Research Journal of Mathematical and Statistical Sciences E-ISSN 2320-6047
Vol. **4(9),** 1-3, October **(2016)** *Res. J. Mathematical and Statistical Sci.*

Short Communication

A Characterization of the Zero-One Inflated Negative Binomial Distribution

Rafid Saeed Abdulrazak Alshkaki
Ahmed Bin Mohammed Military College, Doha, Qatar
rafid@abmmc.edu.qa

Available online at: www.isca.in, www.isca.me
Received 15th September 2016, revised 29th September 2016, accepted 7th October 2016

Abstract

In this paper, we introduce a characterization of the zero-one inflated negative binomial distributions through a linear differential equation of its probability generating function.

Keywords: Negative Binomial Distribution, Zero-One Inflated Negative Binomial Distribution, Probability Generating Function, Linear Differential Equation.

Introduction

The negative binomial distribution (NBD), which is a well known discrete distribution, has so many real live applications. It is the distribution in which we are looking for the number of failures observed until a specific number of successes occur[1]. Banik and Kibria[2] reported recent real-life applications of the standard NBD and its inflated form. In particular, Sharma and Landge[3] used the zero inflated negative binomial (ZINB) regression for modeling heavy vehicle crash rate on Indian rural highway.

Saengthong et al[4] introduce the ZINB – crack distribution, consists of a mixture of Bernoulli distribution and NBD, which is an alternative distribution for the excessive zero counts and over dispersion, and studied some of its properties and parameter estimates. Fang et al[5] considered a hierarchical regression-based approach using a ZINB mixed model to evaluate associations with disease state while adjusting for potential confounders for two organisms of interest from a study of human microbiota sequence data in oesophagitis.

Preisser et al[6] proposed a marginalized ZINB regression model to the evaluation of a school-based fluoride mouth rinse program on dental caries in 677 children. Recently, Suresh et al[7] characterized the ZINBD via a linear differential equation of its probability generating function (pgf). In this paper, we introduce in Section 2, the definition of the NBD and its zero-one inflated form with their pgf, followed in Section 3 by a characterize of the zero and zero-one inflated negative binomial distribution (ZOINBD) through a linear differential equation.

The Negative Binomial Distributions and its Zero-One Inflated Form

Let $k > 0$ *and* $p \in (0,1)$, then the discrete random variable (rv)X having probability mass function (pmf); is said to have a NBD with parametersk and p. We will denote that by writing $X \sim \mathrm{NBD}(k, p)$. Johnson et al[1] reported other forms and parameterizations of the NBD.

$$P(X = x) = \begin{cases} \binom{k+x-1}{x} p^x (1-p)^k, & x = 0, 1, 2, 3, \dots \\ 0, & \end{cases} \quad (1)$$

Otherwise

The pgf of the rv X, $G_X(t)$, is given by;

$$G_X(t) = E(t^X)$$

$$= (1-p)^k \sum_{x=0}^{\infty} \binom{k+x-1}{x} p^x t^x$$

$$= \left(\frac{1-p}{1-pt}\right)^k \quad (2)$$

Let $X \sim \mathrm{NBD}(k, p)$ as given in (1), let $\alpha \in (0,1)$ be a proportion of zero added to the rv X, and let $\beta \in (0,1)$ be an extra proportion added to the proportion of ones of the rv X, such that $0 < \alpha + \beta < 1$, then the rv Z defined by;

$$P(Z = z) = \begin{cases} \alpha + (1-\alpha-\beta)(1-p)^k, & z = 0 \\ \beta + k(1-\alpha-\beta)p\,(1-p)^k, & z = 1 \\ (1-\alpha-\beta)\binom{k+z-1}{z} p^z (1-p)^k, & z = 2, 3, 4, \dots \\ 0, & \end{cases} \quad (3)$$

Otherwise

is said to have a ZOINBD, and we will denote that by writing $Z \sim \text{ZOINBD}(k, p; \alpha, \beta)$.
Note that, if $\beta \to 0$, then (3) reduces to the form of the zero-inflated NBD. Similarly, the case with $\alpha \to 0$ and $\beta \to 0$, reduces to the standard case of NBD.

