Para r Inf nce for Zero-Inf ated Po istri tion anc Its Varia nts

San in rabo , Sujay Da a[2*] and Bas irul A. Po

August 10 2009

[1] Unive as Pa nerican, Edir urg, [2] Fred H tchinson Ca search ttle, [3] theon Netwo k Centric Sys m, * Corres author

1 I tion

Count at ntly a in real life a d are often m delled by a distribu io ver, it und sometin s that the nur ber of zeroes a datas than is allowed a Poisson m del. So, in accomm d xcess n roes, a zero-in lated Poisson istribution introdu ed eratu ee, for examp e, Cohen (196 , Singh (196 (1965), 977), p (1986) an Heilborn (198). Some in applica pr can e be found in H ilborn and Gi son (1990), (1992), al. (1 , Saei and M Gilchrist (19 7), Li et al. and Gl 999). ere have bee attempts to g neralize the distribu o bast, as the one in Consul and Jai (1973), and urn ha eraliz ersions of ZIP Angers and Bi was (2003) p a Bayes of su generalized IP model. Th Poisson dist s well- he equ of its mean nd variance— feature not or ma asets e variances e ceed their me s. This nec the intr an o spersed Poiss n (OP) model See Cox (19 Scollni ong o . Overdisper d Poisson mo els have bee used fo liseas a and one ca show that a P model is exampl mueli l. (2005) re ved another v riant of the distribu the C y-Maxwell-P isson (CMP) istribution, been in rlier hers and de onstrated its sefulness as for disc Later dane et al. 2006) provide a Bayesian of a C sing c gate priors. he CMP dist bution can extende inflat rsion.

Here para ic inference sults for a ZI distribution variant above e discuss poi estimation, h pothesis tes Bayesia Mu the existing work cited ab ve came as velopments and are somewhat scatt red ... To our knowledge, ere is no suc unified presentation ... inf ... gy for this class of tributions. However, this is not m int ... ticle. Most of the ults present d here are ne y (except ome ... e quoted from the erences cited). Detailed p oofs and deriv ... ed wherever possi- . We believe this article will serve s a ... future theoretical d applied res arch on this interesting fami ... . The layout of the per is as follows. In section 2, we pr vide ... on the ZIP family distributions. In section 3 we derive the ... conditional distri- tions of sum of ZIP varia les. Sect n 4 ... on and hypothesis ting for ZIP data. Section 5 provid s som ... derivations for the nway-Maxw ll-Poisson (CMP) dist buti ... h, section 6 takes the zero-inf ated version of CMP. Secti ... the zero-inflated neralized Poi son distribu on and ... co ... ion (section 8) on yesian infere ce for the ZIP model. Som ... ils are provided in pendix 1, w ile Appendix 2 descri s so ... dies related to the thodology o tlined in sec ion 4.

Prelim inaries

e basic idea behind a ZIP distributi n wi ... nd λ is to assign a her probability to zero an lower pr babi ... es as compared to orresponding Poisson distr bution w h p ... a ZIP distribution h two param eters is form lly defin by:

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

$$P(X = \ldots) = (1 - \ldots) e^{- \ldots}$$

ere ϕ is a nu nber betwee 0 and 1 nd ... so write this as:

$$P(X = x) = \phi I(x = 0) \ldots \tag{1}$$

any non-neg ative integer , where $I(A)$ i ... ction of the event

The above efinition is t e most standa ... g a ZIP distribu- n. However, here are sev ral other ways ... ll of them use two ameters and are based on simple-m hec ... n ... . We shall discuss w of them here.

The first alt ernative defi ition is:

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

$$P(X = x) = \left(1 - \frac{\ldots}{1 - e^{-\lambda}} \ldots\right) \ldots 0$$

So, once again, this ZIP distributio [illegible] as two [illegible] rame[illegible]s γ and λ and it is clear that γ is nothing but $\phi(1-e^{-\lambda})$. [illegible] s de[illegible] n is [illegible] ded by the following simple-minded intuition: increase the [illegible] abil[illegible]ty [illegible] 0 by [illegible] constant number γ. One has to keep in mind that γ should [illegible] less th[illegible] $1-$ [illegible] λ.

A second way of defining the ZIP [illegible] ibut[illegible]

$$P(X=0) = (\quad \delta)\, e^{-\lambda} \quad \text{[illegible]d}$$

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

This time, the two parameters are δ a[illegible] and [illegible] can [illegible]ow that δ is nothing but $\frac{\phi(1-e^{-\lambda})}{e^{-\lambda}}$. This definition is gu[illegible] by t[illegible] [illegible]llow[illegible] intuition: increase the probability at 0 by multiplying the [illegible] higina[illegible] [illegible] abil[illegible] by a number bigger than 1. Of course, one has to keep in [illegible] o[illegible] th[illegible] [illegible] oul[illegible] e less than $e^{\lambda}-1$.

