Wrapped Zero-inflated Poisson Distribution and Its Properties

Anita Brobbey[1], Aerambamoorthy Thavaneswaran[1] & Saumen Mandal[1]

[1] Department of Statistics, University of Manitoba, Winnipeg, Canada

Correspondence: Saumen Mandal, Department of Statistics, University of Manitoba, Winnipeg, MB, R3T 2N2, Canada. Tel: 1-204-474-9661. E-mail: saumen.mandal@umanitoba.ca

Abstract

Recently, there has been a growing interest in discrete valued wrapped distributions and the trigonometric moments. Characteristic functions of stable processes have been used to study the estimation of the model parameters using estimating function approach (Thavaneswaran et al., 2013). In this paper, we introduce a new discrete circular distribution, the wrapped zero-inflated Poisson distribution and derive its population characteristics.

Keywords: Circular distribution, zero-inflated Poisson, characteristic function, trigonometric moments, skewness, kurtosis

1. Introduction

Directional data is a natural occurence in many scientific fields such as earth science, meteorology, biology, medicine and astronomy where data is a set of observations on directions. Such directions may be in two or three dimensions. Specifically two-dimensional directions can be represented with respect to some suitable starting point and a "sense of rotation". Observations from such directions are called circular data (see more details in Chapter 2, Jammalamadaka and SenGupta, 2001).

Circular models have been studied by various researchers. A good overview on directional and circular models can be found in Mardia (1975). Mardia and Jupp (2000) published a comprehensive text on directional statistics including wrapped Poisson model. In discrete circular models, Girija et al. (2014a, 2014b) derived characteristics functions of the wrapped Poisson and wrapped binomial distributions. Pewsey et al. (2013, p.160) fits a circular regression model to the number of monthly deaths attributed to lung disease (cylindrical data collected over six consecutive years) in a given area. Recently, Thavaneswaran and Ravishanker (2015) studied estimating functions for circular models with continuous circular distributions, and discussed the related information issues. In this present paper, we introduce a new discrete circular distribution, the wrapped zero-inflated Poisson distribution and derive its population characteristics.

2. Circular Probability Distributions

A circular distribution is a probability distribution of a random variable, θ, whose values are angles measured in radians in the range $[0, 2\pi)$. The total probability of a circular distribution is concentrated on the circumference of a circle of unit radius. Circular distributions may be continuous or discrete-assigning probability masses only to a countable number of directions. The probability density function, $f(\theta)$, of a continuous circular random variable satisfies the following properties.

1. $f(\theta) \geq 0, \quad (0 \leq \theta < 2\pi)$

2. $\int_0^{2\pi} f(\theta) \, d\theta = 1$

3. The random variable θ has the same distribution as $(\theta + 2\pi k)$. Hence θ is a periodic random variable for any integer k. That is, $f(\theta) = f(\theta + 2\pi k)$.

3. Wrapped Discrete Distribution

In general, if X is a linear random variable on the real line with density $f(x)$, the corresponding wrapped random variable θ is defined by

$$\theta = x \pmod{2\pi}.$$

The operation corresponds to wrapping the real line around the unit circle accumulating probability over all the overlapping points $x = \theta, \theta \pm 2\pi, \theta \pm 4\pi, \cdots$. In particular, if θ has a distribution concentrated on the points $x = \frac{r}{2\pi m}$, $r = 0, 1, 2, \cdots, m-1$ and m is an integer, we have a wrapped discrete circular random variable θ, such that the probability

mass function of θ is

$$p_w(\theta = \frac{2\pi r}{m}) = \sum_{k=-\infty}^{\infty} p(r + km), \quad r = 0, 1, 2, \cdots, m - 1.$$

The probability mass function satisfies the following properties

1. $p_w(\theta = \frac{2\pi r}{m}) \geq 0$

2. $\sum_{r=0}^{m-1} p_w(\theta = \frac{2\pi r}{m}) = 1$

3. The random variable θ has the same distribution as $(\theta + 2\pi k)$. Hence θ is a periodic random variable for any integer k. That is, $p_w(\theta) = p_w(\theta + 2\pi k)$.

Similarly, if $f(x)$ is the probability mass function of X, the corresponding pmf of θ is

$$f_w(\theta) = \sum_{k=-\infty}^{\infty} f(\theta + 2\pi k).$$

In general, if $\phi_X(t)$ is the characteristic function of an unwrapped variable, X, then the characteristic function of the wrapped variable, X_w, at p is φ_p, which is given below (see page 53, Mardia, 1972). Moreover, let $F_w(\theta)$ and $F(x)$ are distribution functions of θ and X respectively. Then

$$\varphi_p = \int_0^{2\pi} e^{ip\theta} \, dF_w(\theta) = \sum_{k=-\infty}^{\infty} \int_{2\pi k}^{2\pi(k+1)} e^{ip\theta} \, dF(\theta)$$

$$= \int_{-\infty}^{\infty} e^{ipx} \, dF(x) = \phi_X(p).$$

4. Wrapped Zero-inflated Poisson Distribution

Data with excess zeros are often observed in applied science and public health studies. Zero-inflated Poisson (ZIP) regression model is a common statistical tool for analyzing such data. The *zero-inflated Poisson* distribution is given by

$$p(x) = \begin{cases} w + (1 - w)e^{-\lambda} & x = 0 \\ (1 - w)\frac{e^{-\lambda}\lambda^x}{x!} & x = 1, \cdots, \infty. \end{cases}$$

