Wrapped Geometric Distribution:
A new Probability Model for Circular Data

Sophy Jacob [1] and K. Jayakumar [2]

[1] *MES Asmabi College, P. Vemballur, Kerala, India*

sophyjacob@rediffmail.com

[2] *University of Calicut, Kerala, India*

jkumar19@rediffmail.com

We propose a new family of circular distributions, obtained by wrapping geometric distribution on $Z_+ = 0, 1, \ldots,$ around a unit circle. The properties of this new family of distributions are studied.

Keywords: Circular distributions, trigonometric moments, wrapped geometric distribution.

1. Introduction

Circular data arise in various ways. Two of the most common correspond to two circular measuring instruments, the compass and the clock. Data measured by compass usually include wind directions, ocean current directions, the direction and orientations of birds and animals, and orientation of geological phenomena like rock cores and fractures. Data measured by clock includes times of arrival of patients at a hospital emergency room, incidences of a disease throughout the year, and the number of tourists (daily or monthly) in a city within a year, where the calendar is regarded as a one-year clock. Circular or directional data also arise in many scientific fields, such as Biology, Geology, Meteorology, Physics, Psychology, Medicine and Astronomy (see Mardia and Jupp (2000)).

Study on directional data can be dated back to the 18th century. In 1734 Daniel Bernoulli proposed to use the resultant length of normal vectors to test for uniformity of unit vectors on the sphere (see Mardia (1972)). In 1918 von Mises introduced a distribution on the circle by using characterization analogous to the Gauss characterization of the normal distribution on a line (see Mardia (1972)). Later, interest was renewed in spherical and circular data by Fisher (1953), Greenwood and Durand (1955), and Watson and William (1956).

Wrapped distributions provide a rich and useful class of models for circular data. The special cases of the wrapped normal, wrapped Cauchy, wrapped Poisson are widely discussed in Mardia (1972). The wrapped Exponential and Laplace distributions have been introduced and studied by Jammalamadaka and Kozubowski (2004). In Section 2, we introduce and study wrapped geometric

distribution. Properties of the wrapped geometric distribution are studied in Section 3. Section 4 deals with estimation of parameters.

2. Wrapped Geometric Distribution

One of the common methods to analyze circular data is known as wrapping approach. In this method, given a known distribution on the real line, we wrap it around the circumference of a circle with unit radius. Technically this means that if X is a random variable on the real line with distribution function $F(x)$, the random variable X_w of the the wrapped distribution is given by

$$X_w = X \pmod{2\pi} \tag{2.1}$$

and the distribution function $F_w(\theta)$ is given by

$$F_w(\theta) = \sum_{k=-\infty}^{\infty} \{F(\theta + 2\pi k) - F(2\pi k)\}. \tag{2.2}$$

More precisely, if X is a random variable on the integers, the X_w, defined by

$$X_w = 2\pi x \pmod{2\pi m}$$

is a random variable on the lattice $\left\{ \frac{2\pi r}{m}, \ r = 0, 1, \ldots, m-1 \right\}$ on the circle.

Now, consider the geometric distribution on Z_+ with parameter δ. Its probability mass function is given by

$$p(x; \delta) = (1-\delta)^x \delta, \ \ x = 0, 1, \ldots; \ \ \delta > 0.$$

The probability function of X_w is given by

$$P\left(X_w = \frac{2\pi r}{m}\right) = \sum_{k=-\infty}^{\infty} p(r + km; \delta), \quad r = 0, 1, \ldots, m-1.$$

That is,

$$P\left(X_w = \frac{2\pi r}{m}\right) = \sum_{k=0}^{\infty} (1-\delta)^{r+km} \delta = \frac{\delta(1-\delta)^r}{1-(1-\delta)^m} \tag{2.3}$$

where $m \in \mathbb{N}$ and $r = 0, 1, \ldots, m-1$. Again

$$\sum_{r=0}^{m-1} \frac{\delta(1-\delta)^r}{1-(1-\delta)^m} = 1.$$

Therefore, $P(\cdot)$ represents a probability mass function.