A third alternative definition of th[illegible] [illegible] utio[illegible]:

$$P(X=0) = 1 - [illegible] \quad \text{anc}$$

$$P(X=x) = e^{-\mu} [illegible] \quad \text{for } [illegible], 2, [illegible]$$

with $0<r<1$. Here the two paramet[illegible] are [illegible] λ. [illegible] e can easily observe that the two parameters of the origin[illegible] PD [illegible] ly, [illegible] λ) can be recovered from (ρ,μ) as: $\lambda = \mu\rho$, $\phi = 1 - e^{-\mu+\mu}$ [illegible]

The mean and variance of ZIPD ar[illegible] given by

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

The moment generating function $M_X($[illegible] nd the [illegible] arac[illegible] istic function $\Phi_X(t)$ are given by

$$M_X(t) = \phi + (1-\phi)\, e^{-\lambda+\lambda e^t} \quad [illegible]_X(t) = [illegible] + (1-\phi)\, e^{-\lambda+\lambda e^{it}}$$

Now it should be clear why a ZIP distr[illegible] ion is [illegible] xam[illegible] of an overdispersed Poisson model. Next we look at a weig[illegible] [illegible] vers[illegible] f a [illegible] sson distribution. It is given by

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

where Y is a usual Poisson random va[illegible] e. S[illegible] [illegible] e w[illegible] to consider the ZIP distribution as a weighted version of a [illegible] sson d[illegible] ibution, then

$$\frac{w(y)}{E[w(Y)]} \quad \Phi \quad 1-[illegible]$$

for $y>0$ and

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

where $\delta = \dfrac{\phi(1-e^{-\lambda})}{e^{-\lambda}}$. 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 $\delta = \dfrac{(1-e^{-\lambda})}{e^{-\lambda}}$. 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 $X_1, X_2, ..., X_n$ are independent ZIP variables with X_i having parameters (ϕ_i, λ_i) for $i = 1, 2, ..., n$ and if $Z = X_1 + \cdots + X_n$, then using induction on n,

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

In case the X_is 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

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

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 $X_1, X_2, ..., X_n$ 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 $X_1, ..., X_n$ given $n_0 = j$ is

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

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

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

Notice 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

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

where $Z_n \sim$ Poisson$(n\lambda)$ and $\xi_n(k)$ is a known, parameter-free function. For $n = 2$, $\xi_2(k) = 2^k - 2$; for $n = 3$, $\xi_3(k) = 3^k - 3.2^k + 3$, 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

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

where W has a usual Poisson distribution with parameter $(n-k)\lambda$. This is exactly the same as $P(\sum_{i=1}^{n-k} X_i = a \mid n_0 = 0)$. 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 $m \le n$) 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 $m \le n$) 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 $m \le n$) and $n_0 = k^*$, then it turns out to be parameter-free, although not binomial. For $k^* = 0$, it is actually symmetric. For example,

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

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

$$P(X_1 = 0 | X_1 + X_2 + X_3 = 2,\, n_0 = 1) = \frac{1}{3},\, P(X_1 = 1 | X_1 + X_2 + X_3 = 2,\, n_0 = 1) = \frac{2}{3};$$

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

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

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

In general, for $m < n$,

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

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

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

So, if $X_1, X_2, ..., X_n$ 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

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

The sample likelihood function is given by

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

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

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

For this pair of equations no analytic solution exists and it must be solved numerically, which involves the Lambert's W function (see, for example, Corless 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 × 1 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 $X_1, \ldots, X_n$ are i.i.d. following a ZIP(ϕ_1, λ_1) distribution, $Y_1, \ldots, Y_n$ are i.i.d. following a ZIP(ϕ_2, λ_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, $n_0 + m_0$ has 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 $(1 - \alpha/2)^{th}$ percentile of, or falls below the $\alpha/2^{th}$ 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 partition 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 P(Type I error $| H_0$) is equal to

$$\sum_{k=0}^{2n} \sum_{k^*=0}^{\infty} P(\text{Type I error} | H_0, n_0+m_0 = k, \sum_{i=1}^{n}(X_i+Y_i) = k^*).P(n_0+m_0 = k, \sum_{i=1}^{n}(X_i+Y_i) = k^*)$$

This test is, in some sense, a generalization of the one introduced by Przyborowski and Wilenski (1940) for the equality of two Poisson means.

5 The Conway-Maxwell-Poisson Distribution

The CMP distribution w[illegible] the abbreviated form of th[illegible] distribution [illegible]troduced by Conway and [illegible]ell (1962) also looks like a v[illegible]nt of the Z[illegible] distribution under certai[illegible]ptions. He[illegible] the probability [illegible]ass function [illegible] as follows:

$$[illegible]\,x) = \frac{\lambda^x}{(x!)^{\zeta}\sum_{k=0}^{\infty}\frac{\lambda^k}{(k!)^{\zeta}}}$$

