[Volume 3 issue 11 November 2015]
PageNo.1187-1190 ISSN : 2320-7167

INTERNATIONAL JOURNAL OF MATHEMATICS AND COMPUTER RESEARCH

A Characterization of Zero-Inflated Binomial Model

[1]G. Nanjundan , [2]Sadiq Pasha
Department of Statistics, Bangalore University, Bangalore 560056, India

Abstract:
Zero-inflated probability models have been applied to a variety of situations in the recent years. Especially they are found very useful in count regression modeling. The zero-inflated binomial model is characterized in this paper through a differential equation which is satisfied by its probability generating function.

Keywords: Zero-inflated binomial model, probability generating function, linear differential equation.

Introduction:
A subfamily of power series distributions, whose probability generating function (pgf) *f(s)* satisfies the differential equation $(a+bs)f'(s)=cf(s)$ with *f'(s)* being the first derivative of *f(s)*, has been characterized by Nanjundan (2010). Binomial, Poisson, and negative binomial distributions are members of this family. Also, Nanjundan and Sadiq Pasha (2015) have characterized zero-inflated Poisson distribution through a linear differential equation. Along the same lines, Nagesh et al (2015a, 2015b) have characterized zero-inflated geometric distribution and further extended the characterization to zero-inflated negative binomial distribution. In this paper, the zero-inflated binomial distribution is chracterized using a differential equation satisfied by its pgf.
A random variable *X* is said to have a zero-inflated binomial distribution if its probability mass function is given by

$$p(x) = \begin{cases} \varphi + (1-\varphi)q^n, x = 0 \\ (1-\varphi)\binom{n}{x} p^n q^{n-x}, \ x = 1, 2, \ldots, n \end{cases} \tag{1}$$

$$= \varphi\, p_o(x) + (1-\varphi)p_1(x), \ 0 < \varphi < 1,$$

where $p_0(x) = \begin{cases} 1, & x = 0 \\ 0, & x \neq 0 \end{cases}$ and $p_1(x) = \binom{n}{x} p^x q^{n-x}, x = 1, 2, \ldots, n;$

$$0 < p < 1, \ p + q = 1.$$

Hence the distribution of *X* is a mixture of a distribution degenerate at zero and a binomial distribution. The probability generating function (pgf) of *X* is given by

$$f(s) = E(S^X)$$

$$= \sum_{x=0}^{\infty} p(x)s^x, 0 < s < 1$$

$$f(s) = \varphi + (1-\varphi)(q+ps)^n. \tag{2}$$

Characterization:

The following theorem characterizes a random variable X having a zero-inflated binomial distribution.

Theorem: Let X be a random variable taking only a finite number of non-negative integer values $0, 1, \ldots, n$ with $n \geq 1$. Then X has a zero-inflated binomial distribution if and only if its pgf $f(s)$ is such that

$$f(s) = a + b(c+ds)f'(s) \tag{3}$$

where $a \neq 0$, b, c, d are con stants and $f'(s)$ is the derivative of $f(s)$.

Proof: 1) Suppose that X has a zero-inflated binomial distribution with the probability mass function (pmf) specified in (1). On differentiating its pgf, we get

$$f'(s) = (1-\varphi)np(q+ps)^{n-1}.$$

Note that $f'(s)$ satisfies (3) with $a = 1-\varphi, b = \frac{1}{np}, c = q,$ and $d = p$.

2) Suppose that the pgf $f(s)$ of X satisfies (3). Writing the differential equation (3) as $y = a + b(c+dx)\frac{dy}{dx}$, we see that $\frac{dy}{y-a} = \frac{1}{bd}\frac{d}{(c+dx)}dx$. Integrating both sides,we obtain

$log(y-a) = \frac{1}{bd}log(c+dx) + constant$. That is $y = k\,log(c+dx)^{\frac{1}{bd}}$, where k is a constant. Hence the solution of the differential equation (2) becomes

$$f(s) = a + k(c+ds)^{\frac{1}{bd}}.$$

Since $f(1) = 1$, we get $k = (1-a)(c+d)^{-\frac{1}{bd}}$. Further, either $b \to 0$ or $d \to 0$ implies that $f(s) \to 0$ and hence $b, d \neq 0$. Therefore, (3) can be writtten as

$$f(s) = a + (1-a)(c+d)^{-\frac{1}{bd}}(c+ds)^{\frac{1}{bd}}. \tag{4}$$

If $c = 0$, then $f'(s) = (1-a)\frac{1}{bd}s^{\frac{1}{bd}-1}$ and $f'(0) = 0$. Since $f(s)$ is a pgf, $f'(0) = P(X=1) > 0$. Hence $c \neq 0$. Since X takes the values $0, 1, \ldots, n$, its pgf is such that

$$f(s) = p_0 + p_1 s + p_2 s^2 + \ldots + p_n s^n, \tag{5}$$

where $p_x = P(X=x)$. Note that $f(s)$ in (4) matches with that in (5) if and only if $\frac{1}{bd}$ is a positive integer.

Take $\frac{1}{bd} = m$. Then the equation (4) can be expressed as

$$f(s) = a + (1-a)(c+d)^{-m}(c+ds)^m.$$

Note that $(c+ds)^m = f_1(s)$ on the RHS of $f(s)$ is the pgf of a binomial distribution and $f_0(s) = 1$ is the pgf of a random variable degenerate at 0. Therefore $f(s) = af_0(s) + (1-a)(c+d)^{-m} f_1(s)$ can be identified as a convex combination of these two pgfs. Hence $c+d = 1$and the pgf of X becomes

$$f(s) = a + (1-a)(c+ds)^m. \tag{6}$$

Hence *f(s)* of (6) satisfies (2) with $a = \varphi, \ c = q, d = p,$ and $m = n$ and this completes the proof of the theorem.

References:

1. Nanjundan, G., 2010. A characterization of the members of a subfamily of power series distributions, Applied Mathematics, Vol. 2, 750 – 51.
2. Nanjundan, G. and Pasha, S., 2015. A Note on the characterization of zero-inflated Poisson model, Open Journal of Statistics, 5, 140 – 142. http://dx.doi.org/10.4236/ojs. 2015.52017.
3. Nagesh, S., Nanjundan, G., Suresh, R., and Sadiq Pasha, 2015a. A characterization of zero-inflated geometric model, Int. Journal of Mathematical Trends and Technology, Vol. 23, 71 – 73.
4. Suresh ,r.,nanujdan ,g., nagesh ,s.and pasha,s.,2015b ,on a characterization of zero-inflated negative binomial distribution , open journal of statics,5,511-513 http://dx.doi.org/10.4236/ojs.2015.56053