A GROUP SEQUENTIAL TEST OF CIRCULAR DATA USING THE VON MISES DISTRIBUTION

K. Özgür Peker *† Sevil Bacanlı ‡

Abstract

In this study, the group sequential test is suggested for the mean direction parameter of the von Mises distribution when the concentration parameter is known and unknown. An application of the proposed test is illustrated by using a medical data of the patients, who were complained about internal rotation angles of the shoulder and treated in a rehabilitation and physical therapy center in Eskiehir, Turkey. It is shown that the results of the study demonstrate that the group sequential test can provide a great advantage not only for linear data but also for circular data in terms of sample size.

Keywords: Sequential test, Circular data, Von Mises distribution, Mean direction
2010 AMS Classification: Primary 62H11, Secondary 62L10

1 Introduction

Circular data often arise in many scientific disciplines like meteorology, geography, biology, geology and medicine etc. As an example, meteorological events are periodical, that's why it is convenient to analyze them by using directional methods. It is shown that the distribution of the wind direction can be approximated by a specific circular model.

Ecologists consider the prevailing wind direction as an important factor in many studies including those of which involve pollutant transport. In Geology, geologists study paleocurrents to find out about the direction of flow of rivers in the past [16] and analyze paleomagnetic directions of the earth's magnetic pole to investigate the phenomenon of pole-reversal as well as in support of the hypothesis of continental drift. In Biology, biologists who study bird-migrations record the flight directions of just-released birds as they disappear over the horizon. Batschelet [2] presented a number of noteworthy applications of circular statistics in Biology. Also, any periodic phenomenon which is known and may be a day, a month or a year, can be represented on a circle by aggregating the necessary data of several individuals or periods if the circumferences corresponds to this period. Examples include arrival times of patients to a hospital over the day, or the time of patients at a hospital in the day. As a last example, the circle may represent the 365 days in the year and could be plotted the occurrence of crash accidents in a specific roadway junction to see if they are uniformly distributed over the different seasons of the year [8].

*Corresponding Author.
†Anadolu University, Faculty of Science, Department of Statistics, 26470, Eskisehir, Turkey, Email: (K.Ö.Peker) opeker@anadolu.edu.tr
‡Hacettepe University, Faculty of Science, Department of Statistics, 06800, Ankara, Turkey. Email:(S.Bacanlı) sevil@hacettepe.edu.tr

Circular data take values on the circumference of a circle and they form the angles in the range $(0^o, 360^o)$ or $(0, 2\pi)$ radians [7]. The circular probability distributions are used to fit the distribution of circular data. The von Mises distribution is the most common probability distribution for circular data. A comprehensive discussion of circular statistics as well as examples of the applications and general properties of the von Mises distribution can be found in [11] and [8].

There are many practical situations in which it is desirable to update the decision with every incoming observation, by sequentially, either in the temporal or in the spatial mode of collecting the circular data.

As an example of using a sequential test for circular data, observations on the imbalanced directions of individually produced wheels can provide for the information of whether the procedure is under control.

Gadsden & Kanji [5] developed a sequential probability ratio test (SPRT) of Wald [17] for the mean direction (μ_{VM}) of the von Mises distribution with a known and an unknown concentration parameter (κ). Gadsden & Kanji [6] represents the applications of SPRT for circular data.

The sample size is a predetermined fixed value in fixed sample size test procedure. In practice, this test cause, the practitioner, to spend more resources such as money and time. When the sequential tests are used, these difficulties can be removed. The test begins with a single observation value and stops when there is sufficient data for statistical comparison and for making a decision on the hypothesis. Thus, it leads to a great saving in the sample size [17].

However, in some cases, when a new data is obtained, testing the data by grouping is an easier way than applying SPRT. A test which is performed sequentially by grouping data is called a group sequential test (GST). Various group sequential testing procedures have been proposed to achieve the desired levels of type I error. Pocock [14], O'Brien & Fleming [12] and Lan & DeMets [10] were among the first scholars to develop group sequential test. A great part of the progress of group sequential tests are reviewed in detail by Jennison & Turnbull [9].

