



Goodness of Fit Test For Wrapped Normal Distribution

Anahita Nodehi

Department of Statistics, Computer Science, Applications (DiSIA), Florence University, Florence, Italy.

Abstract. *One of the main difficulties in any statistical method is whether the data could have actually been drawn from that fitted distribution or not. To extend in circular data, it is necessary to consider nature feature of this data. The Wrapped Normal and Von Mises are two most important and famous distributions in circular. Based on the author's knowledge, there is no Goodness-of-Fit test for Wrapped Normal. To enhance this issue, in this paper, we present an appropriate test and compare its performance based on simulation study.*

Keywords. *Goodness-of-Fit Test; Wrapped Normal; Von Mises; Circular Data*

1 Introduction

Directional data is a type of data that has a wide range of applications in applied sciences. For consideration of value to directions, it is common to specify an angle on a unit circle since an initial direction and the orientation of the circle have been chosen. Therefore, having such periodic feature makes one to consider the topological feature of the non-Euclidean space. Accordingly, many methods and statistical techniques have been developed to analyze and understand this type of data. The popular approaches have been embedding, wrapping and intrinsic approaches. Based on every approach, great number of distributions are proposed which are in non-Euclidean space. The Wrapped Normal and the Von Mises are two important distributions on the circle, which resemble on circle the Normal distribution on Euclidean space. To provide a better sense of this phenomenon, certain overview of the embedding and intrinsic approaches can be found in Jammalamadaka and SenGupta (2001), Mardia (1972) and Mardia and Jupp (2000).

One of the fundamental questions that arise in every statistical application is whether the data could have actually been drawn from that fitted distribution. This is the so-called Goodness-of-Fit problem. The importance of this problem arises especially in some estimation methods as the first step is to check if the assumption is hold or not. For example, Nodehi et al (2018) proposed two algorithms which estimate the parameters of Wrapped Normal distribution. With regard to that, it is necessary to check whether the data is Wrapped Normal or not.

To do so, in circular data, one should consider the periodic feature of data. The Goodness-of-Fit testing for a Von Mises distribution fitted using maximum likelihood estimation (without bias correction for

the estimation of κ), were obtained by Lockhart and Stephens (1985). Based on the author's knowledge, there are no contribution to find the same test for Wrapped Normal distribution. In that sense, the main goal in this paper is to propose a Goodness-of-Fit test based on Wrapped Normal distribution.

The remainder of this paper is organized as follows. In section 2, the review of circular densities is presented. Afterwards, Section 3 is based on Goodness-of-Fit procedure. Section 4 provides simulation study while Section 5 gives final comments and remarks.

2 Statistical Modeling

As mentioned in Introduction, there are three approaches to modeling circular data: embedding, wrapping and intrinsic approaches. In the embedding approach the sample space is considered as part of larger space and the distributions on the S^{p-1} (the circular sample space) can be obtained by radial projection of the in line distributions on R^p . In general, most of the literature is focused on developing statistical methods for the projected Normal distribution, which is, the only a significant limitation of the embedding approach.

In the intrinsic approach, the circle is used as the sample space. The directions are represented as points on the circle and probability distributions are defined on the circle directly. The main probability distributions obtained from this approach are the Uniform, Cardioid and Von Mises distributions.

The wrapping approach consists to wrap a known distribution in the real line around a circumference of a circle with a unit radius. In that sense, the main characteristic of this approach is flexibility. Elaborating on, it is a rich class of distributions on the circle that can be obtained using the wrapping technique, as it is possible to wrap any known distribution in the real line onto the circle. Therefore, the most famous probability distribution based on this approach is Wrapped Normal which resembles Normal distribution in Euclidean space. Since the main contribution of this paper is to propose a Goodness-of-Fit for Wrapped Normal, it is necessary to revivie certain features of this distribution.

Any linear random variable X may be transformed to a circular random variable by reducing its modulo 2π . i.e.

$$\theta = X(\text{mod } 2\pi)$$

This operation is equal to taking a line random variable and wrapping around circle of unit radius, accumulating probability over all points $X = (\theta + 2K\pi)$ where $K \in \mathbb{Z}$. If F represents the circular distribution function and G distribution function of line random variable, we have

$$F(\theta) = \sum_{K=-\infty}^{+\infty} \{G(\theta + 2K\pi) - G(2K\pi)\}, \quad 0 \leq \theta \leq 2\pi.$$

In particular, if θ has a circular density function f and g is density function of X then

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

A Wrapped Normal distribution is obtained by wrapping a $N(\mu, \sigma^2)$ distribution around the circle. Its

pdf is given by

$$f(\theta) = \sum_{K=-\infty}^{+\infty} g(\theta - \mu + 2K\pi) = \frac{1}{\sqrt{2\pi\sigma^2}} \sum_{K=-\infty}^{+\infty} \exp\left\{-\frac{(\theta - \mu + 2K\pi)^2}{2\sigma^2}\right\}$$

An alternate and more useful representation of this density using Fourier expansion and properties of characteristic function can be shown to be

$$f(\theta) = \frac{1}{2\pi} \left(1 + 2 \sum_{p=1}^{+\infty} \rho^{p^2} \cos p(\theta - \mu)\right)$$

where $\rho = \exp\left(-\frac{\sigma^2}{2}\right)$ (Jammalamadaka and SenGupta, 2001). In this regard, some properties of Wrapped Normal are as follows: it is unimodal and symmetric about the value μ , the mean resultant length is $\rho = \exp\left(-\frac{\sigma^2}{2}\right)$, as $\rho \rightarrow 0$, the distribution converges to the uniform distribution, as $\rho \rightarrow 1$, it tends to a point distribution at μ , it appears in the central limit theorem and Brownian Motion, the convolution of two Wrapped Normal variables is also Wrapped Normal, unlike the Von Mises distribution

