Journal of Statistical Planning and Inference 46 (1995) 77–91

journal of statistical planning and inference

An almost sure representation of sample circular median

Sumitra Purkayastha*

Division of Theoretical Statistics and Mathematics, Indian Statistical Institute, Calcutta 700035, India

Received 3 March 1993; revised 7 March 1994

Abstract

We consider arc distance median for circular data. It is shown to admit an almost sure asymptotic linear representation. In the process of deriving this representation, an algorithm to obtain a suitable version of this statistic for a set of observations from the circle is also developed.

AMS Subject Classifications: 62G20, 62H11

Key words: Circular median; Bahadur representation; Asymptotic normality

1. Introduction

For a set of observations on the circle, one can define a location of its central tendency by minimizing the sum of geodesic distances of an arbitrary point on the circle from these points and by simultaneously searching for the point where this minimum is attained. Liu and Singh (1992) have called it the arc distance median for circular data. This is also known as the Mardia–Fisher median for circular data (Small, 1990) and is in spirit the same as Mardia's (1972) median. This notion of circular median (henceforth abbreviated to c-median) has been studied by Liu and Singh (1992) in the context of their notion of depth for directional data, and also by Purkayastha (1991a) in characterization theory. Also see Lenth (1981), Wehrly and Shine (1981), and Ko and Guttorp (1988) for other works in c-median.

In this paper we undertake the study of some aspects of asymptotics for the c-median. One of the major tools for studying the large sample behaviour of a statistic consists in deriving an asymptotic linear representation of the same. For sample quantiles, such a representation was obtained by Bahadur (1966), which came to be known later as 'Bahadur representation of quantiles'. Also see Kiefer (1967), Sen (1968), Ghosh (1971), Dutta and Sen (1971), Sen and Ghosh (1971), Sen (1972), and de Haan and Taconis-Haantjes (1979) for related works in linear data. It is therefore worthwhile to investigate whether sample c-median admits any such representation. This investigation is the subject matter of this paper.

The organization of the paper is as follows. In Section 2, we state the set of assumptions under which the main representation theorem is proved and also record some immediate consequences of these assumptions. In Section 3, we state a property of population c-median which is analogous to the one connecting minimum mean absolute deviation with the linear median. In Section 4, we develop an algorithm to obtain a suitable version of c-median and study this version accordingly. This algorithm is also worth recording in view of the fact that our definition of sample c-median involves minimizing a function (depending on the data) over the circle. Employing the major findings of Sections 2–4, we finally establish strong consistency of sample c-median in Section 5 and prove the main representation theorem in Section 6.

2. Preliminaries

We begin by introducing the definition of an interval on the circle and also of the antipodal point of an arbitrary point on the circle.

Note. For two numbers $a, b \in [0, 2\pi)$ with $a < b$, we denote the set $[0, a] \cup [b, 2\pi)$ by $[b, a]$. If, however, $a > b$, $[b, a]$ denotes the usual closed interval of $\mathbb{R}$. Observe that $[b, a]$ indeed denotes a closed interval over the circle. The notation for an open, or a half-open, interval over the circle is similar. We also define for any $x \in [0, 2\pi)$, its antipodal point, denoted x^*, by

$$x^* = x + \pi((\text{mod } 2\pi)).$$

Let $\{X_i\colon i \geqslant 1\}$ be a sequence of independent and identically distributed circular random variables defined over some probability space $\{\Omega, \mathscr{A}, \mathscr{P}\}$. This means, $0 \leqslant X_i(\omega) < 2\pi$ for all $\omega \in \Omega$ and for every $i \geqslant 1$ Denote the distribution function of X_1 by F. We make the following assumptions about F:

(A.1) F is continuous.

(A.2) There exists a unique $\theta \in [0, \pi)$, denoted θ_0, such that $G(\theta) = \frac{1}{2}$, where $G\colon [0, 2\pi) \to [0, 1]$ is defined as

$$G(x) = P(X_1 \in (x, x^*]). \tag{2.1}$$

(A.3) There exists $\varepsilon > 0$ such that $F(x)$ is differentiable either (a) at every $x \in (\theta_0 - \varepsilon, \theta_0 + \varepsilon) \cup (\theta_0^* - \varepsilon, \theta_0^* + \varepsilon)$, if $0 < \theta_0 < \pi$, or (b) at every $x \in (2\pi - \varepsilon, \varepsilon) \cup$

$(\pi-\varepsilon, \pi+\varepsilon)\cup\{2\pi\}$, if $\theta_0=0$. In case (b), $F'(x)$ is to be understood as the right-derivative when $x=0$, and as the left-derivative when $x=2\pi$. In this case, we stipulate further that $F'(0)=F'(2\pi)$. We denote $F'(x)$, whenever it exists, by $f(x)$.