Group sequential tests are widely used in medicine. On the other hand, medical events are convenient to be analyzed using directional methods since many of them are periodical. The occurrences of deaths caused by some disease in several times of year is a typical example for circular data observations. However, none of these studies consider group sequential test for von Mises distribution. In this study, a group sequential test is suggested for the mean direction of the von Mises distribution with known and unknown concentration parameter.

This article is organized as follows: The von Mises distribution and the sequential probability ratio test ($SPRT$) are briefly reviewed in the second and the third sections, respectively. In the fourth section, Pocock's group sequential test is described for the mean of the normal distribution. In the fifth section, it is indicated that Pocock's group sequential test can be used for the mean direction of the von Mises distribution. An application of medical data and conclusions are given in the sixth and the seventh sections, respectively.

2 The Von Mises Distribution

The von Mises distribution is a symmetric distribution which is the most important model for unimodal samples of circular data and it plays the same role in circular statistical inference as the normal distribution on the line.

If a circular random variable θ has a von Mises distribution ($\theta \sim VM(\mu, \kappa)$), its probability density function (pdf) is given by

$$f(\theta; \mu, \kappa) = \frac{1}{2\pi I_o(\kappa)} e^{\kappa cos(\theta - \mu)} \quad , 0 \leq \theta < 2\pi \tag{2.1}$$

where $\kappa \geq 0$ and $0 \leq \mu < 2\pi$. Here, $I_o(\kappa)$ is a particular function of κ and it denotes the modified Bessel function of the first kind and order zero, and is defined by

$$I_o(\kappa) = \frac{1}{2\pi}\int_0^{2\pi} e^{\kappa cos\theta} d\theta = \sum_{r=0}^{\infty} (\frac{1}{r!})^2 (\frac{\kappa}{2})^{2r} \tag{2.2}$$

This function has the effect of scaling the distribution.

For sufficiently large κ, the von Mises distribution is related to the normal distribution. If $\kappa \to \infty$ and $\xi = \kappa^{1/2}(\theta - \mu)$, ξ is approximately distributed as standard normal distribution $(N(0,1))$ [11], [8].

Several properties of the von Mises distribution are similar to those of the normal distribution. For instance, it is completely determined by two parameters. The parameter μ is the mean direction. The von Mises density is unimodal and symmetrical about the mean direction μ. The mode of the distribution is at $\theta = \mu$ and antimode is at $\theta = \mu + \pi$. The parameter κ is the concentration parameter which measures the concentration around the mean direction. As κ approaches zero, the von Mises pdf approaches a uniform distribution and as κ increases, the distribution increasingly concentrated at μ. Due to these properties, the concentration parameter is similar to the variance of a normal distribution.

By giving a random sample $\theta_1, \theta_2, ..., \theta_n$ from $VM(\mu, \kappa)$, the log-likelihood function is given by

$$logL(\mu, \kappa; \theta_1, \theta_2, ..., \theta_n) = n[log2\pi + \kappa \bar{R} cos(\bar{\theta} - \mu) - logI_o(\kappa)]. \tag{2.3}$$

Then the maximum likelihood estimate $\hat{\mu}$ of μ is

$$\hat{\mu} = \bar{\theta} \tag{2.4}$$

where

$$\bar{\theta} = \begin{cases} tan^{-1}(\frac{\sum_{i=1}^n sin\theta_i}{\sum_{i=1}^n cos\theta_i}), & \sum_{i=1}^n cos\theta_i \geq 0 \\ tan^{-1}(\frac{\sum_{i=1}^n sin\theta_i}{\sum_{i=1}^n cos\theta_i}) + \pi, & \sum_{i=1}^n cos\theta_i < 0. \end{cases} \tag{2.5}$$

