x=0}^{\infty} \frac{1}{x!} \left(\frac{z}{f(z)} \right)^x \quad (9)$$

This is true for all z . So, substituting $z = 1$ we get

$$e^{\lambda_1} = \sum_{x=0}^{\infty} \frac{1}{x!} (1 + \lambda_2 x)^x e^{-\lambda_2 x}$$

which implies

$$1 = \sum_{x=0}^{\infty} \frac{1}{x!} (1 + \lambda_2 x)^x e^{-\lambda_2 x}$$

so that the right hand side is 1, which is $\sum_{x=0}^{\infty} P(X=x|\lambda_1, \lambda_2)$. Now substituting $\lambda_1 = \lambda$ and $\frac{\lambda_2}{\lambda_1} = \alpha$ in (9), we get

$$P(X=x|\lambda, \alpha) = \frac{\lambda^x}{x!} (1 + \alpha x)^x e^{-(1+\alpha x)\lambda}$$

which is nothing but $U \sim GP(\lambda, \alpha)$. Our next step is to prove that

$$\sum_{x=0}^{\infty} P(X=x|\lambda, \alpha) = \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} (1 + \alpha x)^x e^{-(1+\alpha x)\lambda} = 1$$

Therefore, if we go to zero-inflated distribution where

$$P(X=0) = \phi + (1-\phi)e^{-\lambda}; P(X=x) = (1-\phi) \frac{\lambda^x}{x!} (1 + \alpha x)^x e^{-(1+\alpha x)\lambda}$$

for $x > 0$, then we get

$$\sum_{x=0}^{\infty} P(X=x) = \phi + (1-\phi) \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} (1 + \alpha x)^x e^{-(1+\alpha x)\lambda} = 1$$

Appendix 1B: Expectation of zero-inflated Poisson distribution

Here we derive the mean of a zero-inflated generalized Poisson distribution. For this, we go back to the paper by Consul and Jain (1973) and work with (9). Consul and Jain have their own way of dealing with negative values of λ_2 . In our case, we can simply differentiate (9) and arrive at the following:

$$\lambda_1 e^{\lambda_1 z} = \sum_{x=1}^{\infty} \frac{\lambda_1}{x!} (1 + \lambda_2 x)^{x-1} z^x e^{-\lambda_2 z} (1 - \lambda_2 z) \quad (10)$$

which is same as

$$\frac{\lambda_1}{1 - \lambda_2 z} = \sum_{x=1}^{\infty} \frac{\lambda_1}{x!} (1 + \lambda_2 x)^{x-1} z^x e^{-\lambda_2 z}$$

Now putting $z = 1$ once again,

$$\frac{1}{1-\lambda_2} = \sum_{x=1}^{\infty} \frac{\lambda_1(\lambda_1 + \lambda_2 x)^{x-1}}{(x-1)!} e^{-(\lambda_1 + \lambda_2 x)}$$

Observe that the right hand side is the expectation of X so that $E(X) =$
 Substituting $\lambda_2 = 0$, one can see that it is the mean of the usual
 Poisson distribution. Now, coming back to the generalized Poisson distribution
 (10), we get $E(X) = \sum_{x=1}^{\infty} \frac{\lambda^x (1 + \alpha x)^{x-1}}{x!} e^{-(1+\alpha x)\lambda} =$
 For the generalized Poisson distribution the expectation is given by

$$E(X) = (1-\phi) \sum_{x=1}^{\infty} \frac{\lambda^x (1 + \alpha x)^{x-1}}{x!} e^{-(1+\alpha x)\lambda} = \frac{\lambda(1-\phi)}{1-\alpha\lambda}$$

Setting $\phi = 0$, we get the mean of the usual generalized Poisson distribution

Appendix 1C: Variance of a GZIP distribution

The next goal is to get the variance of the zero-inflated generalized Poisson distribution. For this we again go back to the equation (10). We differentiate it to get

$$\lambda_1^2 \sum_{x=2}^{\infty} \frac{\lambda_1(\lambda_1 + \lambda_2 x)^{x-1}}{(x-2)!} z^{x-2} e^{-\lambda_2 x z} (1-2z)^2 - \sum_{x=1}^{\infty} \frac{\lambda_1(\lambda_1 + \lambda_2 x)^{x-1}}{(x-1)!} z^{x-1} e^{-\lambda_2 x z} \lambda_2 (2-\lambda_2 z)$$

which simplifies to

$$\lambda_1^2 \sum_{x=2}^{\infty} \frac{\lambda_1(\lambda_1 + \lambda_2 x)^{x-1}}{(x-2)!} z^{x-2} e^{-(\lambda_1 + \lambda_2 x)z} (1-2z)^2 - \sum_{x=1}^{\infty} \frac{\lambda_1(\lambda_1 + \lambda_2 x)^{x-1}}{(x-1)!} z^{x-1} e^{-(\lambda_1 + \lambda_2 x)z} \lambda_2 (2-\lambda_2 z)$$