(A.4) There exists a positive constant K such that either (a) $|f(x)-f(\theta_0)|<K|x-\theta_0|^{1/2}$, if $|x-\theta_0|<\varepsilon$ and $|f(x)-f(\theta_0^*)|<K|x-\theta_0^*|^{1/2}$, if $|x-\theta_0^*|<\varepsilon$, whenever $0<\theta_0<\pi$, or (b) $|f(x)-f(0)|<K|x|^{1/2}$, if $0\leqslant x<\varepsilon$, $|f(x)-f(\pi)|<K|x-\pi|^{1/2}$, if $|x-\pi|<\varepsilon$ and $|f(x)-f(2\pi)|<K|x-2\pi|^{1/2}$, if $2\pi-\varepsilon<x\leqslant 2\pi$, whenever $\theta_0=0$.

(A.5) $f(\theta_0)\neq f(\theta_0^*)$.

Let us now define

$$M=\begin{cases}\theta_0 & \text{if } f(\theta_0)>f(\theta_0^*),\\ \theta_0^* & \text{if } f(\theta_0)<f(\theta_0^*).\end{cases}$$

Thus,

$$f(M)>f(M^*). \tag{2.2}$$

In consideration of (A.2) and (2.2), we therefore have the following fact.

Fact 2.1. *M is the unique population c-median (in the sense of Mardia* (1972, *p.* 28)) *of X_1.*

Remark 2.1. It should be pointed out that in order to define c-median of an arbitrary circular random variable X having distribution function F, it is not necessary to have such a strong requirement as the existence of density of X (see Mardia, 1972, p. 28). It is enough to have differentiability of F at M and M^*, where $G(M)=\frac{1}{2}$, and then demand (2.2). Fact 2.1 indeed asuumes this modified definition of Mardia's (1972) median. We shall see later (Theorem 3.1) that M will also be the unique arc distance median of X_1.

We conclude this section with the following result that studies the behaviour of the function $G(x)$. This result will be needed to prove the main representation theorem in Section 6.

Lemma 2.1. *Assume* (A.1)–(A.5). *Then we have the assertion* (a) *below if $0<M<\pi$, and* (b) *if $M=0$.*

(a) $G(x)>\frac{1}{2}$ *if* $x\in(M^*, M)$, $=\frac{1}{2}$ *if* $x\in\{M, M^*\}$, $<\frac{1}{2}$ *if* $x\in(M, M^*)$.

Moreover, there exists $\delta>0$ such that G is strictly increasing on $[M-\delta, M+\delta]$ and strictly decreasing on $[M^*-\delta, M^*+\delta]$.

(b) $G(x)>\frac{1}{2}$ *if* $x\in(0,\pi)$, $=\frac{1}{2}$ *if* $x\in\{0,\pi\}$, $>\frac{1}{2}$ *if* $x\in(\pi, 2\pi)$.

Moreover, there exists $\delta>0$ such that G is strictly decreasing on $[2\pi-\delta, \delta]$ and strictly increasing on $[\pi-\delta, \pi+\delta]$.

Proof. We shall provide the proof for part (a) only. The proof for part (b) is similar.

The first assertion follows from the facts that G is continuous on $[0, 2\pi)$, $G(x) = \frac{1}{2}$ if and only if $x = M$ or M^*, $G'(M)(= f(M^*) - f(M))$ is negative, and $G(x) + G(x^*) = 1$ for any $x \in [0, 2\pi)$. These facts are easy to verify. Details can be found in Purkayastha (1991b).