Differentiating (2.3) with respect to κ gives

$$\frac{logL(\mu, \kappa; \theta_1, \theta_2, ...\theta_n)}{\partial \kappa} = n\{\bar{R} cos(\bar{\theta} - \mu) - A(\kappa)\} \tag{2.6}$$

where $A(\kappa) = I_1(\kappa)/I_o(\kappa)$ is the ratio of two modified Bessel functions and $I_1(\kappa)$ is the imaginary Bessel function of order one. The maximum likelihood estimate $\hat{\kappa}$ of κ is the solution of

$$A(\hat{\kappa}) = \bar{R} \tag{2.7}$$

i.e.

$$\hat{\kappa} = A^{-1}(\bar{R}) \tag{2.8}$$

where $\bar{R}$ is the mean resultant length of the sample and is given by;

$$\bar{R} = \sqrt{\left(\frac{1}{n}\sum_{i=1}^{n} cos\theta_i\right)^2 + \left(\frac{1}{n}\sum_{i=1}^{n} sin\theta_i\right)^2}. \tag{2.9}$$

Values of functions A and A^{-1} are taken from the tables, such as Mardia and Jupp (2000, p. 362-363) and Fisher (1993, p. 224-225). A reasonable approximation to the solution of (2.8) can, also, be obtained by

$$\hat{\kappa} = \begin{cases} 2\bar{R} + \bar{R}^3 + 5\bar{R}^5/6, & \bar{R} < 0.53 \\ -0.4 + 1.39\bar{R} + 0.43/(1-\bar{R}), & 0.53 \leq \bar{R} < 0.85 \\ 1/(\bar{R}^3 - 4\bar{R}^2 + 3\bar{R}), & \bar{R} \geq 0.85 \end{cases} \tag{2.10}$$

[4, 11].

3 Sequential Probability Ratio Test for the Mean Direction

Let θ be a von Mises distributed random variable with a mean direction μ_0 and a concentration parameter. For testing $H_0 : \mu = \mu_0$ against $H_1 : \mu = \mu_1$, sequential probability ratio test is defined as follows; If the values of θ random variable is defined as $\theta_1, \theta_2, ..., \theta_n$, likelihood ratio is defined as,

$$L_n = \prod_{i=1}^{n} \frac{f(\theta_i; \mu_1)}{f(\theta_i; \mu_0)} = \frac{\frac{1}{[2\pi I_0(\kappa)]^n} e^{\kappa \sum_{i=1}^{n} cos(\theta_i - \mu_1)}}{\frac{1}{[2\pi I_0(\kappa)]^n} e^{\kappa \sum_{i=1}^{n} cos(\theta_i - \mu_0)}}. \tag{3.1}$$

Then by taking logarithm and simplifying, (3.1) can be written as;

$$lnL_n = \sum_{i=1}^{n} Z_i = 2\kappa \sum_{i=1}^{n} sin(\theta_i - v_1) sin(-v_2) \tag{3.2}$$

where $v_1 = \frac{\mu_0 + \mu_1}{2}$ and $v_2 = \frac{\mu_0 - \mu_1}{2}$.

At each stage of the test process, the value of $\sum_{i=1}^{n} Z_i$ is computed and compared with lnA and lnB critical values which depend on type-1(α) and type-2(β) errors. A and B values are computed as $A = \frac{1-\beta}{\alpha}$, $B = \frac{\beta}{1-\alpha}$. Then, one of the following decision is made.

1. If $\sum_{i=1}^{n} Z_i \leq lnB$, the process is terminated with the acceptance of H_0.
2. If $\sum_{i=1}^{n} Z_i \geq lnA$, the process is terminated with the rejection of H_0.
3. If $lnB < \sum_{i=1}^{n} Z_i < lnA$, the experiment is continued by taking an additional observation.

[17].

When μ is the test parameter for the von Mises distribution, the approximate formula for the operating characteristic (OC) function $P(\mu)$ is given by;

$$P(\mu) = \frac{A^h - 1}{A^h - B^h} \tag{3.3}$$

where $h = \frac{sin(\mu - v_1)}{sinv_2}$ [5, 6].