Setting $z = 1$, we get

$$\lambda_1^2 \sum_{x=2}^{\infty} \frac{\lambda_1(\lambda_1 + \lambda_2 x)^{x-1}}{(x-2)!} e^{-(\lambda_1 + \lambda_2 x)} (1-\lambda_2)^2 - \sum_{x=1}^{\infty} \frac{\lambda_1(\lambda_1 + \lambda_2 x)^{x-1}}{(x-1)!} e^{-\lambda_2 x} \lambda_2 (2-\lambda_2) \quad (11)$$

Now the second term without the negative sign is just $E(X)$ multiplied by λ_2 and hence (11) becomes

$$= \sum_{x=2}^{\infty} \frac{\lambda_1(\lambda_1 + \lambda_2 x)^{x-1}}{(x-2)!} e^{-(\lambda_1 + \lambda_2 x)} (1-\lambda_2)^2 - \frac{\lambda_1 \lambda_2 (2-\lambda_2)}{1-\lambda_2}$$

From (10) one shows that

$$\sum_{x=2}^{\infty} \frac{\lambda_1(\lambda_1 + \lambda_2 x)^{x-1}}{(x-2)!} e^{-(\lambda_1 + \lambda_2 x)} = \lambda_1 \left(\frac{\lambda_1}{1-\lambda_2} + \frac{\lambda_2 (2-\lambda_2)}{(1-\lambda_2)^2} \right)$$

Now, substituting $\lambda_2 = 0$ in the above equation, we get

$$E[X] = \frac{\lambda_1^2}{(1 - \lambda_2)^2} + \frac{\lambda_1}{(1 - \lambda_2)^3}$$

Finalizing the above equation, we get

$$E[X] = \frac{\lambda_1}{(1 - \lambda_2)^3}$$

Substituting $\lambda_2 = 0$ in the above equation, we get the variance of the usual Poisson distribution. Now, if $U \sim GP(\alpha, \lambda)$, we get

$$E[U] = \lambda \left[\frac{\lambda}{(1 - \alpha\lambda)^2} + \frac{\alpha\lambda(2 - \alpha\lambda)}{(1 - \alpha\lambda)^3} \right]$$

so that $E[U^2] = \lambda^2 \left[\frac{\lambda}{(1 - \alpha\lambda)^3} + V(U) \right]$

Next, if U is a generalized Poisson distribution, we get

$$E[U] = \lambda \left[\frac{\lambda}{(1 - \alpha\lambda)^2} + \frac{\alpha\lambda(2 - \alpha\lambda)}{(1 - \alpha\lambda)^3} \right]$$

Adding the above two equations, we get

$$E[U] = \lambda \left[\frac{\lambda}{(1 - \alpha\lambda)^2} + \frac{1}{(1 - \alpha\lambda)^3} \right]$$

Finally, multiplying both sides, we get

$$E[U] = \frac{(1 - \phi)\lambda^2}{(1 - \alpha\lambda)^2} + \frac{(1 - \phi)\lambda}{(1 - \alpha\lambda)^3}$$

Substituting $\phi = 0$ in the above equation, we get the variance of the usual generalized Poisson distribution.

Appendix 2: Simulation results for biases and variances of estimators
For the method-of-moments estimators of λ and ϕ that we mention in Section 2, we report the results of some simulation studies. For a fixed value of $\phi \in (0, 1)$, we generated $n = 100$ from a $ZIP(\phi, \lambda)$ distribution with $\lambda = 1$ and computed the bias in the method-of-moments estimator. For each λ , we computed the sample variance of this average bias value. We repeated this exercise for 9 different values of ϕ .

by different colors, were superimposed. Similarly, the 9 ' λ vs. $\text{var}(\hat{\lambda}_{MOM})$ ' graphs were superimposed. In order to implement the same procedure with the roles of λ and ϕ switched, we now fixed λ at an integer value between 2 and 10 and in each case created a ' ϕ vs. average bias' graph and a ' ϕ vs. $\text{var}(\hat{\phi}_{MOM})$ ' graph. For $M = 1000$ random samples of size $n = 100$ drawn from a $\text{ZIP}(\phi, \lambda)$ with $\phi = 0.01, 0.02, \dots, 0.09$. On each graph, again, each of the two resulting curves were superimposed. Similarly, the 9 ' λ vs. $\text{var}(\hat{\lambda}_{MLE})$ ' graphs were superimposed. For the MLEs of λ and ϕ , the 4 sets of graphs which are λ vs. $\text{bias}(\hat{\lambda}_{MLE} - \lambda)$, λ vs. $\text{var}(\hat{\lambda}_{MLE})$, ϕ vs. $\text{bias}(\hat{\phi}_{MLE} - \phi)$ and ϕ vs. $\text{var}(\hat{\phi}_{MLE})$. All of the above were carried out again with $M = 10000$. It is clear that the MLEs of both parameters are asymptotically unbiased. For $\hat{\lambda}_{MLE}$ the bias is close to zero even for moderate values of λ such as 6 or 7. In the case of $\hat{\phi}_{MLE}$, the bias is negligible irrespective of ϕ for λ upwards of 5. The variance of $\hat{\lambda}_{MLE}$, for the value-range of λ considered here, are below 0.4 except for extreme values of ϕ . Also, it shows a slight increasing trend with λ . As for the variance of $\hat{\phi}_{MLE}$, the only one that stands out is the case $\lambda = 2$. When this aberrant behavior was caused by computational error needs to be investigated. Moving on to the bias and the variance of $\hat{\lambda}_{MOM}$, the bias shows more fluctuations in this case than for the MLE, but it is close to zero except for the extremely small values of ϕ . The variance of $\hat{\lambda}_{MOM}$ once again shows a slight increasing trend with λ and is below 0.4 over this value-range of λ . For some extremely small values of ϕ , the bias in $\hat{\phi}_{MOM}$ also shows more fluctuations compared to that in the MLE and the variance of $\hat{\phi}_{MOM}$ shows a quadratic trend with the maximum occurring around the mid-range of ϕ (or some small values of λ).