We define the corresponding wrapped probability function as

$$p_w(\theta = \frac{2\pi r}{m}) = \sum_{k=-\infty}^{\infty} p(r + km), \quad r = 0, 1, 2, \cdots, m - 1.$$

For a zero-inflated Poisson random variable, X, the characteristic function is given by

$$\phi_X(u) = E\left[e^{iuX}\right] = \sum_{k=0}^{\infty} p(k)e^{iuk}$$

$$= \left(w + (1 - w)e^{-\lambda}\right) + (1 - w)\sum_{k=1}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!} e^{iuk}$$

$$= w + (1 - w)e^{-\lambda}e^{\lambda e^{iu}}.$$

Applying the property of wrapped distributions, the characteristic function of the wrapped zero-inflated Poisson random variable, θ, for any interger p is

$$\varphi_p = E\left[e^{ip\theta}\right] = w + (1-w)e^{-\lambda}e^{\lambda e^{i\frac{2\pi p}{m}}} = w + (1-w)e^{-\lambda(1-\cos\frac{2\pi p}{m})}e^{i\lambda\sin\frac{2\pi p}{m}}$$

$$= w + (1-w)e^{-\lambda(1-\cos\frac{2\pi p}{m})}\left[\cos\left(\lambda\sin\frac{2\pi p}{m}\right) + i\sin\left(\lambda\sin\frac{2\pi p}{m}\right)\right].$$

The pth trigonometric moments of the wrapped zero-inflated Poisson distribution are

$$\alpha_p = w + (1-w)e^{-\lambda(1-\cos\frac{2\pi p}{m})}\cos\left(\lambda\sin\frac{2\pi p}{m}\right)$$

$$\beta_p = (1-w)e^{-\lambda(1-\cos\frac{2\pi p}{m})}\sin\left(\lambda\sin\frac{2\pi p}{m}\right).$$

The pth circular mean is

$$\mu_p = \tan^{-1}\left[\frac{\beta_p}{\alpha_p}\right]$$

$$= \tan^{-1}\left[\frac{(1-w)e^{-\lambda(1-\cos\frac{2\pi p}{m})}\sin\left(\lambda\sin\frac{2\pi p}{m}\right)}{w + (1-w)e^{-\lambda(1-\cos\frac{2\pi p}{m})}\cos\left(\lambda\sin\frac{2\pi p}{m}\right)}\right].$$

The pth circular mean resultant length is

$$\rho_p = \sqrt{\alpha_p^2 + \beta_p^2}$$

$$= \sqrt{w^2 + (1-w)^2 e^{-2\lambda(1-\cos\frac{2\pi p}{m})} + 2w(1-w)e^{-\lambda(1-\cos\frac{2\pi p}{m})}\cos\left(\lambda\sin\frac{2\pi p}{m}\right)}.$$

So, the mean of the wrapped zero-inflated Poisson distribution is

$$\mu = \mu_1 = \tan^{-1}\left[\frac{(1-w)e^{-\lambda(1-\cos\frac{2\pi}{m})}\sin\left(\lambda\sin\frac{2\pi}{m}\right)}{w + (1-w)e^{-\lambda(1-\cos\frac{2\pi}{m})}\cos\left(\lambda\sin\frac{2\pi}{m}\right)}\right]$$

and the mean resultant length is defined by

$$\rho = \rho_1 = \sqrt{\alpha_1^2 + \beta_1^2}$$

$$= \sqrt{w^2 + (1-w)^2 e^{-2\lambda(1-\cos\frac{2\pi}{m})} + 2w(1-w)e^{-\lambda(1-\cos\frac{2\pi}{m})}\cos\left(\lambda\sin\frac{2\pi}{m}\right)}$$

$$= \sqrt{a_1^2 + b_1^2 + 2a_1 b_1 \cos\left(\lambda\sin\frac{2\pi}{m}\right)}$$

where $a_1 = w$ and $b_1 = (1-w)e^{-\lambda(1-\cos\frac{2\pi}{m})}$.