In linear data, acceptance probabilities are computed for the various values of h. Apart from the linear data, minimum and maximum values of operating characteristic function are obtained in circular data. Differentiating OC function with respect to μ, it is obtained that $\mu = 90^0 + v_1$ and $\mu = 270^0 + v_1$, and these can be shown to be a minimum and maximum, respectively.

An approximation to the average sample number function ASN(μ), which is the expected number of observations, is given by;

$$ASN(\mu) = \frac{P(\mu)lnB + [1 - P(\mu)]lnA}{2A(\kappa)sinv_1 sinv_2}. \tag{3.4}$$

It is possible to compute maximum and minimum values of the average sample number in circular data. Therefore, the average sample numbers, which are obtained when H_0 or H_1 is true in linear data, are computed for the maximum and minimum values in circular data. Differentiating the average sample number with respect to v_2 and setting that equal to zero gives;

$$ASN(\mu)_{min} = \frac{P(\mu)lnB + [1 - P(\mu)]lnA}{2A(\kappa)sinv_1}. \tag{3.5}$$

Since a minimum can be obtained in only one turning point, the ends of the range of v_2 will give the maximum. This leads to the point 0^0 and it gives

$$ASN(\mu)_{max} = \infty \tag{3.6}$$

[5, 13].

4 Pocock's Group Sequential Test

The basic concepts of Pocock's group sequential test in one sample are described as follows. Consider K groups (stages) of normally distributed observations with an unknown mean μ and a known variance σ^2, where in group k,$k = 1, 2, ..., K$ and $n_1 = n_2 = ...n_K = n$ observations are obtained. It is planned as a test of the null hypothesis $H_0 : \mu = \mu_0$ against the two sided alternative $H_1 : \mu \neq \mu_0$. Let $\bar{x}_j$ denote the mean response of the sample in the jth. group of n observations. In the jth stage, the normal score Z_j is given by

$$Z_j = \sqrt{n}(\bar{x}_j - \mu_0)/\sqrt{\sigma^2}. \tag{4.1}$$

The cumulative normal score

$$S_k = \textstyle\sum_{j=1}^{k} Z_j \quad , k = 1, 2, ...K \tag{4.2}$$

is the usual statistic for testing the hypothesis of the mean at type-I error probability α. Z_j is $N(0, 1)$ and $N(\Delta, 1)$ distributed, under H_0 and H_1 respectively. Where Δ is given as

$$\Delta = E(Z_j) = \sqrt{n}(\mu_1 - \mu_0)/\sqrt{\sigma^2} \tag{4.3}$$

[1, 9]. Formally the test process is as follows:

1. After group $k = 1, 2, ..., K-1$
   If $|S_k| \geq z_p(K,\alpha)\sqrt{k}$, stop, reject H_0
   otherwise, continue to group$k+1$

2. After group K
   If $|S_K| \geq z_p(K,\alpha)\sqrt{K}$, stop, reject H_0
   otherwise, stop, acceptH_0.

Where $z_p(K,\alpha)$ is the Pocock's critical value as in Table 1. The sample size per group is obtained as

$$n = \Delta^2 \left(\frac{\sqrt{\sigma^2}}{\mu_1 - \mu_2} \right)^2 \tag{4.4}$$

where Δ is the value of noncentrality parameter and it can be determined by a given value of $1-\beta$. The maximum sample size is $n_{max} = nK$. If $K = 1$ is taken as fixed sample size design (4.4) becomes the familiar sample size for a normal response. The average sample number, under H_1 is $ASN = n\bar{K}^*$, where $\bar{K}^*$ is the average number of stages.

$z_p(K,\alpha)$, Δ and $\bar{K}^*$ values are given in Table 1 for $k = 1, 2, .., 5$, $\alpha = 0,05$, $1-\beta = 0,95$. More complete tabulations of various values can be found in [14] and [9].