3 Goodness-of-fit test

According to Pewsay et al (2013), considering the circular analogue of the probability integral transformation, it follows implicitly that the Goodness-of-Fit of a posited distribution with distribution function $F(\theta)$ that can be tested by calculating the values of $2\pi F(\theta_1), \dots, 2\pi F(\theta_n)$ and applying any test of circular uniformity to them. If the data do come from the postulated distribution, then we would expect circular uniformity not to be rejected. The problem with this approach is that the usual critical values of the tests for circular uniformity do not apply if the parameters of the distribution have been estimated from the data. The difference between the correct critical values and those for the usual tests of circular uniformity should not be great, however, for large sample sizes. Lockhart and Stephens (1985) proposed a Goodness-of-Fit testing for a Von Mises distribution fitted using maximum likelihood estimation based on the critical values of Watson's U^2 test which is implemented within the function `watson.test` available in R's circular package (Agostinelli and Lund, 2017) if its argument `dist` is specified as `vonmises`. Since, Wrapped Normal and Von Mises have close relationships, it is possible to use the same procedure by calculating the values of $2\pi F(\theta_1), \dots, 2\pi F(\theta_n)$ and applying any test of circular uniformity to them. In this regard, it is expected that under some conditions ($\sigma \rightarrow 0$), the two distributions have the same behavior.

4 Simulation study

To compare the performance of the proposed method we consider simulation study. To do so, we consider sample size $n = 50, 100$, $\mu_0 = 0$, $\sigma_0 = (\pi/8, \pi/4, \pi/2, \pi, 3/2\pi, 2\pi)$, and the number of Monte Carlo replications 100. As can be seen in Table 1, the values (within 100 replications) are based on number of times, the test has been accepted. In other words, we generate the data of Wrapped Normal and Von Mises distribution and see whether the data could have actually been drawn from that fitted distribution or not. According to Kent (1978), any Von Mises distribution can be approximated by a Wrapped Normal distribution when σ is small or $\kappa \rightarrow \infty$; i.e.

$$f_{VM}(\theta, \mu, \kappa) - f_{WN}(\theta, \mu, A_1(\kappa)) = O(\kappa^{-1/2})$$

Real Distribution	σ	Fitted distribution							
		WN				VM			
		Kuper		Watson		Kuper		Watson	
		$n = 50$	$n = 100$	$n = 50$	$n = 100$	$n = 50$	$n = 100$	$n = 50$	$n = 100$
WN	$\frac{\pi}{8}$	100	92	99	93	100	92	99	93
	$\frac{\pi}{4}$	97	96	96	96	98	94	96	92
	$\frac{\pi}{2}$	95	88	98	92	97	95	95	92
	π	85	77	78	65	90	96	92	96
	$\frac{3\pi}{2}$	86	74	85	76	96	92	96	92
	2π	81	78	80	69	91	100	94	98
	VM	$\frac{\pi}{8}$	85	93	83	91	92	99	91
$\frac{\pi}{4}$		89	86	91	81	94	97	96	96
$\frac{\pi}{2}$		95	92	96	94	96	92	98	96
π		94	86	93	79	95	97	94	96
$\frac{3\pi}{2}$		93	83	92	76	97	96	96	96
2π		94	78	93	73	96	96	97	95

Table 1: Results of the Monte Carlo simulation based on 100 replications.

where f_{VM} and f_{WN} are the densities of the Von Mises (μ, κ) and the Wrapped Normal $(\mu, A_1(\kappa))$ distribution, respectively. Therefore, when σ is sufficiently small, the test cannot distinguish between these two distribution, as expected. Moreover, by $\sigma \rightarrow \infty$, the both tend to circular Uniform distribution. Thus, it is better to increase the number of replications and sample sizes and add another distribution with different features in simulation study to see more distinctions obtained by the test.

5 Conclusion

Based on the author's knowledge, there is yet no test for Goodness-of-Fit to Wrapped Normal distribution. According to periodic feature of circular data and close relationship between Wrapped Normal and Von Mises distribution, in simulation study, we show the performance of this test.

References

- [1] Agostinelli, C. and Lund, U. (2017). R package `circular`: Circular Statistics (version 0.4-93), <https://r-forge.r-project.org/projects/circular/>.
- [2] Jammalamadaka, S. R. and SenGupta, A. (2001). *Topics in Circular Statistics*, World Scientific, Singapore.
- [3] Kent, J. T. (1978). Limiting Behaviour of the Von Mises-Fisher Distribution. *Math. Proc. Cambridge Phil. Soc.*, **84**, 531-536.
- [4] Lockhart, R. A. and Stephens, M. A. (1985). Tests of Fit for the Von Mises Distribution. *Biometrika*, **72**, 647-52.
- [5] Mardia, K. V. (1972). *Statistics of Directional Data*, Academic Press, London.
- [6] Mardia, K. V. and Jupp, P. (2000). *Directional Statistics*, John Wiley, Chichester.
- [7] Nodehi, A., Gosalizadeh, M., Maadooliat, M. and Agostinelli, C. (2018). Estimation of Multivariate Wrapped Models for Data in Torus, *arXiv preprint arXiv:1811.06007*.
- [8] Pewsey, A., Neuhauser, M., and Ruxton, G. D. (2013). *Circular Statistics in R*. Oxford University Press, England.