for $x = 0, 1, 2, \ldots$ and λ an[illegible] positive. For $x = 0$, therefore, [illegible] probability [illegible] $\left(\sum_{k=0}^{\infty}\frac{\lambda^k}{(k!)^{\zeta}}\right)^{-1}$. One ca[illegible]ve that for $\zeta = 1$, this probab[illegible]y distributi[illegible] is nothing but Poisson wit[illegible]heter λ. For ζ [illegible] 1, the probab[illegible]y at 0 is bigg[illegible] than $e^{-\lambda}$, which may lea[illegible] believe that the CMP distrib[illegible]on with $\zeta >$ [illegible] is an example of a usual Z[illegible]ibution. This, howeve[illegible], is not [illegible]e case becau[illegible] for values of $x > 0$, the v[illegible]unction associated with the [illegible] is depende[illegible] on x, unlike the constant [illegible] function o[illegible] a ZIP distributi[illegible] However, [illegible] still discuss it here as a g[illegible]ation of the usual Poisson di[illegible]bution, wh[illegible] zero-inflated version will [illegible]duced later. We begin our di[illegible]ssion with t[illegible] distribution of the convol[illegible] $X_1, X_2, \ldots, X_n$ which are i.i.[illegible] CMP (λ, ζ). [illegible] is given by

$$P(\sum_{i=1}^{n} X_i = [illegible]\ \frac{\frac{\lambda^k}{(k!)^{\zeta}}}{[illegible]^{\infty}\frac{\lambda^j}{(j!)^{\zeta}}}\sum_{i_1}\cdots\sum_{i_n}\left(\frac{k!}{i_1!\cdots i_n!}\right)[illegible]$$

for $k = 0, 1, 2, \ldots$ etc. This [illegible]so be proved by induction on [illegible] Calculation [illegible] the mean and the varianc[illegible]plicated in this case and so is [illegible] derivation [illegible] the method-of-moments ([illegible]timators. It may be worth poi[illegible]ng out that [illegible] $\zeta > 1$, the CMP distributi[illegible]nderdispersed, as opposed to [illegible] overdispersi[illegible] ZIP distribution.

Now let us look at the [illegible]ional distributions associated [illegible]th the CM[illegible] The simplest one, namely [illegible] $= l | X_1 + X_2 = k)$ is given by

$$P(X_1 = [illegible] X_2 = k) = \frac{\binom{k}{l}^{\zeta}}{\sum_{i=0}^{k}\binom{k}{i}^{\zeta}}.$$

Going one step further, w[illegible]

$$P(X_1 = l | X_1 + X[illegible] = k) = \frac{\sum_{m=[illegible]}^{k-l}\left(\frac{k!}{l!m!(k-l-[illegible])!}\right)^{\zeta}}{\sum_{i_1}\sum_{i_2}\sum_{i_3}\left(\frac{k}{i_1!i_2[illegible]}\right)^{\zeta}},$$

where i_1 - $_3 = k$. In general, we have,

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

where $\sum_i^{\prime}$ $k-l$ and $\sum_{=1} x_i = k$.

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

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

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

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

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

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

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

So, we ca ate $P(Y = y$ where y is the product of n uniquely determined factorials, e factorials o $x_1, \ldots, x_n$ (it is not necessary that all $x_1, \ldots, x_n$ are differ et $y_i = x_i!$ for $i = 1, 2, ..., n$. Suppose that the only distinct numbers $x_1, \ldots, x_n$ are $z_1, \ldots, z_k$. Let n_i be the number of times z_i is present fo 2, ..., k, so that $\sum n_i = n$. Then we have

$$Y = y) = \frac{n!}{n_1! \, {}_2!\cdots n_k!}\prod_{j=1}^{k}\left(\frac{\lambda^{z_j}}{(z_j!)^{\zeta}\sum_{l=0}^{\infty}\frac{\lambda^l}{(l!)^{\zeta}}}\right)^{n_j}$$

which is t e as

$$P(Y = y) = \frac{n!}{_1!n_2!\cdots n_k!}\frac{\lambda^{\sum x_i}}{(\prod_{i=1}^{n} x_i!)^{\zeta}\Psi^n(\zeta,\lambda)}$$

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 $k=1,2,\cdots$, where $\Psi(\zeta,\lambda)$ is as in the previous section. Then the convolution 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 $G_{m,n}(k,\zeta) = \sum_{i_{m+1}}\cdots\sum_{i_n}\left(\frac{k!}{i_{n+1}!\cdots i_n!}\right)^\zeta$, then we have

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

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,

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

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

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

and, as was the case in Section ... turns out ... the unconditional joint pmf of n i.i.d. ... ero-trunc... CMP ran... ext, as in the case of a usual ZIP distribution, we ... ve the co... X_1 given $\sum_{i=1}^{n} X_i$ which is

$$P(X_1 = l_1 | \sum_{i=}^{n} X_i = k, \ldots 0) = \frac{\sum_{l_2} \ldots \left(\frac{k!}{\ldots l_n!}\right)^{\zeta}}{\sum_{x} \ldots \left(\frac{k!}{\ldots x_n!}\right)^{\zeta}},$$

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

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

$$p_{m,n}(l,k) = \frac{\binom{k}{l} \sum_{l_1} \cdots \sum \ldots \frac{l!}{l_1! \cdots l_m!} \ldots \left(\frac{(k-l)!}{j_1! \cdots j_{n-m}!}\right)^{\zeta}}{\ldots x_1 \cdots \sum_{x} \ldots}$$

for $\sum_{i=1}^{m} l_i = l$, $\sum_{i=1}^{n-m} j_i = k \ldots$, $\sum_{i=1}^{n} x_i$ ... the ... ulas are exactly same as that of usual COM-P... n.