Table 1: Pocock's Critical Values $z_p(K,\alpha)$, Δ and $\bar{K}^*$ for $k = 1, 2, ..., 5$, $\alpha = 0,05$, $1-\beta = 0,95$

k	$z_p(K,\alpha)$					Δ	K^*
1	1,645					3,290	1
2	1,875	1,875				2,445	1,282
3	1,993	1,993	1,993			2,035	1,656
4	2,067	2,067	2,067	2,067		1,782	2,056
5	2,122	2,122	2,122	2,122	2,122	1,605	2,460

When the variance σ^2 is unknown, group sequential t-test is used. Test procedure is the same as the one with known σ^2. Since σ^2 is unknown, the pooled sample variance is estimated of n observations and is used for σ^2 in (4.1). Furthermore, sample size per group can not be calculated with (4.4) in group sequential t-test. Thus, the researcher supposed that each group contains n observations, in this case [9].

5 Group Sequential Test for the Mean Direction of the Von Mises Distribution

In this section, it is shown that Pocock's group sequential test can be used for the mean direction of the von Mises distribution both for known κ and unknown κ cases.

It is assumed that $\theta_1, \vartheta_2, ..., \vartheta_n$ is a random sample from a von Mises distribution $VM(\mu,\kappa)$.

Let the concentration parameter be known as $\kappa = \kappa_0 (\kappa_0 \geq 2)$. Then, the population mean resultant length of a von Mises distribution is ρ. The hypothesis to be tested is $H_0 : \mu = \mu_0$ against $H_1 : \mu \neq \mu_0$. From (2.3), the score statistic is defined as

$$\left.\frac{\partial logL(\mu,\kappa;\theta_1,\theta_2,...,\theta_n)}{\partial\mu}\right|_{\mu=\mu_0} = n\kappa\bar{R}sin(\bar{\theta}-\mu_0). \tag{5.1}$$

[3]. Under H_0, the score statistic is equal to

$$\sqrt{n\kappa_0\rho}sin(\bar{\theta}-\mu_0) \tag{5.2}$$

and it has approximately the distribution $N(0,1)$, for large n. The circular standard error of the mean direction for the von Mises distribution is

$$\sigma_{VM} = \frac{1}{\sqrt{n\kappa_0\rho}}. \tag{5.3}$$

Thus, the test statistic for the score test is given by

$$Z_{VM} = \frac{sin(\bar{\theta}-\mu_0)}{\sigma_{VM}} \tag{5.4}$$

[4]. Let $z_{\alpha/2}$ indicates the upper $100(\alpha/2)\%$ point and z_α indicates $100(\alpha)\%$ point of the standard normal distribution. Then the test of $H_0 : \mu = \mu_0$ against the alternatives are at the $100\alpha\%$ level are follows:

1. When $H_1 : \mu \neq \mu_0$: if $|Z_{VM}| > z_{\alpha/2}$, then reject H_0.
2. When $H_1 : \mu < \mu_0$: if $\mu_0 - \pi < \bar{\theta} < \mu_0$ and $Z_{VM} < -z_\alpha$, then reject H_0.
3. When $H_1 : \mu > \mu_0$: if $\bar{\theta} < \mu_0 + \pi$ and $Z_{VM} > -z_\alpha$, then reject H_0.

In the sense of the information given above, the group sequential test statistic for the von Mises distribution can be defined as:

$$S_{VMk} = \sum_{j=1}^{k} Z_{VMj}, \quad k = 1,...,K \tag{5.5}$$

where

$$Z_{VMj} = \sqrt{n\kappa_0\rho}sin(\bar{\theta}_j - \mu_0) \tag{5.6}$$

where $\bar{\theta}_j$ is computed from the data of n observations for the jth group. For $K = 1$, the test statistic (5.5) transforms into fixed sample test in the von Mises distribution. Therefore, since Z_{VMj} has approximately the distribution $N(0,1)$ under H_0, the group sequential test can be used for testing the mean direction of the von Mises distribution with the known concentration parameter. The test statistic S_{VMk} is compared with $z_p(K,\alpha)$ as follows:

After group $k = 1,2,...,K-1$

For $H_1 : \mu \neq \mu_0$, if $\lvert S_{VMk}\rvert \geq z_p(K,\alpha)\sqrt{k}$,	stop, reject H_0
For $H_1 : \mu < \mu_0$ and $\mu_0 - \pi < \bar{\theta_k} < \mu_0$, if $\lvert S_{VMk}\rvert < -z_p(K,\alpha)\sqrt{k}$,	stop, reject H_0
For $H_1 : \mu > \mu_0$ and $\bar{\theta_k} < \mu_0 + \pi$, if $\lvert S_{VMk}\rvert > -z_p(K,\alpha)\sqrt{k}$,	stop, reject H_0
otherwise continue to group $k+1$	

After group K

For $H_1 : \mu \neq \mu_0$, if $\lvert S_{VMk}\rvert \geq z_p(K,\alpha)\sqrt{K}$,	stop, reject H_0
For $H_1 : \mu < \mu_0$ and $\mu_0 - \pi < \bar{\theta_k} < \mu_0$, if $\lvert S_{VMk}\rvert < -z_p(K,\alpha)\sqrt{K}$,	stop, reject H_0
For $H_1 : \mu > \mu_0$ and $\bar{\theta_k} < \mu_0 + \pi$, if $\lvert S_{VMk}\rvert > -z_p(K,\alpha)\sqrt{K}$,	stop, reject H_0
otherwise stop, accept H_0.	

For this test, the group size n_{VM} is obtained from the expected value of the test statistic (5.6) under H_1;

$$\Delta = E(Z_{VMj|H_1}) = \sqrt{n_{VM}\rho\kappa}sin(\mu_1 - \mu_0). \tag{5.7}$$

Therefore, the value of n_{VM} for this test is

$$n_{VM} = \Delta^2 \frac{1}{[sin(\mu_1 - \mu_0)]^2\kappa\rho}. \tag{5.8}$$

The maximum sample size can be defined as

$$n_{max} = n_{VM}N \tag{5.9}$$

and the average sample number is

$$ASN_{VM} = n_{VM}\bar{K}^*. \tag{5.10}$$

Now, let the concentration parameter κ be unknown, and then the test statistic for the score test can be defined as

$$Z_{VM} = \frac{sin(\bar{\theta} - \mu_0)}{\hat{\sigma}_{VM}} \tag{5.11}$$

where

$$\hat{\sigma}_{VM} = \frac{1}{\sqrt{n\hat{\kappa}\bar{R}}}. \tag{5.12}$$

Therefore, Z_{VM} is approximately distributed as $N(0,1)$ under H_0.This approximation is satisfactory for the values of estimated concentration parameter $(\hat{\kappa})$ and sample size (n) which are given in Table 2 [4, 11].

Table 2: $\hat{\kappa}$ and n values for the test

$\hat{\kappa}$	n
$0,4 \leq \hat{\kappa} < 1$	$n \geq 25$
$1,0 \leq \hat{\kappa} < 1,5$	$n \geq 15$
$1,5 \leq \hat{\kappa} < 2,0$	$n \geq 10$
$\hat{\kappa} \geq 2,0$	All n

Then, as for group sequential test statistic, it can be defined as

$$Z_{VMj} = \sqrt{n\bar{R}_j\hat{\kappa}_j}sin(\bar{\theta}_j - \mu_0) \tag{5.13}$$

where $\bar{\theta}_j$,$\bar{R}_j$ and $\hat{\kappa}_j$ values are computed from the data of n observations for the jth group. The test proceeds as in the same way of known κ. Since κ is unknown, group size can not be calculated in (5.8). Therefore, group size is supposed by researchers.

To give an instance for the application of real-life data on wind directions, the following example compares the group sequential test for the von Mises distribution with known κ, with fixed sample test and SPRT.

Example 5.1: Wind directions, in Anadolu University Airport Eskisehir, are measured sequentially (hourly) in university's weather station. For this data set, κ is known as $\kappa = 4,58$ (corresponding $\rho = 0,88263$) and $\alpha = \beta = 0,05$ is supposed and the hypothesis is tested $H_0 : \mu = 141^0$ against $H_1 : \mu = 130^0$. Table 3 gives the maximum sample sizes and the expected sample sizes for the fixed sample, the sequential probability ratio, and the group sequential test.

