A Characterization of Zero-inflated Geometric Model

Nagesh, S[1], G. Nanjundan[2], Suresh, R[3], and Sadiq Pasha[4]

[1, 3]*Assistant Professor, Department of Statistics, Karnatak University, Dharwad-580003, India*
[2,] *Professor,* [4] *Research Scholar, Department of Statistics, Bangalore University, Bangalore 560056, India*

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

Keywords: *Zero-inflated geometric model, probability distribution, linear differential equation.*

I. Introduction

Nanjundan (2011) has characterized a subfamily of power series distributions whose pgf *f(s)* satisfies the differential equation $(a+bs)f'(s)=cf(s)$, where $f('s)$ is the first derivative of *f(s)*. This subfamily includes binomial, Poisson, and negative binomial distributions. Also, Nanjundan and Sadiq Pasha (2015) have characterized zero-inflated Poisson distribution through a differential equation. In this paper, zero-inflated geometric distribution is characterized by a differential equation satisfied by its pgf.

A random variable *X* is said to have a zero-inflated geometric distribution, if its probability mass function (pmf) is given by

$$p(x) = \begin{cases} \varphi+(1-\varphi)p, & x=0 \\ (1-\varphi)pq^x, & x=1,2,3,\ldots \end{cases} \quad (1.1)$$

$$= \varphi p_0(x) + (1-\varphi)pq^x, \ x=1,2,3,\ldots$$

where $p_0(x) = \begin{cases} 1, & x=0 \\ 0, & x\neq 0 \end{cases}$ and

$p_1(x) = pq^x$, $x = 0,1,2,\ldots$ with $0 < p < 1$ and $p+q = 1$. Thus, the distribution of *X* is a mixture of a distribution degenerate at zero and a geometric distribution.

The pgf of *X* is given by

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

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

$$= \varphi+(1-\varphi)p + (1-\varphi)p\sum_{x=1}^{\infty}(qs)^x$$

$$f(s) = \varphi+(1-\varphi)p + (1-\varphi)\frac{pqs}{1-qs}. \quad (1.1)$$

II. Characterization

A random variable having a zero-inflated geometric distribution is characterized in this section. **Theorem:** A non-negative integer valued random variable *X* with $0<P(X=0)<1$ and the pgf *f(s)* has a zero-inflated geometric distribution if and only if

$$f(s) = a + s(1-bs)f'(s), \quad (2.1)$$

where *a* and *b* are constants.

Proof:

1) Suppose that *X* has a zero-inflated geometric distribution with the pmf specified in (1.1). Then its pgf is

$$f(s) = \varphi+(1-\varphi)p + (1-\varphi)\frac{pqs}{1-qs}.$$

On differentiating w.r.t. s, we get

$$f'(s) = (1-\varphi)\frac{pq}{(1-qs)^2}.$$

By expressing $f(s)$ in terms of $f'(s)$, we see that

$$f(s) = \varphi + (1-\varphi)p + (1-\varphi)\frac{pqs}{1-qs}$$

$$= \varphi + (1-\varphi)p + s(1-qs)\frac{(1-\varphi)pq}{(1-qs)^2}$$

$$f(s) = \varphi + (1-\varphi)p + s(1-qs)f'(s).$$

Therefore *f(s)* satisfies the differential equation

$$f(s) = a + s(1-bs)f'(s),$$

where $a = \varphi + (1-\varphi)p$ and $b = q$.

2) Suppose that the pgf *f(s)* of *X* is such that

$$f(s) = a + s(1-bs)f'(s).$$

If *a = 0*, then $f(0) = P(X=0) = 0$, which is not possible because *P(X = 0) > 0*.

If *b = 0*, then $f(s) = a + sf'(s)$ and $f(1) = a + f'(1)$. Hence $a = 1 - E(X)$. Since *X* is a non-negative integer valued, it is very much possible that *E(X) > 1* and in that case *a = P(X = 0) < 0* which is against our assumption. Therefore $b \neq 0$.

Now the differential equation (2.1), for convenience, can be written as

$$\frac{dy}{dx} = \frac{y-a}{x(1-bx)},$$ with *y = f(s)* and *x = s*.

$$\Rightarrow \frac{dy}{y-a} = \frac{dx}{x(1-bx)}.$$

By resolving $\frac{1}{x(1-bx)}$ into partial fractions, we get

$$\frac{dy}{y-a} = \frac{dx}{x} + \frac{bdx}{1-bx}.$$

$$\Rightarrow \int \frac{dy}{y-a} = \int \frac{dx}{x} - \int \frac{bdx}{1-bx}$$

Hence

$$\log(y-a) = \log x - \log(1-bx) + \log c,$$

where *c* is a constant.

$$\Rightarrow y = \frac{cx}{1-bx} + a.$$

The solution of the differential equation (2.1) becomes

$$f(s) = \frac{cs}{1-bs} + a. \quad (2.2)$$

Since $f(1) = 1$, $\frac{c}{1-b} + a = 1$. Thus $c = (1-a)(1-b)$.

We shall now extract the probabilities $P(X=k) = p_k,\ k = 0,1,2,\ldots$ using the above solution. Since $f(s)$ is a pgf, $p_0 = f(0)$ and $p_k = \frac{f^{(k)}(0)}{k!},\ k = 1,2,3,\ldots,$ where $f^{(k)}(s)$ is the k-th derivative of *f(s)*.

Note that

$$f^{(k)}(s) = k!(1-a)(1-b)b^{k-1}(1-bs)^{k+1}, k = 1, 2, 3, \ldots.$$

Hence we get

$p_0 = a$ and

$p_k = (1-a)(1-b)b^{k-1},\ k = 1, 2, 3, \ldots.$

Therefore the pmf of X is given by

$$p_k = \begin{cases} a,\ k = 0 \\ (1-a)(1-b)b^{k-1},\ k = 1, 2, 3, \ldots \end{cases}$$

Clearly, $0 < a = P(X=0) < 1$ and hence $0 < p_1 = (1-a)(1-b) < 1$. Thus $0 < 1-b < 1$ and we can take $a = \alpha + (1-\alpha)(1-b)$, with $0 < \alpha < 1$.

Then

$$p_k = \begin{cases} \alpha + (1-\alpha)(1-b),\ k = 0 \\ (1-\alpha)(1-b)b^k,\ k = 1, 2, 3, \ldots \end{cases}$$

Therefore X has the pgf specified in (1.1) with $\alpha = \varphi$ and $1-b = p$.

REFERENCES

[1] G. Nanjundan, "A characterization of the members of a subfamily of power series distributions," *Applied Mathematics*, Vol. **2**, pp. 750–751, 2010.

[2] G. Nanjundan and Sadiq Pasha, "A note on the characterization of zero-inflated Poisson distribution," *Open Journal of Statistics*, Vol. **5**, pp. 140–142, 2015.