If $\Phi(t)$ is the characteristic function of a linear random variable X, then the characteristic function of X_w is $\Phi(p)$. For the wrapped geometric distribution, we have

$$\Phi(p) = \sum_{r=0}^{m-1} \frac{\delta(1-\delta)^r e^{i2\pi \frac{rp}{m}}}{1-(1-\delta)^m}, \quad p = 0, \pm 1, \pm 2, \ldots$$

$$= \frac{\delta}{1-(1-\delta)e^{\frac{i2\pi p}{m}}}$$

$$= \frac{\delta}{x-iy},$$

where $x = 1 - (1-\delta)\cos\left(\frac{2\pi p}{m}\right)$ and $y = (1-\delta)\sin\left(\frac{2\pi p}{m}\right)$, where $p \neq 0 \pmod{m}$.

On expanding $\Phi(p)$, we have

$$\Phi(p) = \delta\left\{\left(1-(1-\delta)\cos\left(\frac{2\pi p}{m}\right)\right)^2 + \left((1-\delta)\sin\left(\frac{2\pi p}{m}\right)\right)^2\right\}^{\frac{-1}{2}} e^{i\arctan\left\{\frac{(1-\delta)\sin\left(\frac{2\pi p}{m}\right)}{1-(1-\delta)\cos\left(\frac{2\pi p}{m}\right)}\right\}}$$

$$= \rho_p e^{i\mu_p}, \tag{2.4}$$

where

$$\rho_p = \delta\left\{\left(1-(1-\delta)\cos\left(\frac{2\pi p}{m}\right)\right)^2 + \left((1-\delta)\sin\left(\frac{2\pi p}{m}\right)\right)^2\right\}^{\frac{-1}{2}}$$

and

$$\mu_p = \arctan\left\{\frac{(1-\delta)\sin\left(\frac{2\pi p}{m}\right)}{1-(1-\delta)\cos\left(\frac{2\pi p}{m}\right)}\right\}.$$

Thus the pth trigonometric moment of the wrapped geometric distribution is given by

$$\Phi_p = \alpha_p + i\beta_p,$$

where

$$\alpha_p = E(\cos p\theta) = \rho_p \cos\mu_p$$

and

$$\beta_p = E(\sin p\theta) = \rho_p \sin\mu_p.$$

The central trigonometric moments are given by

$$\overline{\alpha}_p = \rho_p \cos(\mu_p - p\mu_1), \quad \overline{\beta}_p = \rho_p \sin(\mu_p - p\mu_1).$$

Sophy Jacob and K. Jayakumar

The length of the resultant vector, $\rho = \rho_1$ is

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

$$= \frac{\delta}{\left(1 + (1-\delta)^2 - 2(1-\delta)\cos\frac{2\pi}{m}\right)^{\frac{1}{2}}}$$

$$= \frac{\delta}{\left((\delta)^2 + 2(1-\delta)\left(1 - \cos\frac{2\pi}{m}\right)\right)^{\frac{1}{2}}} \qquad (2.5)$$

and the mean direction,

$$\mu = \arctan\frac{(1-\delta)\sin\left(\frac{2\pi}{m}\right)}{1 - (1-\delta)\cos\left(\frac{2\pi}{m}\right)}.$$

The circular variance, V_0 is given by

$$V_0 = 1 - \rho$$

$$= 1 - \frac{\delta}{\left(1 + (1-\delta)^2 - 2(1-\delta)\cos\frac{2\pi}{m}\right)^{\frac{1}{2}}}.$$

The circular standard deviation,

$$\sigma_0 = \sqrt{-2\ln\rho}$$

$$= \sqrt{\ln\frac{(\delta)^2 + 2(1-\delta)\left(1 - \cos\frac{2\pi}{m}\right)}{(\delta)^2}}.$$

The measure of skewness is determined by the formula

$$\gamma_1^0 = \frac{\overline{\beta}_2}{V_0^{\frac{3}{2}}}$$

$$= \frac{\rho_2\sin(\mu_2 - 2\mu_1)}{\left(1 - \frac{\delta}{\left((\delta)^2 + 2(1-\delta)\left(1-\cos\frac{2\pi}{m}\right)\right)^{\frac{1}{2}}}\right)^{\frac{3}{2}}}$$

