Advances and Applications in Mathematical Sciences
Volume 20, Issue 10, August 2021, Pages 2291-2298

NOTE ON THE CHARACTERIZATION OF ZERO-INFLATED CHI-SQUARE DISTRIBUTION

SADIQ PASHA and T. RAVEENDRA NAIKA

Department of Statistics
Bangalore University
Bengaluru 560056, India

Assistant Professor
Department of Statistics
Maharani's Science College for Women
Bengaluru 560 001, India
E-mail: sadiqbujc@gmail.com
raviisec@yahoo.co.uk

Abstract

Probability distributions are useful to model random phenomena. New random experiments are conducted and novel data sets are encountered. In turn, new distributions emerge. Zero-inflated discrete and zero-inflated continuous distributions are such cases. They have found many applications in the recent past. Characterization of a distribution is studying a unique property enjoyed by it. A probability distribution can be characterized through various methods. Zero-inflated discrete distributions are characterized through differential equation [see Nanjundan and Sadiq Pasha (2018)]. In this paper, zero-inflated chi-square distribution is characterized via a differential equation satisfied by its moment generating function.

1. Introduction

Characterization of a distribution is studying a unique property enjoyed by it. A probability distribution can be characterized through various methods. Lack of memory property characterizes exponential distribution in the continuous case and geometric distribution in the discrete case. Distributions have been characterized through the lower bounds of certain

2010 Mathematics Subject Classification: 62K20.
Keywords: Zero-inflated continuous distributions, moment generating function, linear differential equation.
Received March 29, 2021; Accepted April 12, 2021

functions of their variances [see Srivastava and Sreehari [11]. Also, life time distributions are characterized by mean residual life function [see Sankaran and Unnikrishnan [10]]. Further, the characterization of continuous distributions by truncated moments has been addressed by Ahsanullah et al. [1].

It is highly infeasible in a paper like this to summarize all types of characterizations of distributions. Since zero-inflated chi-square distribution is characterized through a differential equation, a brief review of literature in this direction is presented.

If a general insurance policy is under the detectable agreement, a claim will not be recorded and honored unless the loss exceeds a prescribed detectable limit. Also, if an insured does not claim for any loss, there will be a bonus for the next year's premium. Naturally there is a huge hunger for such a bonus.

Hence the policy holders do not report all their losses in order to gain bonus for the next year. Therefore, a data set on the claim counts of an insurance portfolio will have excess zeros. Yip and Yau [13] have used zero-inflated Poisson and zero-inflated negative binomial distributions to model claim count data, Boucher et al. [2] have employed zero-inflated distributions to model insurance panel count data. They have also illustrated zero-inflated models based on the data on the claims reported to a Spanish insurance company.

Nanjundan [5] has characterized a subfamily of power series distributions through a differential equation satisfied by the probability generating functions (pgfs) of the distributions. Nanjundan and Sadiq Pasha [7-8] have characterized zero-inflated Poisson and zero-inflated binomial distributions through a linear differential equation. Suresh et al [12] have identified a linear differential equation that characterizes zero-inflated negative binomial distribution. Along the same lines, Nanjundan and Sadiq Pasha [9] have characterized a subfamily of zero-inflated power series distributions via a differential equation satisfied by their pgfs.

Continuous distributions play an important role in statistical theory and applications. Many entities that we encounter in real life can be modeled by continuous distributions.

Such distributions are found useful in many fields such as Economics, Medical, Insurance, and etc. Also zero-inflated random variables are common in economic surveys. In a real situation, the response variable can take any non-negative value but has positive probability of zero outcomes. This results in zero-inflated continuous distributions. For example, when each observation is a record of the rainfall in a day, it takes positive values and also the value zero with positive probability because some days may have no rain. This type of a situation can be modeled by a zero-inflated continuous distribution.

In the context of both life and nonlife insurance, the claim sizes are usually modeled by continuous distributions with the positive support. But due to the hunger for bonus on no claim, a data set on the claim sizes of an insurance portfolio will have excess zeros. This leads to distributions which are mixtures of a distribution degenerate at zero and another one that is continuous with positive support. Such distributions are called perturbed continuous distributions or zero-inflated continuous distributions [see Nanjundan [4], Mills [3]]. In this paper, we are interested in zero-inflated chi-square distribution.

In this paper, we study the characterization of zero-inated chi-square distribution. This is mixtures of a degenerate distribution at zero and chi-square distribution with positive support.