The variance, V_0, and the standard deviation, σ_0, of the wrapped zero-inflated Poisson are given by

$$V_0 = 1 - \rho = 1 - \sqrt{a_1^2 + b_1^2 + 2a_1 b_1 \cos\left(\lambda\sin\frac{2\pi}{m}\right)}$$

$$\sigma_0 = \sqrt{-2\log(1 - V_0)}$$

$$= \sqrt{\log\left(\frac{1}{a_1^2 + b_1^2 + 2a_1 b_1 \cos\left(\lambda\sin\frac{2\pi}{m}\right)}\right)}.$$

5. Central Trigonometric Moments of Wrapped Zero-inflated Poisson Distribution

The characteristic function of the central wrapped zero-inflated Poisson distribution is

$$
\begin{aligned}
\varphi_p^c = E\left[e^{ip(\theta-\mu)}\right] &= e^{-ip\mu}E\left[e^{ip\theta}\right] \\
&= e^{-ip\mu}\left[w + (1-w)e^{-\lambda}e^{\lambda e^{i\frac{2\pi p}{m}}}\right] \\
&= e^{-ip\mu}\left[w + (1-w)e^{-\lambda}e^{\lambda\left(\cos\frac{2\pi p}{m}+i\sin\frac{2\pi p}{m}\right)}\right] \\
&= we^{ip\mu} + (1-w)e^{-\lambda}e^{\lambda\cos\frac{2\pi p}{m}}e^{i\left(\lambda\sin\frac{2\pi p}{m}-p\mu\right)} \\
&= (1-w)e^{-\lambda(1-\cos\frac{2\pi p}{m})}\left(\cos\left(\lambda\sin\frac{2\pi p}{m}-p\mu\right)+i\sin\left(\lambda\sin\frac{2\pi p}{m}-p\mu\right)\right)+ \\
&\quad w\left(\cos p\mu + i\sin p\mu\right).
\end{aligned}
$$

The pth central trigonometric moments are

$$
\alpha_p^c = w\cos p\mu + (1-w)e^{-\lambda(1-\cos\frac{2\pi p}{m})}\cos\left(\lambda\sin\frac{2\pi p}{m}-p\mu\right)
$$

$$
\beta_p^c = w\sin p\mu + (1-w)e^{-\lambda(1-\cos\frac{2\pi p}{m})}\sin\left(\lambda\sin\frac{2\pi p}{m}-p\mu\right).
$$

The circular skewness, s, and circular kurtosis, k, of the wrapped zero-inflated Poisson distribution are given by

$$
s = \frac{\beta_2^c}{V_0^{3/2}}
$$

$$
= \frac{\left[w\sin 2\mu + (1-w)e^{-\lambda(1-\cos\frac{4\pi}{m})}\sin\left(\lambda\sin\frac{4\pi}{m}-2\mu\right)\right]}{\left[1-\sqrt{a_1^2+b_1^2+2a_1b_1\cos\left(\lambda\sin\frac{2\pi}{m}\right)}\right]^{3/2}},
$$

$$
k = \frac{\alpha_2^c - (1-V_0)^4}{V_0^2}
$$

$$
= \frac{\left[w\cos 2\mu + (1-w)e^{-\lambda(1-\cos\frac{4\pi}{m})}\cos\left(\lambda\sin\frac{4\pi}{m}-2\mu\right)-\left(a_1^2+b_1^2+2a_1b_1\cos\left(\lambda\sin\frac{2\pi}{m}\right)\right)^2\right]}{\left[1-\sqrt{a_1^2+b_1^2+2a_1b_1\cos\left(\lambda\sin\frac{2\pi}{m}\right)}\right]^2}.
$$

6. Conclusions

A significant contribution of this paper is the introduction of a new wrapped zero-inflated Poisson distribution and its population characteristics which are useful in studying the parameter estimation using the estimating function method.

References

Girija, S. V. S., Rao, A. V. D., & Srihari, G. V. L. N. (2014). On characteristic function of wrapped Poisson distribution. *International Journal of Mathematical Archive, 5*(5), 168-173.

Girija, S. V. S., Rao, A. V. D., & Srihari, G. V. L. N. (2014). On wrapped binomial model characteristics. *Mathematics and Statistics, 2*(7), 231-234.

Jammalamadaka, S. R., & SenGupta, A. (2001). *Topics in circular statistics*. New Jersey: World Scientific.

Mardia, K. V. (1975). Statistics of directional data. *Journal of the Royal Statistical Society, Series B, 37*(3), 349-393.

Mardia, K. V. (1972). *Statistics of directional data*. London: Academic Press.

Mardia, K.V., & Jupp, P.E. (2000). *Directional statistics*. Chichester: John Wiley & Sons.

Pewsey, A., Neuhauser, M., & Ruxton, G. (2013). *Circular statistics in R*. Oxford: Oxford University Press.

Thavaneswaran, A., & Ravishanker, N. (2015). Estimating functions for circular models. *Technical Report #2015-07, Department of Statistics, University of Connecticut, Storrs, USA*.

Thavaneswaran, A., Ravishanker, N., & Liang, Y. (2013). Inference for linear and nonlinear stable error processes via estimating functions. *Journal of Statistical Planning and Inference, 143*(4), 827-841. http://dx.doi.org/10.1016/j.jspi.2012.10.014