Characterization of the Zero-One Inflated Negative Binomial Distribution

The pgf of the rv $Z\sim \text{ZOINBD}(k, p; \alpha, \beta)$, can be shown to be;

$$G_Z(t) = \alpha + \beta t + (1-\alpha-\beta)G_X(t)$$

$$= \alpha + \beta t + (1-\alpha-\beta)\left(\frac{1-p}{1-pt}\right)^k \quad (4)$$

Theorem 1: The discrete rv Z taking non-negative integer values, has a ZOINBD if its pgf, $G(t)$ satisfies for some arbitrary number c and non-zeros numbers a, b, f and h, satisfying that $bf< 0$, that;

$$(a+bt)\frac{\partial}{\partial t}G(t) = c + ht + fG(t) \quad (5)$$

Proof: Assume first that $f > 0$, then $b < 0$, and therefore, without loss of generality, we can assume that $b = -1$, hence (5) becomes;

$$(a-t)\frac{\partial}{\partial t}G(t) = c + ht + fG(t) \quad (6)$$

Now, (6), with using of the derivative reserve product rule, can be rewritten in the following equivalent form;

$$\frac{\partial}{\partial t}\left[(a-t)^f G(t)\right] = (a-t)^{f-1}(c+ht)$$

Hence,

$$(a-t)^f G(t) = \int (a-t)^{f-1}(c+ht)\,dt \quad (7)$$

Now by making the substituting $x= a - t$in the integral given in (7) and evaluated it, we get that;

$$(a-t)^f G(t) = -\left(c\frac{(a-t)^f}{f} + ah\frac{(a-t)^f}{f} - h\frac{(a-t)^{f+1}}{f+1}\right) + k$$

Where: k is arbitrary constant. Hence;

$$G(t) = -\left(\frac{c}{f} + \frac{ah}{f} - h\frac{(a-t)}{f+1}\right) + k(a-t)^{-f}$$

Or equivalently;

$$G(t) = -\left(\frac{c+cf+ah+fht}{f(f+1)}\right) + k(a-t)^{-f}$$

Since 1 = G(1);

$$1 = -\left(\frac{c+cf+ah+fh}{f(f+1)}\right) + k(a-1)^{-f}$$

And hence;

$$k = \left[1 + \left(\frac{c+cf+ah+fh}{f(f+1)}\right)\right](a-1)^f$$

Therefore;

$$G(t) = -\left(\frac{c+cf+ah+fht}{f(f+1)}\right) + \left[1 + \left(\frac{c+cf+ah+fh}{f(f+1)}\right)\right](a-1)^f(a-t)^{-f}$$

Or equivalently,

$$G(t) = -\left(\frac{c+cf+ah+fht}{f(f+1)}\right) + \left[1 + \left(\frac{c+cf+ah+fh}{f(f+1)}\right)\right]\left(\frac{a-t}{a-1}\right)^{-f}, \quad (8)$$

Where;

$$p = \frac{1}{a} \quad (9)$$

$$\alpha = -\frac{c}{f} - \frac{ah}{f(f+1)} \quad (10)$$

$$\beta = -\frac{h}{f+1} \quad (11)$$

Thus, G(t) given in (8) can be written in the following form;

$$G_Z(t) = \alpha + \beta t + (1-\alpha-\beta)\left(\frac{1-pt}{1-p}\right)^{-f}, \quad (12)$$

Which is the same form given (4).

Now we need to check possible values of the parameters of (12); namely, p, f, α and β.

If $a > 1$, then p, given by (9) satisfies that $0 < p < 1$. Therefore, if $0 < \alpha < 1, 0 < \beta < 1$ *and* $0 < \alpha + \beta < 1$, then the G(t) is the pgf of $\text{ZOILSD}(f, p; \alpha, \beta)$.

If $-(1+f) < h < 0$, and recall that $f > 0$, then $0 < -\frac{h}{f+1} < 1$, hence $0 < \beta < 1$. If; $-f - \frac{ah}{f+1} < c < -\frac{ah}{f+1}$, then $0 < -\frac{c}{f} - \frac{ah}{f(f+1)} < 1$, and hence $0 < \alpha < 1$. If c is also satisfying that $-f - \frac{h(a+f)}{f+1} < c < -\frac{h(a+f)}{f+1}$, then $0 < -\frac{c+cf+ah+fh}{f(f+1)} < 1$ and hence $0 < \alpha + \beta < 1$. Since the intervals $(-f - \frac{ah}{f+1}, -\frac{ah}{f+1})$ and $(-f - \frac{h(a+f)}{f+1}, -\frac{h(a+f)}{f+1})$ is overlapping and their intersection is $(-f - \frac{h(a+f)}{f+1}, -\frac{ah}{f+1})$, it follows that if $-f - \frac{h(a+f)}{f+1} < c < -\frac{ah}{f+1}$, then $0 < \alpha < 1$, $0 < \beta < 1$ *and* $0 < \alpha + \beta < 1$, therefore G(t) is the pgf of $\text{ZOILSD}(f, p; \alpha, \beta)$.