2. Zero-inflated Chi-square Distribution

Let X be a random variable with $P(X = 0) = \varphi,\ 0 < \varphi < 1$ and have a probability density function (pdf) $(1-\varphi)f(x)$ for $x > 0$, where $f(x) = \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}},\ n > 0$. Then, the distribution function of X is given by

$$F(x) = P(X \le x) = \begin{cases} 0, & x < 0, \\ \varphi, & x = 0, \\ (1-\varphi)\int_0^x \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} y^{\frac{n}{2}-1} e^{-\frac{y}{2}} dy, & x > 0. \end{cases} \tag{2.1}$$

Thus the probability density function of X is a zero-inflated chi-square distribution with parameters $\varphi, \frac{n}{2}$, and $\frac{1}{2}$.

3. Moment Generating Function

The moment generating function (mgf) of X is given by

$$M(t) = \int_0^\infty e^{tx}\, dF(x),\ -\infty < t < \infty$$

$$= \int_0^\infty e^{tx} f(x)\, dx$$

$$= \varphi + (1-\varphi)\frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\int_0^\infty e^{tx} x^{\frac{n}{2}-1} e^{-\frac{x}{2}} dx$$

$$= \varphi + (1-\varphi)\frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\int_0^\infty x^{\frac{n}{2}-1} e^{-x\left(\frac{1}{2}-t\right)} dx$$

$$M(t) = \varphi + (1-\varphi)\frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\int_0^\infty \frac{u^{\frac{n}{2}-1}}{\left(\frac{1}{2}-t\right)^{\frac{n}{2}}} e^{-u} du$$

$$M(t) = \varphi + (1-\varphi)\frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)\left(\frac{1}{2}-t\right)^{\frac{n}{2}}}\int_0^\infty u^{\frac{n}{2}-1} e^{-u} du$$

$$= \varphi + (1-\varphi)\frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)\left(\frac{1}{2}-t\right)^{\frac{n}{2}}}\Gamma\left(\frac{n}{2}\right)$$

$$= \varphi + (1-\varphi)\frac{1}{2^{\frac{n}{2}}\left(\frac{1-2t}{2}\right)^{\frac{n}{2}}}$$

$$M(t) = \varphi + (1 - \varphi)\frac{1}{(1 - 2t)^{\frac{n}{2}}}.$$

4. Characterization

A characterization of a zero-inflated chi-square distribution is given by the following Theorem.

Theorem. *Let X be a random variable with the pdf* $f(x)$ *as specified in* (2.1). *Then, X has a zero-inflated chi-square distribution if and only if its mgf is such that*

$$M(t) = a + \frac{1}{b}(1 - 2t)M'(t), \tag{4.1}$$

where $0 < a < 1,\ b > 0$ *are constants,* $t < \frac{1}{2}$, *and* $M'(t)$ *is the derivative of* $M(t)$.

Proof. It is straight forward to verify that the mgf of zero-inflated chi-square distribution satisfies (4.1).

(1) Suppose that *X* has zero-inflated chi-square distribution with the distribution function specified in (2.1). Then, the mgf of *X* is given by

$$M(t) = \varphi + \frac{(1 - \varphi)}{(1 - 2t)^{\frac{n}{2}}}.$$

On differentiating the mgf with respect to *t*, we get

$$M'(t) = (1 - \varphi)\frac{n}{(1 - 2t)^{\frac{n}{2}+1}}$$

and

$$M(t) = \varphi + \frac{1}{n}(1 - 2t)M'(t). \tag{4.2}$$

Hence $M(t)$ in (4.2) satisfies (4.1) with $a = \varphi,\ b = n$, and $t < \frac{1}{2}$.

(2) Suppose that the mgf of *X* satisfies the linear differential equation in (4.1).

Now let us have a close look at the possible value of b and its consequences.

(a) If $b = 0$, then (4.1) turns out to be $M(t) = \infty, \forall t \in R$ and hence $M(t)$ has no meaning. Therefore $b \neq 0$.