7 Generalized Zer... nflate... istribution

Angers and Biswas (2003) into... d the foll... generalized zero-inflated Poisson (henceforth G... distribut...

$$P(X \ldots) = \phi + (1 \ldots$$

$$\text{and } P(X = k) \ldots - \phi) \frac{(1 - \ldots}{\ldots}$$

for $k = 1, 2, \ldots$ with the parame... satisfying ... $(1 - e^{-\lambda})^{-1} < \phi <$ 1, $0 \leq \alpha < \lambda^{-1}$ and $\lambda > 0$. So, ... w param... the picture. It is easy to see that choosing $\alpha = \ldots$ ngs us ba... ZIP distribution. See the appendix for a proof o... fact that ... $= 1$ and that the mean and the variance are giv...

$$\ldots) = \frac{\lambda(1 - \ldots}{1 - \ldots}$$

and

$$V(X) = \frac{(1 - \phi)\lambda^2}{(1 - \alpha\lambda)^2} \ldots \frac{\lambda}{\ldots)^3}$$

The sample likelihood function ... given by

$$L(\alpha, \phi, \lambda | x_1, \ldots, x_n) = (\phi + (1 - \ldots \lambda)^{n_0} [(1 - \ldots]^{n_0} \ldots k \ldots)^{\sum x_i} \prod_{i=1}^{n} \frac{(1 + \alpha \ldots)^{x_i - 1}}{\ldots}$$

The 3×1 v[illegible] $\sum_{i=1}^{n} X_i$, $\prod_{i=}^{n}$ [illegible] $1+\alpha x_i)^{x_i-1}$) is jointl[illegible] [illegible]nimally sufficient for the thr[illegible] [illegible]me[illegible]s (α, ϕ, λ).
Unde[illegible] this [illegible], th[illegible] pmf of $X_1 +$ [illegible]$_2$ will be

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

, whe[illegible]e U_j [illegible] $(\alpha,$ [illegible]$)$ for $j = 1,$ [illegible]and $GP(a,b)$ is the p[illegible] given by

$$\text{[illegible]} J = \text{[illegible]}) = e^{-b};\ P(\text{[illegible]} = k) = \frac{(1+ak)^{k-1}}{k!} \text{[illegible]} b^k$$

for positive [illegible] values of k.

An alte[illegible] ex[illegible]ession for th[illegible] pmf of $X_1 + X_2$ is

$$P(X \text{[illegible]} = \text{[illegible]} = 2\phi(1-\phi \text{[illegible]} P(V_1 = k) + (1-\phi)^2 \text{[illegible]} + V_2 = k)$$

where $V_j \sim$ [illegible]$,\lambda)$ [illegible]or $j = 1, 2$. [illegible]e can go two steps f[illegible]r and obtain the pmfs of X_1 [illegible] $+ X$[illegible] and $X_1 + X$[illegible] $+ X_3 + X_4$ as follows

$$\text{[illegible]} + \text{[illegible]}_2 + X_3 = k) \text{[illegible]} 3P(X_1 = 0)P(X_2 + \text{[illegible]} = k)$$

$$\text{[illegible]} (1- \text{[illegible]})^3 \sum_{k_1,k_2,k_3\neq 0} \prod \frac{(1+k_i\alpha)^{k_i-1}}{k_i!} e^{-\lambda(3+\text{[illegible]}} \text{[illegible]} \lambda^k,$$

where the [illegible] ove[illegible] all positive i[illegible]tegers k_1, k_2, k_3 addi[illegible] to k. This simplifies to

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

Next,

$$P(X \text{[illegible]} + \text{[illegible]}_3 + X_4 = k) \text{[illegible]} 4P(X_1 = 0)P(X_2 + \text{[illegible]} + X_4 = k)$$

$$\text{[illegible]} - \phi \text{[illegible]} \sum_{k_1,k_2,k_3,k_4\neq 0} \prod_{i=1} \frac{(1+k_i\alpha)^{k_i-1}}{k_i!} e^{-\lambda(4\text{[illegible]}} \lambda^k,$$

where the [illegible] ove[illegible] all positive [illegible]tegers k_1, k_2, k_3, k_4 a[illegible]g up to k. This simplifies t[illegible]

$$P(X_1 \text{[illegible]} X\text{[illegible]} + X_4 = k) = \text{[illegible]} 4\phi(1-\phi)[\phi + (1-\phi \text{[illegible]}]^2 P(V_1 = k)$$
$$\text{[illegible]} 12 \text{[illegible]} 1-\phi)^2[\phi + \text{[illegible]} - \phi)e^{-\lambda}]P(V_1 + V_2 \text{[illegible]}$$
$$+4\phi(1 \text{[illegible]} P(V\text{[illegible]} + V_2 + V_3 = \text{[illegible]}) + (1-\phi)^4 P(V_1 + V\text{[illegible]} V_3 + V_4 = k).$$