Acknowledgment

The second author (Datta) was partially supported by the Department of Statistics, Texas A & M University, through a National Cancer Institute grant (R01CA0301).

References

- [1] Biswas, Jean-Francois, Biswas, Atanu, 2013: A Bayesian analysis of zero-inflated generalized Poisson model. *Communications in Statistics & Data Analysis*, 2, p.37-46.
- [2] del Pino, J. del, Perez-Casany, Marta, 2014: Overdispersed and underdispersed Poisson generalizations. *Journal of Statistical Planning and Inference*, 34, p.486-500.
- [3] Datta, A.C., 1960: Estimating the parameters of a modified Poisson distribution. *J. Amer. Statist. Assoc.*, 55, p.143-144.

- [4] Consul, P.C., Jain, G.C., 1973: A Generalization of the Poisson distribution. *Technometrics*, **15**(4), p.791-799.
- [5] Cox, D.R., 1933: Some remarks on overdispersion. *Biometrika*, **26**, p.269-274.
- [6] Dahiya, R.C., Gross, A.J., 1973: Estimating the parameters of a Poisson sample. *J. Amer. Statist. Assoc.*, **68**, p.1733.
- [7] Ghosh, S.K., Mukhopadhyay, P., Lu, J.C., 2006: Bayesian analysis of zero-inflated regression models. *J. Statist. Plan. Inf.*, **36**, p.1360-1373.
- [8] Gupta, P.L., Gupta, R.C., Tripathi, R.C., 1990: Analysis of zero count data. *Comput. Statist. Data Anal.*, **33**, p.1547.
- [9] Heilborn, D.C., 1989: Generalized linear models for altered zero counts and overdispersion in count data. *Technical Report*, Department of Epidemiology and Biostatistics, University of California, San Francisco.
- [10] Heilborn, D.C., Gibson, D.R., 1990: Shared model use and heterogeneity concerning AIDS: regression modeling of zero count data. *Proceedings of the Sixth International Conference on Biostatistics*, San Francisco, CA.
- [11] Kadane, J.B., Shmiele, G., Minka, T.P., Berle, S., 1998: Boatwright Conjugate analysis of the Conway-Maxwell-Poisson distribution. *Analysis*, **1**(2), p.363-374.
- [12] Kemp, A.W., 1986: Weighted discrepancies and maximum likelihood estimation for discrete distribution. *Commun. Statist. Theory*, **15**, p.783-803.
- [13] Kokonendji, C.C., Mizere, D., Balakrishnan, N., 2008: Connecting Poisson weight function to overdispersion and underdispersion. *Plan. Inf.*, **138**(5), p.1287-1296.
- [14] Lambert, D., 1992: Zero inflated Poisson regression with an application to defects in manufacturing. *Technometrics*, **34**, p.1.
- [15] Li, C.S., Lu, J.C., Park, J., Kim, K.M., Brinkle, P.A., Petersen, A., 1999: A multivariate zero-inflated Poisson distribution and its inference. *Technometrics*, **41**, p.29-38.
- [16] Martin, D.C., Katti, S.K., 1965: Fitting of some contagious distributions to some available data by the maximum likelihood method. *Biometrika*, **52**, p.34-48.
- [17] Przyborowski, J., Wilenski, H., 1940: Homogeneity of results from samples from Poisson Series. *Biometrika*, **31**, p.323.

- [18] Saei, A., McGilchrist, C.A., 1997: Random threshold models applied to zero class data. *Austral. J. Statist.*, **39**, p.5-16.
- [19] Scollnik, D.P.M., 1995: Bayesian analysis of two overdispersed Poisson models. *Biometrics*, **51(33)**, p.1117-1126.
- [20] Shmueli, G., Minka, T.P., Kadane, J.B., Borle, S. and Boatwright, P., 2005: A useful distribution for fitting discrete data: Revival of the Conway-Maxwell-Poisson distribution. *J. Roy. Statist. Soc. Series C*, **54(1)**, p.127-142.
- [21] Singh, S.N., 1963: A note on inflated Poisson distribution. *J. Indian. Statist. Assoc.*, **1**, p.140-144.