The measure of kurtosis,

$$\gamma_2^0 = \frac{\rho_2\cos(\mu_2 - 2\mu_1) - \rho^4}{V_0^2}.$$

Proposition 2.1. *Let $\Theta_1,\ldots,\Theta_n$ be n independent and identically distributed wrapped geometric random variables with parameters δ and m. Then $\Theta = \Theta_1 + \cdots + \Theta_n \pmod{2\pi}$ follows wrapped negative binomial distribution with parameters p, m, n where $p = \frac{1-\delta}{\delta}$, $m \in \mathbb{N}$.*

Proof. $\Theta_i \sim$ Wrapped geometric distribution with parameters δ and m. Therefore,

$$\Phi_{\Theta_i}(p) = \frac{\delta}{1 - (1 - \delta) e^{\frac{i2\pi p}{m}}}.$$

Let $\Phi_{\Theta}(p)$ denotes the characteristic function of Θ. Then

$$\Phi_{\Theta}(p) = \Phi_{\Theta_1 + \Theta_2 + \cdots + \Theta_n \pmod{2\pi}}(p)$$
$$= \prod_{j=1}^{n} \Phi_{\Theta_j}(p)$$
$$= \left(Q - P e^{\frac{i2\pi p}{m}} \right)^{-n}$$

where $Q = \frac{1}{\delta}$, $P = \frac{1-\delta}{\delta}$, $Q - P = 1$, $\delta > 0$, which is the characteristic function of wrapped negative binomial distribution. $\square$

3. Divisibility Properties

3.1. *Infinite Divisibility*

Definition 3.1. An angular random variable Θ is said to be *infinitely divisible* if for any integer $n \geqslant 1$ there exists independently and identically distributed (i.i.d.) angular random variables $\Theta_1, \Theta_2, \ldots, \Theta_n$ such that

$$\Theta_1 + \Theta_2 + \cdots + \Theta_n \pmod{2\pi} \overset{d}{=} \Theta.$$

Since a circular random variable obtained by wrapping an infinitely divisible random variable is infinitely divisible (see Mardia and Jupp, 2000), the wrapped geometric random variable is infinitely divisible.
We have, the characteristic function of wrapped geometric random variable Θ,

$$\Phi_{\Theta}(p) = \frac{\delta}{1 - (1 - \delta) e^{\frac{i2\pi p}{m}}}$$
$$= \left(\left(Q - P e^{\frac{i2\pi p}{m}} \right)^{-\frac{1}{n}} \right)^{n},$$

for every $n \geqslant 1$, where $Q = \frac{1}{\delta}$, $P = \frac{1-\delta}{\delta}$, $Q - P = 1$, $\delta > 0$.

3.2. *Geometric Infinite Divisibility*

Definition 3.2. An angular random variable Θ is said to be geometrically infinitely divisible if for any $q \in (0, 1)$ there exists i.i.d. angular random variables $\Theta_1, \Theta_2, \ldots$ such that

$$\Theta_1 + \Theta_2 + \cdots + \Theta_{\nu_q} \mod (2\pi) \overset{d}{=} \Theta,$$

where ν_q has the geometric distribution

$$P(\nu_q = k) = (1 - q)^{k-1} q, \quad k = 1, 2, \ldots$$

Let $\Theta_1, \Theta_2, \ldots$ be i.i.d. wrapped geometric random variables. Conditioning on the distribution of v_q, we have

$$E\left[e^{ip(\Theta_1+\Theta_2+\cdots+\Theta_{v_q})\,(\text{mod }2\pi)}\right] = \sum_{k=1}^{\infty} E\left[e^{ip(\Theta_1+\Theta_2+\cdots+\Theta_k)}\right] (1-q)^{k-1}q$$

$$= \sum_{k=1}^{\infty} \widetilde{\Phi}_p^k (1-q)^{k-1}q$$

$$= \frac{q\widetilde{\Phi}_p}{1-(1-q)\widetilde{\Phi}_p}$$

$$= \frac{q}{\left(Q - Pe^{\frac{i2\pi p}{m}}\right) - (1-q)}$$

$$= \left(A - Be^{\frac{i2\pi p}{m}}\right)^{-1}$$

where $A = \frac{Q-1+q}{q}$, $B = \frac{P}{q}$ and $A - B = 1$. Hence the wrapped geometric distribution is geometrically infinitely divisible.