In general, by [illegible] pro [illegible]

$$P(X_1+\cdots+X \ldots -\phi) \ldots e^{-\lambda}]^{n-j-1}P(V_1+\cdots+V_j=k)$$

$$\ldots V_1+ \ldots k).$$

Suppose we h[illegible] then, [illegible] ve this for $n+1$ as follows:

$$P(X_1+\cdots \ldots +1 \ldots)P(X_2+\cdots+X_{n+1}=k)$$

$$+ \ldots)^{k_i-} \ldots {}^{1}e^{-\lambda[(n+1)+k\alpha]}\lambda^k$$

$$=(n+1)[\phi+(\ldots -\phi) \ldots e^{-\lambda}]^{n-j-1}P(V_1+\cdots+V_j=k)$$

$$+(1-\phi)^nP(V_1+ \ldots \frac{k_i\alpha)^{k_i-1}}{k_i!}(1-\phi)^{n+1}e^{-\lambda[(n+1)+k\alpha]}\lambda^k$$

$$=\sum_{j=1}^{n-1} \ldots (1- \ldots P(V_1+\cdots+V_j=k)$$

$$\ldots P(\ldots V_n=k)$$

$$\ldots e^{-\lambda} \ldots +V_n=k)$$

$$+ \ldots)^{k_i-} \ldots {}^{1}e^{-\lambda[(n+1)+k\alpha]}\lambda^k$$

$$=\sum_{j=1}^{n} \ldots (1- \ldots P(V_1+\cdots+V_j=k)$$

$$\ldots +\cdots \ldots {}_{+1}=k)$$

8 Bayes[illegible] for [illegible] Model

Bayesian anal[illegible] Poisson [illegible] es a Gamma density as a conjugate pri[illegible] r λ. [illegible] nalysis of a CMP model has been discussed [illegible]) an[illegible] generalized Poisson model can be found [illegible] (200[illegible] et al. (2006) introduced a zero-inflated [illegible] mode[illegible] cludes the ZIP model as a special case an[illegible] ana[illegible] a generalized linear model setup with cov[illegible] ta p[illegible] zero-inflation parameter ϕ and an appro[illegible] conju[illegible] e power series model. Here we adopt a di[illegible] ve d[illegible] ayesian hierarchical model that leads naturally to a method of two-sample (or multi-sample) comparison. Then we present a conjugate Bayesian analysis for the sum of (conditionally) i.i.d. ZIP random variables.

Suppose we have independent samples from J ZIP populations, say, representing counts of individuals who respond in a particular way to J different treatments administered independently K times. Let Y_{jk} be the k^{th} replicate count observed under the j^{th} treatment ($j = 1, \ldots, J$ and $k = 1, \ldots, K$). We assume $Y_{jk} \sim ZIP(p, \lambda_{jk})$. In other words

$$\mathrm{pr}(Y_{jk} = y) = \phi I(y = 0) + (1 - \phi) P(Y_{jk}^{*} = y) \tag{3}$$

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

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

where ϵ_{jk} is the random residual component following Normal$(0, \sigma_\epsilon^2)$. The use of a residual 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 explanatory variables that were not recorded at first. This is particularly appropriate for Poisson data sets with over-dispersion. The use of residual effects in GLMs is discussed in Sun et al. (2000) and is a special case of the class of generalized linear mixed models (Zeger and Karim, 1991; Breslow and Clayton, 1993). Equation (3) boils down to

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

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., 1995). 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 u and v having the appropriate subscripts. In other words,

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

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

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

However, the specification of the zero-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 $L(y, z \mid \phi, \lambda) =$

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

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

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

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

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

$$\begin{array}{rcl}
Y_{jk} & \sim & ZIP(\phi, \lambda_{jk}) \\
\phi & \sim & \text{Beta}(a, b) \\
\log(\lambda_{jk}) & \sim & \text{Normal}(\beta_j, \sigma_\epsilon^2) \\
\sigma_\epsilon^2 & \sim & \mathcal{IG}(u_{\epsilon,\pi}, v^2_{\epsilon,\pi}) \\
\beta_j, \sigma_\beta^2 & \sim & \mathcal{NIG}(\mu, \sigma^2_\beta, u_{\beta,\pi}, v_{\beta,\pi}) \\
\mu, \sigma_\mu^2 & \sim & \mathcal{NIG}(\mu_0, \sigma^2_\mu, u_{\mu,\pi}, v_{\mu,\pi})
\end{array}$$

Sampling from the posterior distributions 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 $X_1, \ldots, X_n$ are conditionally i.i.d. ZIP (λ, ϕ) 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* 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 $\{X_1, X_2, \cdots, X_n\}$ [illegible] ple of [illegible] from the p[illegible] in (1), which implies that $Z = \sum_{i=1}^{n} X_i$, wi[illegible] e p[illegible] $Z = z | \lambda,$ [illegible] (2), one does the usual posterior calculatio[illegible] ts th[illegible] fo[illegible] wing. The [illegible] onditional or marginal distribution of Z is

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

where N_r is negative bino[illegible] par[illegible] et[illegible] $\left(\alpha, \frac{\mu}{\mu + \cdots}\right.$ [illegible] The posterior joint distribution of λ and ϕ i[illegible] y

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

END

Appendix 1A: Probability [illegible] uncti[illegible] d the GZIP [illegible] tribution