In order to verify the second assertion, first we observe that $G'(x) = f(x^*) - f(x)$ for $x \in (M - \varepsilon, M + \varepsilon) \cup (M^* - \varepsilon, M^* + \varepsilon)$. In consideration of (A.4) and the definition of M, this implies our assertion. □

Remark 2.2. If $M = \pi$ or $\pi < M < 2\pi$, it is possible to establish results analogous to part (a) or part (b) of the above lemma by using the fact that for any $x \in [0, 2\pi)$, $G(x) + G(x^*) = 1$.

3. A property of population circular median

We begin by stating the fact that for any integer $p \geqslant 1$, (S^p, d) is a compact metric space, where $S^p = \{\boldsymbol{x} \in \mathbb{R}^{p+1}: \|\boldsymbol{x}\| = 1\}$, and $d(\boldsymbol{x}, \boldsymbol{y}) = \cos^{-1}(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y})$ [=the unique angle $\theta \in [0, \pi]$ such that $\cos\theta = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{y}$], $\boldsymbol{x}, \boldsymbol{y} \in S^p$. This metric d is indeed the geodesic distance on the sphere. For $p = 2$, the relevance of 'd' in the study of spherical median can be found in Fisher (1985). See also Cabrera and Watson (1990).

We shall, however, focus our attention on the case corresponding to $p = 1$. Observe that if we identify $S^1 = \{(\cos\theta, \sin\theta): 0 \leqslant \theta < 2\pi\}$ with $[0, 2\pi)$ through the identification $(\cos\theta, \sin\theta) \equiv \theta$, the metric d^*, induced by d on $[0, 2\pi)$, has the form $d^*(\theta_1, \theta_2) = \pi - |\pi - |\theta_1 - \theta_2||$, $\theta_1, \theta_2 \in [0, 2\pi)$. We agree to write d in place of d^*.

Consider now the function $D: [0, 2\pi) \to \mathbb{R}$ defined as follows:

$$D(y) = E\{d(X_1, y)\}. \tag{3.1}$$

We state below a few properties of the function D. These properties are easy to verify. Details can be found in Purkayastha (1991b). See also Mardia (1972).

Lemma 3.1. (a) *$D(y) + D(y^*) = 1$ for every $y \in [0, 2\pi)$.*

(b) *$\lim_{y \to 2\pi-} D(y) = D(0)$ and D is uniformly continuous on $[0, 2\pi)$.*

(c) *D is differentiable on $[0, 2\pi)$ with $D'(y) = 1 - 2G(y)$, $0 \leqslant y < 2\pi$. (For $y = 0$, $D'(y)$ is indeed the right derivative.)*

Remark 3.1. It should be mentioned that parts (a) and (b) of the above lemma are true for any circular random variable, i.e., without any restriction of F. Part (c), however, requires the additional assumption of continuity of F, i.e., (A.1).

We shall conclude this section by stating the following theorem, which is indeed the goal we have set out with. Before we state it, we note that in view of part (b) of the preceding lemma it makes sense to talk if $\min\{D(y): 0 \leqslant y < 2\pi\}$. For the result to hold we need the assumptions (A.1), (A.2), (A.4), (A.5) and differentiability of F at θ_0 and θ_0^*. The proof follows by employing Lemmas 2.1 and 3.1(c).

Theorem 3.1. $\{z \in [0, 2\pi): D(z) = \min_{0 \leqslant y < 2\pi} D(y)\} = \{M\}$.

Remark 3.2. In Mardia (1972, p. 31), a result similar to the above theorem can be found. See also Liu and Singh (1992).

4. A definition of sample circular median

We develop our definition of sample circular median in this section. It is indeed a version of sample arc distance median for circular data (Liu and Singh, 1992), or equivalently of Mardia–Fisher median for circular data (Small, 1990). This definition is obtained by minimizing the sample analogue of the function D, defined in (3.1). We shall address the issue of minimization of the sample analogue, and obtain our definition accordingly. We shall see that sample circular median so defined enjoys a natural property (Theorem 4.1). This property will, moreover, be used to prove the main representation theorem in Section 6. It should be mentioned here that our definition coincides with the one proposed in Purkayastha (1991a), where we restricted our attention only to small sample sizes, viz. $n = 2, 3$, and 4.