3.3. *A Generalization of Wrapped Geometric Distribution and its Divisibility Properties*

The probability mass function of the wrapped geometric random variable Θ with parameters δ and m is given by

$$P\left(\Theta = \frac{2\pi r}{m}\right) = \frac{\delta(1-\delta)^r}{1-(1-\delta)^m}; \;\; 0 < \delta < 1, \;\; m \in \mathbb{N}, \;\; r = 0, 1, \ldots.$$

The probability generating function (p.g.f.) of Θ is given by

$$P(s) = \sum_{r=0}^{m-1} \frac{\delta(1-\delta)^r s^r}{1-(1-\delta)^m}$$

$$= \frac{\delta}{1-(1-\delta)^m} \left(\sum_{r=0}^{m-1} (s(1-\delta))^r\right)$$

$$= \frac{\delta\left(1-(1-\delta)^m(s)^m\right)}{\left(1-(1-\delta)^m\right)\left(1-(1-\delta)s\right)}.$$

We have $P(1) = 1$. The series $P(s)$ converges for $-1 \leqslant s \leqslant 1$. When $s = e^{\frac{i2\pi p}{m}}$, we get the characteristic function of wrapped geometric distribution.

As a generalization of geometric distribution, we now consider the discrete Mittag-Leffler distribution introduced by Pillai and Jayakumar (1995). The p.g.f. of the discrete Mittag-Leffler distribution with parameter α is given by

$$P(s) = \frac{\delta}{1-(1-\delta)(1-(1-s)^\alpha)}.$$

Therefore, its characteristic function,

$$\Phi_t = \frac{\delta}{1-(1-\delta)(1-(1-e^{it})^\alpha)}.$$

Hence the wrapped discrete Mittag-Leffler distribution has characteristic function,

$$
\begin{aligned}
\Phi_p &= \frac{\delta}{1 - (1-\delta)(1 - (1 - e^{\frac{i2\pi p}{m}})^\alpha)} \\
&= \frac{\delta}{\delta + (1-\delta)(1 - e^{\frac{i2\pi p}{m}})^\alpha} \\
&= \frac{1}{1 + c \left(1 - e^{\frac{i2p\pi}{m}}\right)^\alpha},
\end{aligned}
$$

where $c = \frac{1-\delta}{\delta}$, $0 < \alpha < 1$, $p = 0, \pm 1, \pm 2, \ldots$, $m \in \mathbb{N}$, $p \neq 0$, (mod m).

When $\alpha = 1$, we get the characteristic function of wrapped geometric distribution.

We have the characteristic function of wrapped discrete Mittag-Leffler distribution,

$$
\begin{aligned}
\Phi_p &= \frac{1}{1 + c(1 - e^{\frac{i2\pi p}{m}})^\alpha} \\
&= \frac{1}{\left(\left(1 + c \left(1 - e^{\frac{i2\pi p}{m}} \right)^\alpha \right)^{\frac{1}{n}} \right)^n}, \\
&= \widetilde{\Phi}_p^n,
\end{aligned}
$$

for every $n \geqslant 1$ and $\widetilde{\Phi}_p$, is the characteristic function of wrapped discrete Linnik distribution. Again, we have