Now, if $f< 0$, then $b > 0$, and hence we can divided both sides of (5) by (-b) to get;

$$(a_1 - t)\frac{\partial}{\partial t}G(t) = c_1 + h_1 t + f_1 G(t) \quad (13)$$

Where $a_1 = -a, c_1 = -c, h_1 = -h$and $f_1 = -f$.Since$f_1 > 0$, then (13) has the same form as (6) with a positive coefficient of G(t), and therefore gives the same conclusion as above; that is, G is the pgf of a ZOILSD.

Finally, if c = 0, then we arrive simply, either by solving (5) directly or by letting $c \to 0$ in the above proof, which is straightforward, to the same conclusion that the pgf is given (12) with p and β given by (9) and (11), respectively, and that $\alpha = -\frac{ah}{f(f+1)}$, and on same lines as above, when $f > 0$ and $a > 1$, it can be seen that if $-\frac{f(f+1)}{f+a} < h < 0$, then $0 < \alpha < 1,\ 0 < \beta < 1$ and $0 < \alpha + \beta < 1$.This completes the proof.

Theorem 1 and the derivation of (4) lead to the following.

Theorem 2: Let Z be a discrete rv taking non-negative integer values, then $Z \sim \mathrm{ZOINBD}(f, p; \alpha, \beta)$, for some non-zero f, p, α and β if and only if its pgf satisfying (5) for some arbitrary number cand non-zeros numbers a, b, f and h, satisfying that $bf < 0$, that;

$$(a + bt)\frac{\partial}{\partial t}G(t) = c + ht + fG(t).$$

Theorem 2 leads to the following conclusion obtained by Suresh et al[7], that characterize the ZINBD.

Theorem 3: Let Z be a discrete rv taking non-negative integer values, then $Z \sim \mathrm{ZINBD}(f, p; \alpha)$, for some non-zero f, p and αif and only if its pgf G(t) satisfying

$$(a + bt)\frac{\partial}{\partial t}G(t) = c + fG(t)$$

For some arbitrary number c and non-zeros numbers a, b and f, satisfying that $bf < 0$.

Proof: Just let $h \to 0$ in Theorem 2.

Conclusion

We introduced a characterization of the zero-one inflated negative binomial distributions through a linear differential equation of its probability generating function. We would propose an extension of these results to other forms and distributions.

References

1. Johnson N.L., Kemp A.W. and Kotz S. (2005). Univariate Discrete Distributions. Third Edition, John Wiley and Sons, New Jersey.

2. Banik S. and Kibria B.M.G. (2009). On Some Discrete Distributions and their Applications with Real Life Data. *Journal of Modern Applied Statistical Methods*, 8(2), 423-447.

3. Sharma A.K. and Landge V.S. (2013). Zero Inflated Negative Binomial For Modeling Heavy Vehicle Crash Rate On Indian Rural Highway. *International Journal of Advances in Engineering & Technology*, 5(2), 292-301.

4. Saengthong P., Bodhisuwan W. and Thongteeraparp A. (2015). The Zero Inflated Negative Binomial – Crack Distribution: Some Properties And Parameter Estimation. *Songklanakarin J. Sci. Technol*, 37 (6), 701-711.

5. Fang R.,Wagner B.D., Harris J.K. and Fillon S.A. (2016). Zero-inflated negative binomial mixed model: an application to two microbial organisms important in oesophagitis. *Epidemiology and Infection*, 144(11), 2447-2455. DOI: http://dx.doi.org/10.1017/S0950268816000662

6. Preisser J.S., Das K., Long D.L. and Divaris K. (2015). Marginalized zero-inflated negative binomial regression with application to dental caries. *Stat Med*. 35(10), 1722-35. DOI: 10.1002/sim.6804

7. Suresh R., Nanjundan G., Nagesh S. and Pasha S. (2015). On a Characterization of Zero-Inflated Negative Binomial Distribution. *Open Journal of Statistics*, 5, 511-513. http://dx.doi.org/10.4236/ojs.2015.56053.