Consider the paper by Consul [illegible] (19[illegible]). They introdu[illegible] a generalized Poisson distribution as follows

$$P(X = x | \lambda_1, \lambda_2 \cdots {}_1 + x \cdots)^x \cdots e^{-(\lambda_1 + x\lambda_2)}$$

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

$$P \cdots \lambda_1, \lambda_2 \cdots = \cdots$$

for $x \geq m$ if $\lambda_1 + m\lambda_2 \leq 0$. Bu[illegible] case, [illegible] is [illegible] od enough [illegible] onsider $\lambda \geq 0$ so that $P(X = x | \lambda_1, \lambda_2)$ is no[illegible]. On[illegible] an[illegible] observe tha[illegible] tting $\lambda_2 = 0$, we can get back the usual Poi[illegible] ributi[illegible].

Consul and Jain referred to [illegible] (1902) [illegible] ving that [illegible] $= x | \lambda_1, \lambda_2) =$ 1. They used one-dimensional [illegible] e's fo[illegible] u[illegible] to prove th[illegible]

$$\phi(z) = \phi(0) + \sum_{x=1}^{\infty} \cdots \frac{d^{x-1}}{dz^{x-1}} [f(\cdots \phi \cdots)] \Big)_{z=0} \left(\cdots\right)^x \tag{8}$$

where $\phi(z) = e^{\lambda_1 z}$ and $f(z) =$ [illegible] w it [illegible] shown tha[illegible]

$$\left(\frac{d^{x-1}}{dz^{x-1}} [f(z) \cdots \right)_{z=0} = \cdots {}_1(\cdots + \lambda_2 x)^{x-1} \cdots$$

so that (8) becomes

$$e^{\lambda_1 z} = \sum_{x=0}^{\infty} \cdots + \lambda_2 \cdots^{x-} \cdots \left(\frac{z}{f(z)}\right)^x \tag{9}$$

This is true for all z. So, subs[illegible] $z = 1$ [illegible] e [illegible] t

$$e^{\lambda_1} = \sum_{x=[illegible]}^{\infty} \frac{1}{x!} [illegible] + \lambda [illegible] \lambda_2 x}$$

which implies

$$1 = \sum_{x=0}^{\infty} \frac{1}{x} [illegible] (\ [illegible] \lambda_2 x)^{[illegible]} [illegible] + \lambda_2 x)}$$

so that the right hand side is [illegible] thi[illegible] [illegible]t $\sum$ [illegible] $= x | \lambda_1, \lambda_2)$. Now substituting $\lambda_1 = \lambda$ and $\frac{\lambda_2}{\lambda_1} = \alpha$ in [illegible] w[illegible] [illegible]up, w[illegible] [illegible]t [illegible]

$$P(X = x | \lambda, [illegible]) = [illegible] 1 + \alpha [illegible] (1+\alpha x)\lambda}$$

which is nothing but $U \sim GP($ [illegible] $, \lambda)$ [illegible] ur no[illegible] [illegible] So, it [illegible]s proved that

$$\sum_{i=0}^{\infty} P(X = x | \lambda, \alpha [illegible] = \sum_{i=[illegible]}^{[illegible]} [illegible] (1 + [illegible] -(1+\alpha x)\lambda} = 1$$

Therefore, if we go to zero-infl[illegible]ed [illegible]ralize[illegible] [illegible] distribution where

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

for $x > 0$, then we get

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

Appendix 1B: Expectatio[illegible] of [illegible]IP [illegible]ion

Here we derive the mean of a z[illegible]to i[illegible]e[illegible]d ge[illegible] li[illegible] Poisson distribution. For this, we go back to the paper [illegible]y C[illegible]l an[illegible] [illegible]hi [illegible]973) and work with (9). Consul and Jain have their ow[illegible] w[illegible] deal[illegible] w[illegible] this as they are allowing negative values of λ_2. In our c[illegible] [illegible]n sin[illegible] [illegible]rentiate (9) and arrive at the following:

$$\lambda_1 e^{\lambda_1 z} = \sum_{x=1}^{\infty} \frac{\lambda_1 ([illegible] + [illegible]}{x - [illegible])} [illegible]^{x-1} [illegible] ^{xz} (1 - \lambda_2 z) \tag{10}$$

which is same as

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

Now putting $z = 1$ once again[illegible]e [illegible]

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

On [illegible] bserve tha [illegible] he right hand side is th [illegible] xpectation of X so t [illegible] $X) =$ [illegible] Substituti [illegible] g $\lambda_2 = 0$, one can see [illegible] at it is the mean o [illegible] usual Po [illegible] stribution. [illegible] Now, coming back to t [illegible] generalized Poisson [illegible] ution U [illegible] $, \lambda)$, we g [illegible] $E(X) = \sum_{x=1}^{\infty} \frac{\lambda^x(1+ [illegible] x)^{x-1}}{}$ $e^{-(1+\alpha x)\lambda} =$ [illegible]. For zer [illegible] ed general [illegible] d Poisson distribution [illegible] he expectation is giv [illegible]