$$
\begin{aligned}
E\left[e^{ip(\Theta_1 + \Theta_2 + \cdots + \Theta_{v_q}) \,(\text{mod } 2\pi)} \right] &= \sum_{k=1}^\infty E\left[e^{ip(\Theta_1 + \Theta_2 + \cdots + \Theta_k)} \right] (1 - \delta)^{k-1} \delta \\
&= \sum_{k=1}^\infty (\widetilde{\Phi}_p)^k (1 - \delta)^{k-1} \delta \\
&= \frac{\delta \widetilde{\Phi}_p}{1 - (1 - \delta)\widetilde{\Phi}_p} \\
&= \frac{\delta}{\delta + c \left(1 - e^{\frac{i2\pi p}{m}} \right)^\alpha} \\
&= \frac{1}{1 + \sigma \left(1 - e^{\frac{i2\pi p}{m}} \right)^\alpha}
\end{aligned}
$$

where $c = \frac{1-\delta}{\delta}$, $\sigma = \frac{c}{\delta} > 0$. So wrapped discrete Mittag-Leffler distribution is geometrically infinitely divisible.

4. Estimation

4.1. *Method of Moments*

Let $\theta = (\theta_1, \ldots, \theta_n)$ be a random sample of size n taken from the wrapped geometric distribution with parameters δ and m. In this method, estimates of the parameters can be obtained by equating

sample moments to the corresponding population moments. We have, the pth sample trigonometric moment about the zero direction, $m'_p = a_p + i b_p$, where

$$a_p = \frac{1}{n} \sum_{j=1}^{n} \cos(p\,\theta_j) \ \text{ and } \ b_p = \frac{1}{n} \sum_{j=1}^{n} \sin(p\,\theta_j).$$

Equating a_p to α_p and b_p to β_p, we get

$$a_p = \frac{\delta}{\left((\delta)^2 + 2(1-\delta)\left(1-\cos\frac{2\pi}{m}\right)\right)^{\frac{1}{2}}} \times \cos\left\{ \tan^{-1} \frac{(1-\delta)\sin\left(\frac{2\pi p}{m}\right)}{1-(1-\delta)\cos\left(\frac{2\pi p}{m}\right)} \right\} \qquad (4.1)$$

and

$$b_p = \frac{\delta}{\left((\delta)^2 + 2(1-\delta)\left(1-\cos\frac{2\pi}{m}\right)\right)^{\frac{1}{2}}} \times \sin\left\{ \tan^{-1} \frac{(1-\delta)\sin\left(\frac{2\pi p}{m}\right)}{1-(1-\delta)\cos\left(\frac{2\pi p}{m}\right)} \right\} \qquad (4.2)$$

where $p \neq 0 \pmod m$. Solving (4.1) and (4.2) for a fixed value of m and $p = 1$, we get

$$\widehat{\delta} = \frac{\overline{R}^2 \cos\left(\frac{2\pi}{m}\right) \pm \overline{R}\sqrt{\overline{R}^2 \cos\left(\frac{2\pi}{m}\right) - 2(\overline{R}^2 - 1)\cos\left(\frac{2\pi}{m}\right)}}{\overline{R}^2 - 1}$$

where $\overline{R} = \sqrt{a_1^2 + b_1^2}$, the sample mean length of the resultant.

References

[1] R.A. Fisher, "Dispersion on a Sphere", *Proceedings of the Royal Society of London*, A, **217**, 295–305 (1953).
[2] J.A. Greenwood and D. Durand, "The distribution of length and components of the sum of n random unit vectors", *Annals of Mathematical Statistics*, **26**, 233–246 (1955).
[3] Jammalamadaka, S. Rao and T.J. Kozubowski, "New families of wrapped distributions for modeling skew circular data", *Communications in Statistics: Theory and Methods*, **33**, 2059–2074 (2004).
[4] Jammalamadaka, S. Rao and A Sengupta, "Topics in Circular Statistics", *World Scientific, Singapore*,(2001).
[5] K.V. Mardia, "Statistics of Directional Data", *Academic Press, London* (1972).
[6] K.V. Mardia and P.E. Jupp, "Directional Statistics, 2nd Edition", *Wiley , New York* (2000).
[7] R.N. Pillai and K. Jayakumar, "Discrete Mittag-Leffler distributions", *Statistics and Probability Letters*, **23**, 271–274 (1995).
[8] G.S. Watson and E.J. William, "On the construction of significance tests on the circle and the sphere", *Biometrika*, **43**, 344–352 (1956).