To begin with, we choose and fix $x_i \in [0, 2\pi)$, $i = 1, \ldots, n$. We now define

$$\tilde{d}(x) = \sum_{i=1}^{n} d(x, x_i), \quad 0 \leqslant x < 2\pi. \tag{4.1}$$

Observe that apart from a multiplicative constant $(= 1/n)$, $\tilde{d}(x)$ is indeed the sample analogue of $D(x)$, as defined in (3.1). Observe, moreover, that as in the case with the population c-median, it makes sense to talk of $\min\{\tilde{d}(x): 0 \leqslant x < 2\pi\}$. We now prove two lemmas, which settle the problem of minimizing $\tilde{d}(x)$ over $x \in [0, 2\pi)$.

Lemma 4.1. *There exists* $t \in \{1, \ldots, n\}$ *such that*

$$\tilde{d}(x_t) = \min\{\tilde{d}(x): 0 \leqslant x < 2\pi\}.$$

Proof. Let us write $\{x_1, \ldots, x_n\} \cup \{x_1^*, \ldots, x_n^*\} = \{z_1, \ldots, z_l\}$, where $z_i \in [0, 2\pi)$, $1 \leqslant i \leqslant l$, and $z_1 < \cdots < z_l$. Define the sets $A_1, \ldots, A_l$ by $A_j = [z_{j-1}, z_j]$ for $j = 1, \ldots, l$, $(z_0 \equiv z_l)$. It is now easy to verify that $\tilde{d}$ is linear on each of $[0, z_1], [z_1, z_2], \ldots, [z_{l-1}, z_l], [z_l, 2\pi]$ with $(\tilde{d}(0) - \tilde{d}(z_l))/(2\pi - z_l) = (\tilde{d}(z_1) - \tilde{d}(0))/z_1$ in case $z_1 > 0$. Therefore we must have $\min\{\tilde{d}(x): 0 \leqslant x < 2\pi\} = \tilde{d}(z_{j0})$ for $j_0 \in \{1, \ldots, l\}$. The rest of the proof consists in establishing that $z_{j0} = x_{t_0}$ for some $t_0 \in \{1, \ldots, n\}$.

Suppose, on the contrary, that $z_{j0} \neq x_t$ for every $t \in \{1, \ldots, n\}$, and that $z_{j0} = x_p^*$ for some p. This implies in turn that there exists an integer p' such that $x_p \neq x_{p'}$. Therefore it makes sense to define the following function:

$$\tilde{d}_1(x) = \sum d(x, x_i), \quad 0 \leqslant x < 2\pi,$$

where the sum is taken over all those i for which $x_i \neq x_p$. Now define the sets $B_1, \ldots, B_{l-1}$ by $B_j = (z_{j-1}, z_{j+1})$ for $j = 1, \ldots, l-1$, $(z_{l+1} \equiv z_0)$. Observe that $\tilde{d}_1$ is linear on B_{j0} and that $\tilde{d}_1(z_0) \leqslant \tilde{d}_1(z_{j0})$ for some $z_0 \in B_{j0}$ with $z_0 \neq z_{j0}$. Moreover, $d(x.x_p) \leqslant d(z_{j0}, x_p)$ for every $x \in B_{j0}$, with equality if and only if $x = z_{j0}$. It is now easy to see that $\tilde{d}(z_0) < \tilde{d}(z_{j0})$, contradicting the definition of z_{j0}. This contradiction establishes our assertion. □

Now we introduce the following sets of integers which will be needed in the next lemma.

$$\text{(a)} \quad N_i(n) = \begin{cases} \{1, \ldots, n\} - \{1, 2, n\} & \text{if } i = 1, \\ \{1, \ldots, n\} - \{i-1, i, i+1\} & \text{if } i = 2, \ldots, n-1, \\ \{1, \ldots, n\} - \{1, n-1, n\} & \text{if } i = n; \end{cases}$$

whenever n is an even number greater than 2, and

$$\text{(b)} \quad N_i(n) = \{1, \ldots, n\} - \{i\} \quad \text{for } i = 1, \ldots, n,$$

whenever n is an odd number greater than 2.