Table 3: Comparison of the Fixed Sample, Sequential Probability Ratio and Group Sequential Tests for $\kappa = 4,58$, $\alpha = \beta = 0,05$, $H_0 : \mu = 141^0$, $H_1 : \mu = 130^0$ (von Mises response with known κ)

Tests		**Maximum Sample Size**		**Average Sample Number**	
Fixed Sample Test		73,545		73,545	
				minimum	maximum
SPRT		∞		2,380	∞
		Group sizes	n_{max}		
Group Sequential Test	$K = 2$	40,618	81,236	52,072	
	$K = 3$	28,138	84,413	46,596	
	$K = 4$	21,576	86,305	44,361	
	$K = 5$	17,503	87,515	43,057	

Other examples can be presented that have the same general principle with different choices of α,β,μ_0,μ_1 and κ. Pocock [15] compared those tests for the mean of the normal distribution and showed that GST is more advantageous than the fixed sample test and SPRT in terms of sample size; in addition , Bacanlı & Demirhan [1] proposed the group sequential test for the mean of the inverse Gaussian distribution, in a similar way and showed that this test is more advantageous than the others.

Thus, it is seen that these results are, also, valid for circular normal distribution that is known as Von-mises distribution.

6 Application to Medical Data

In this section, the group sequential test is applied to a medical circular data set. The medical data were collected from sequentially patients who was male and female and between the age of 44 and 75 in *Eskiehir Private Fizyomer Rehabilitation and Physical Therapy Center* between the years of 2010 and 2013. These patients were admitted to the center with complaints of pain in their shoulders. After the physical examination, some problems were detected in patients such as shoulder joint motions are painful and, also, partially restrictive and so on. Then, the range of motion the shoulder joints of patients were measured. These measurements include active and passive angular values for *flexion, extension, abduction, internal rotation* and *external rotation* variables. After the patients were diagnosed with the *adhesive capsulitis of shoulder* (also known as *the frozen shoulder*), 30 sessions of physical therapy and rehabilitation were applied to them and the range of motion of the shoulder joints were measured again. After the therapy, it is aimed that the patients will reach a complete joint range of motion in all of the shoulder motions. In this study, the group sequential test is applied for the *internal rotation (passive)* variable which is obtained after the therapy in the data set. In anatomy, *internal rotation* (also known as *medial rotation*) is a term that refers to the rotation towards the center of the body [18] and the term *passive* means that the patient moves with an external support or assistance.

It is theorized that a healthy, "perfect" shoulder should have 90 degrees of internal rotation [19]. Therefore, the group sequential test is applied for $H_0 : \mu = 90^0$ against the alternative $H_1 : \mu = 80^0$ with $\alpha = \beta = 0,05$ and $K = 4$. The concentration

parameter is unknown, so the group sizes are supposed as $n = 5$. GST results are given in Table 4.

Table 4: $\hat{\kappa}$ and n values for the test

j	1	2	3	4
n	5	5	5	5
$\bar{\theta}_j$	84,133	85,031	86,012	85,031
$\hat{\kappa}_j$	11,486	27,181	47,768	27,181
$\bar{R}_j$	0,978	0,991	0,995	0,991
Z_{VMj}	-0,766	-1,005	-1,072	-1,005
S_{VMk}	-0,766	-1,771	-2,843	-3,848
$Z_p(4;0,05)\sqrt{k}$	2,067	2,923	3,580	4,134
Decision	Continue	Continue	Continue	Accept H_0

When Table 4 results are examined, it can be seen that, in stage 4;

$$S_{VM4} = 3,848 > -Z_P(4;0,05)\sqrt{4} = -2,067(2) = -4,134$$

hence we stop and accept H_0.

Therefore, researchers can apply the group sequential test for predetermined α, β, N and n values.

7 Discussion and Conclusions

As in many scientific fields, the most common probability distribution in medical applications of circular data is the von Mises distribution. However, the group sequential tests are often used in medical researches which the data is collected sequentially. Therefore, the group sequential test for the mean of the distributed von Mises data is proposed in this study.