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

Su [illegible] ng $\phi = 0$, [illegible] get the mean of the usu [illegible] generalized Poisson [illegible] ution bac [illegible]

A [illegible] x 1C: Va [illegible] nce of a GZIP distr [illegible] oution

[illegible] ext goal is t [illegible] get the variance of th [illegible] zero-inflated generali [illegible] isson dis [illegible] n. For th [illegible] we again go back to t [illegible] equation (10). We [illegible] tiate it [illegible] re to get

$$\lambda_1^2 [illegible] \sum_{i=2}^{\infty} \frac{\lambda_1(\lambda_1 [illegible] _2 x)^{x-1}}{(x-2)!} z^{x-2} e^{-\lambda_2 x z}(1- [illegible] z)^2 - \sum_{i=1}^{\infty} \frac{\lambda_1(\lambda_1+\lambda_2 [illegible]}{(x-1)} [illegible] ^{x-1} e^{-\lambda_2 x z} \lambda_2(2-\lambda_2 z)$$

wh [illegible] plifies to

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

Su [illegible] ng $z = 1$, [illegible] get

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

No [illegible] second te [illegible] without the negative [illegible] gn is just $E(X)$ m [illegible] d by $\lambda_2($ [illegible] and henc [illegible] 11) becomes

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

Fr [illegible] one show [illegible] hat

$$[illegible] \frac{\lambda_1(\lambda_1+ [illegible] x)^{x-1}}{(x- [illegible])!} e^{-(\lambda_1+\lambda_2 x)} = \lambda_1 [illegible] \frac{\lambda_1}{(1-\lambda_2)^2} + \frac{\lambda_2(2-\lambda [illegible]}{(1-\lambda_2)} [illegible]$$

Now [illegible] nd s[illegible] pt [illegible] ut $E[X(X-1$ [illegible] $]$ in Consul and Jain's set up. Addi[illegible] it,

$$[illegible] X [illegible] \frac{\lambda_1^2}{1-\lambda_2)^2} + \frac{[illegible]_1}{(1 [illegible] \lambda_1)^3}$$

Final[illegible] ting [illegible] $_2$ [illegible] it, we get

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

Subst[illegible] $= 0$ [illegible] he variance of [illegible] he usual Poisson distribution. Now, [illegible] g th[illegible] g[illegible] ized Poisson di[illegible] ribution with $U \sim GP(\alpha,\lambda)$, we ge[illegible]

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

so th[illegible]

$$[illegible] X^2) [illegible] \frac{\lambda^2}{[illegible]} [illegible] \frac{\lambda}{(1-\alpha\lambda)^3};\ V[illegible] X) = \frac{\lambda}{(1-\alpha\lambda)^3}$$

Next, [illegible] g t[illegible] in [illegible] generalized P[illegible] sion distribution, we get

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

Addi[illegible] $\frac{(1 [illegible]}{1 [illegible]}$ [illegible] o[illegible] sides, we get

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

Final[illegible] ing [illegible] $)^2$ [illegible] both sides, we get

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

Subst[illegible] $= 0$ [illegible] he variance of [illegible] he usual generalized Poisson distri[illegible] k.

App[illegible] Sim[illegible] s for biases [illegible] d variances of estimators

For t[illegible] m li[illegible] ethod-of-mom[illegible] ts estimators of λ and ϕ that we m[illegible] Se[illegible] report the res[illegible] lts of some simulation studies regar[illegible] bias[illegible] For a fixed va[illegible] e of $\phi \in (0,1)$, we generated $M =$ [illegible] om [illegible] $n = 100$ fro[illegible] a ZIP(ϕ,λ) distribution with $\lambda =$ [illegible] an[illegible] mputed the bi[illegible] s in the method-of-moments estim[illegible] Fo[illegible] n averaged th[illegible] 1000 bias values and plotted this a[illegible] ag[illegible] A[illegible] r each λ, we co[illegible] puted the sample variance of the 1[illegible] val[illegible] it against λ. [illegible] e repeated this exercise for 9 differ[illegible] of ϕ [illegible] ng 9 'λ vs. av[illegible] age bias' graphs, represented by ... colors, were superimposed. Sin... rly, the 9 'λ vs. var($\hat{\lambda}_{MOM}$)' gr... resented by different colors, were ... erimposed.

... implemented the same procedure ... h the roles of λ and ϕ switched. In ... ords, we now fixed λ at an intege... value between 2 and 10 and in ea... created a 'ϕ vs. average bias' grap... and a 'ϕ vs. var($\hat{\phi}_{MOM}$)' graph ba... $M = 1000$ random samples of size ... $= 100$ drawn from a ZIP(ϕ,λ) di... with $\phi = 0.01, 0.02, \ldots, 0.09$. On... again, each of the two resulting se... aphs (represented by different colo... were superimposed.

... ently, the same procedure was re... ated for the MLEs of λ and ϕ re... 4 sets of graphs which are λ vs. ... ($\hat{\lambda}_{MLE} - \lambda$), λ vs. var($\hat{\lambda}_{MLE}$) ϕ ... $_{MLE} - \phi$) and ϕ vs. var($\hat{\phi}_{MLE}$).