(b) If $b > 0$, then the differential equation (4.1) can be expressed as

$$M(t) = a + \frac{(1-2t)}{b}\frac{dM(t)}{dt}$$

$$M(t) - a = \frac{(1-2t)}{b}\frac{dM(t)}{dt}$$

and we see that

$$\frac{dM(t)}{M(t)-a} = b\frac{dt}{(1-2t)}.$$

On integrating both sides with respect to t, we have

$$\int \frac{dM(t)}{M(t)-a} = \int b\frac{dt}{(1-2t)}$$

$$\Rightarrow \int \frac{dM(t)}{M(t)-a} = b\int \frac{dt}{(1-2t)}.$$

Hence $\log(M(t)-a) = -\frac{b}{2}\log(1-2t) + k$

$$\log(M(t)-a) = \log(1-2t)^{-\frac{b}{2}} + k$$

$$M(t) - a = e^{\log(1-2t)^{-\frac{b}{2}}+k}.$$

Now we get

$$M(t) - a = (1-2t)^{-\frac{b}{2}}e^{k},$$

where k is an arbitrary constant.

That is

$$M(t) = a + (1-2t)^{-\frac{b}{2}}e^{k},$$

where k is an arbitrary constant.

Therefore the solution of the differential equation (4.2) becomes

$$M(t) = a + (1 - 2t)^{-\frac{b}{2}} e^{k}. \tag{4.3}$$

Since $M(0) = 1,$ we get $e^{k} = (1 - a).$

From the equation (4.3),

$$M(t) = a + (1 - a)(1 - 2t)^{-\frac{b}{2}} = a + (1 - a)\frac{1}{(1 - 2t)^{\frac{b}{2}}}, \tag{4.4}$$

which is the mgf of a zero-inflated chi-square distribution with $a = \varphi,\ b = n,$ and $t < \frac{1}{2}.$ This completes the proof of the characterization Theorem.

Therefore X has the df specified in (2.1) with parameters $\varphi,\ \frac{n}{2},$ and $\frac{1}{2}.$

Remark

Also it can be noted that when $a = \phi = 0,$ the zero-inated chi-square distribution reduces to chi-square distribution and the Theorem in Section 4 concurs with the characterization result of chi-square distribution.

Conclusion

The result of this paper is purely of theoretical interest. The consequences and applications of this characterization are yet to be explored.

Acknowledgment

The author is grateful to Late Professor G. Nanjundan for motivating and directing towards the problem discussed in this paper.

References

[1] M. Ahsanullah, M. Shakil and B. M. Golam Kibria, Chracterizations of continuous distributions by truncated moments, Journal of Modern Applied Statistical Methods 15(1) (2016), 316-331.

[2] J. P. Boucher, M. Denuit and M. Guillen, Number of accidents or number of claims? An

approach with zero-inflated Poisson models for panel data, The Journal of Risk and Insurance 76(4) (2009), 821-846.

[3] E. D. Mills, Adjusting for Covariates in Zero-inflated Gamma and Zero-inflated Log-normal models for Semicontinuous Data, Unpublished Doctoral Thesis, The University of Iowa, Iowa City, Iowa. (2013).

[4] G. Nanjundan, A Statistical Analysis of the Number of Errors in a Software and Some Related Problems, Unpublished Doctoral Thesis, Bangalore University, Bangalore, India (2005).

[5] G. Nanjundan, A characterization of the members of a subfamily of power series distributions, Applied Mathematics 2 (2011), 750-751.

[6] S. Nagesh, G. Nanjundan, R. Suresh and Sadiq Pasha, A characterization of zero-inflated geometric model, International Journal of Mathematics Trends and Technology 23(1) (2015), 71-73.

[7] G. Nanjundan and Sadiq Pasha, A note on the characterization of zero-inflated Poisson distribution, Open Journal of Statistics 5 (2015a), 140-142.

[8] G. Nanjundan and Sadiq Pasha, A characterization of zero-inflated binomial model, International Journal of Mathematics and Computer Research 3(11) (2015b), 1187-1190.

[9] G. Nanjundan and Sadiq Pasha, Characterization of a family of zero-inflated discrete distributions, Prob. Stat. Forum 11 (2018), 77-80.

[10] P. G. Sankaran and N. Unnikrishnan, Characterizations by properties of residual life distributions, Statistics 24(3) (1993), 245-251.

[11] D. Srivastava and M. Sreehari, Characterization results based on variance bounds, Sankhya, Series A 52(3) (1990), 297-302.

[12] R. Suresh, S. Nagesh, G. Nanjundan and Sadiq Pasha, On a characterization of a zero-inated negative binomial distribution, Open Journal of Statistics 5 (2015), 511-513.

[13] K. C. H. Yip and K. K. W. Yau, On modeling claim frequency data in general insurance with extra zeros, Insurance: Mathematics and Economics, Elsevier 36 (2005), 153-163.