In medical studies, a significant amount of the collected data is in the form of circular. In the literature, there are fixed sample and sequential probability ratio tests for circular data. However, in medical studies, the use of these tests is very difficult in terms of obtaining required sample sizes. The reason of this is that, when SPRT is used in the studies in which the data is collected sequentially, the expected sample size and the maximum sample size are infinite (see Table 3). Therefore, these values cannot be predetermined before the test. In this study, the group sequential test have been proposed for circular data. An application of this test for a medical data set (shoulder internal rotation angles) is carried out and it is shown that the advantages of the test are also valid for circular data.

In GST, researchers can determine required maximum sample size and expected sample size values for their hypotheses, determined α and β probabilities and K values. In this respect, using GST provides a great advantage. GST was generated for linear data in the literature. In this study, GST is defined for circular data and it is indicated that GST can be used for the mean of the von Mises distribution which is frequently encountered in medical studies.

Acknowledgement
The authors would like to thank to *Dr. Türkan Tünerir* and *Fizyomer Rehabilitation and Physical Therapy Center, Eskisehir, Turkey* for their support and giving the rights of using their medical data.

References

[1] Bacanlı, S. and Parlak Demirhan, Y. *A group sequential test for the inverse Gaussian mean*, Statistical Papers 49, 377-386, 2008.
[2] Batschelet, E. *Circular Statistics in Biology AcademicPress, London*, 1981.
[3] Casella, G. and Berger, R.L. *Statistical Inference, 2nd Ed.* (Duxbury Press, California, 2002).
[4] Fisher, N.I. *Statistical Analysis of Circular Data* (Cambridge University Press, Cambridge, 1993).
[5] Gadsden, R.J. and Kanji, G.K. *Sequential analysis for angular data*, The Statistician 30 (2), 119-129, 1981.
[6] Gadsden, R.J. and Kanji, G.K. *Sequential analysis applied to circular data*, Sequential Analysis 1 (4), 304-314, 1982.
[7] Hussin, A.G., Hassan, S.F. and Zubairi, Y.Z. *A statistical method to describe the relationship of circular variables simultaneously*, Pakistan Journal of Statistics 26 (4), 593-607, 2010.
[8] Jammalamadaka, S.R. and SenGupta, *A Topics in Circular Statistics* (World Scientific Publishing, London, 2001).
[9] Jennison, C. and Turnbull, B.W. *Group Sequential Methods with Applications to Clinical Trials* (Chapman & Hall, New York, 2000).
[10]Lan, K.K.G. and DeMets, D.L. *Discrete sequential boundaries for clinical trials*, Biometrika 70 (3), 659-663, 1983.
[11]Mardia, K.V. and Jupp, P.E. *Directional Statistics* (John Wiley, Chicester, 2000).
[12] O'Brien, P.C. and Fleming, T.R. *A multiple testing procedure for clinical trials*, Biometrics 35, 549-556, 1979.
[13] Peker, K.. and Bacanl, S. *A sequential test for the mean direction applied to circular data and an application*, Eskiehir Osmangazi University Journal of Engineering Architecture Faculty 20 (2), 1-13, 2007.
[14] Pocock, S.J. *Group sequential methods in the design and analysis of clinical trials*, Biometrika 64, 191-199, 1977.
[15] Pocock, S.J. *Interim analysis for randomized clinical trials: the group sequential approach*, Biometrics 38, 153-162, 1982.
[16] SenGupta, S. and Rao, J.S. *Statistical analysis of crossbedding azimuths from the Kamthi formation around Bheemaram, Pranhita Godavari Valley*, Sankhyá 28, 165-174, 1967.
[17] Wald, A. *Sequential Analysis* (John Wiley, New York, 1947).
[18] $https://en.wikipedia.org/wiki/List\_of\_internal\_rotators\_of\_the\_human\_body$
[19] $http://saptstrength.com/2013/08/09/shoulder-range-of-motion-for-the-overhead-athlete/$