... all of the above were carried out ... ain with $M = 10000$. It is clea... th... LEs of both parameters are asymp... tically unbiased. For $\hat{\lambda}_{MLE}$ the bi... e close to zero even for moderate ... ues of λ such as 6 or 7. In the ca... $_{LE}$, the bias is negligible irrespect... e of ϕ for λ upwards of 5. The va... $\hat{\lambda}_{MLE}$, for the value-range of λ co... lered here, are below 0.4 excep... fo... treme values of ϕ. Also, it shows ... slight increasing trend with λ. As ... variance of $\hat{\phi}_{MLE}$, the only one t... stands out is the case $\lambda = 2$. W... is aberrant behavior was caused by ... computational error needs to be in... d. Moving on to the bias and the va... ance of $\hat{\lambda}_{MOM}$, the bias shows a lot... ctuations in this case than for the ... LE, but it is close to zero excep... pe... he extremely small values of ϕ. T... variance of $\hat{\lambda}_{MOM}$ once agai... sh... ht increasing trend with λ and is b... ow 0.4 over this value-range of λ ex... ome extremely small values of ϕ. T... bias in $\hat{\phi}_{MOM}$ also shows more flu... compared to that in the MLE a... the variance of $\hat{\phi}_{MOM}$ shows a co... adratic trend with the maximum ... urring around the mid-range o... ϕ ... or some small values of λ).

Acknowledgr... nt

R... the second author (Datta) was p... ially supported by the Depart... m... tistics, Texas A & M University, th... gh a National Cancer Institute gr... 0301).

R... nces

[1] ..., Jean-Francois, Biswas, Atanu, 2... 3: A Bayesian analysis of zero... generalized Poisson model. *Com... ational Statistics & Data Anal...* **... 2**, p.37-46.

[2] ... o, J. del, Perez-Casany, Marta, 20...: Overdispersed and underdis... Poisson generalizations. *Journal ... Statistical Planning and Infer...* **... 34**, p.486-500.

[3] ... A.C.,1960: Estimating the param... ers of a modified Poisson distri... *J. Amer. Statist. Assoc.*, **55**, p.1... 143.

[4] Consul, P.C., [illegible]ain, G.C., 1973: A Generalizat[illegible] the Poiss[illegible]tion. *Technometrics*, **15(4)**, p.791-799.

[5] Cox, D.R., 1983: Some remarks on overdisp[illegible]i. *Biometr*[illegible] p.269-274.

[6] Dahiya, R.C., Gross, A.J., 1973: Estimating the [illegible] class from [illegible] Poisson sample. *J. Amer. Statist. Assoc.*, **68**, p[illegible]733.

[7] Ghosh, S.K., Mukhopadhyay, P., Lu, J.C., 2006 [illegible]esian analy[illegible] inflated regression models. *J. Statist. Plan. Inf* [illegible], p.1360-1[illegible]

[8] Gupta, P.L., Gupta, R.C., Tripathi, R.C., 199[illegible] Analysis of ze[illegible] count data. *Comput. Statist. Data Anal.*, **23**, p[illegible]547.

[9] Heilborn, D.C., 1989: Generalized linear mode[illegible] altered zer[illegible] ities and over dispersion in count data. *Techn*[illegible] *Report*, Dep[illegible] Epidemiology and Biostatistics, University of C[illegible]fornia, San F[illegible]

[10] Heilborn, D.C., Gibson, D.R., 1990: Shared n[illegible]e use and he[illegible] concerning AIDS: regression modeling of zer[illegible]y count d[illegible] session. *Proceedings of the Sixth International Conference on* [illegible] Francisco, CA.

[11] Kadane, J.B., Shmueli, G., Minka, T.P., Borle, S[illegible] Boatwrigh[illegible] Conjugate analysis of the Conway-Maxwell-Poi[illegible] distributio[illegible] *Analysis*, **1(2)**, p.363-374.

[12] Kemp, A.W., 1986: Weighted discrepancies and maximum like[illegible] mation for discrete distribution. *Commun. Stati*[illegible] *A - Theory a*[illegible] **15**, p.783-803.

[13] Kokonendji, C.C., Mizere, D., Balakrishnan, N[illegible]008: Connect[illegible] Poisson weight function to overdispersion and [illegible]dispersion. [illegible] *PLan. Inf.*, **138(5)**, p.1287-1296.

[14] Lambert, D., 1992: Zero inflated Poisson regression with an ap[illegible] defects in manufacturing. *Technometrics*, **34**, p[illegible]1[illegible]

[15] Li, C.S., Lu, J.C., Park, J., Kim, K.M., Brinkle[illegible] P.A., Peterso[illegible] A multivariate zero-inflated Poisson distribution and its inferen[illegible] *metrics*, **41**, p.29-38.

[16] Martin, D.C., Katti, S.K., 1965: Fitting of so[illegible] contagious di[illegible] to some available data by the maximum likeliho[illegible] method. *Bio*[illegible] p.34-48.

[17] Przyborowski, J., Wilenski, H., 1940: Homog[illegible]y of results [illegible] samples from Poisson Series. *Biometrika*, **31**, p[illegible]3[illegible]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.