Lemma 4.2. *Suppose $x_i \in [0, 2\pi)$, $i = 1, \ldots, n (n > 2)$, are such that $x_1, \ldots, x_n, x_1^*, \ldots, x_n^*$ are all distinct. Denote by $x_{(1)} < \cdots < x_{(n)}$, the x_i's arranged in increasing order. Suppose, moreover, that for every $i = 1, \ldots, n$, $\tilde{d}(x_{(i)}) \neq \tilde{c}(x_{(j)})$ for all $j \in N_i(n)$. Then,*

(a) *whenever n is even,*

$$\left\{ y \in [0, 2\pi) : \tilde{d}(y) = \min_{0 \leqslant x < 2\pi} \tilde{d}(x) \right\} = [x_{(i-1)}, x_{(i)}], \tag{4.2}$$

for some $i (x_{(0)} \equiv x_{(n)})$ and

(b) *whenever n is odd,*

$$\left\{ y \in [0, 2\pi) : \tilde{d}(y) = \min_{0 \leqslant x < 2\pi} \tilde{d}(x) \right\} = \{x_{(i)}\},$$

for some i.

Note. The quantities or sets that appear in the proofs of both Lemmas 4.1 and 4.2 will be assumed to bear the same meaning.

Proof. We shall provide the proof for part (a) only: The proof for part (b) is similar.

We begin with the observation that there exist real constants $g_1, \ldots, g_l, g$ and $c_1, \ldots, c_l, c$ such that $\tilde{d}(x) = g_j(x) + c_j$ for $x \in A_j$, $j = 1, \ldots, l$, and $\tilde{d}_1(x) = gx + c$ for $x \in B_{j0}$. Observe now that for every fixed $y \in [0, 2\pi)$, $x \to d(x, y)$ is a piecewise linear continuous function with every line segment having gradient +1 or −1 which implies that every g_i, being a sum of even(=n) number of +1 or −1, is an even number, positive, negative, or zero. Observe also that $g = g_{j0} + 1 = g_{j0+1} - 1$ $(g_{l+1} \equiv g_1)$, which in turn implies $g_{j0+1} - g_{j0} = 2$.

The next step consists in proving that either

$$g_{j0} = 0 \quad \text{and} \quad g_{j0+1} = 2, \tag{4.3}$$

or

$$g_{j0} = -2 \quad \text{and} \quad g_{j0+1} = 0. \tag{4.4}$$

In view of the last sentence of the preceding paragraph it suffices to show that $g_{j0} \leqslant 0$ and $g_{j0+1} \geqslant 0$. Suppose, if possible, that $g_{j0} > 0$ or $g_{j0+1} < 0$. In the first case we have, moreover, $g_{j0+1} > 0$ and it is easy to obtain $y \in A_{j0}$ with $y \neq z_{j0}$ such that $\tilde{d}(y) < \tilde{d}(z_{j0})$, contradicting the definition of z_{j0}. In the second case we are led again to the same contradiction. Hence either (4.3) or (4.4) holds.

As a consequence of (4.3), it follows that $\tilde{d}$ is constant on A_{j0} and that $\tilde{d}(z_{j0}) < \tilde{d}(z_{j0+1})$. It is now easy to see that C_{t_0} is a subset of the set appearing on the left-hand side of (4.2). We omit the details. Similarly, from (4.4) we arrive at the same conclusion with C_{t_0} replaced by C_{t_0+1}. The rest of the proof indeed follows from the conditions of the lemma. □

Remark 4.1. It should be mentioned that in the study of linear data also we have assertions similar in spirit to the assertions of the above lemma and, moreover, the relationship between such assertions and the linear median is well known. We shall mimic that idea in order to define sample c-median.

We now introduce the following sets which will be needed in our definition:

$$U_n = \{(x_{(1)}, \ldots, x_{(n)}) \in \mathbb{R}^n \colon 0 \leqslant x_{(1)} \leqslant \cdots \leqslant x_{(n)} < 2\pi\},$$

$$A_n(\boldsymbol{x}^{(n)}) = \left\{ y \in [0, 2\pi) \colon \sum_{i=1}^{n} d(y, x_{(i)}) = \min_{0 \leqslant x < 2\pi} \sum_{i=1}^{n} d(x, x_{(i)}) \right\}, \quad \boldsymbol{x}^{(n)} \in U_n,$$

$$C_{i,n}(\boldsymbol{x}^{(n)}) = \begin{cases} [x_{(i-1)}, x_{(i)}], & i = 1, \ldots, n, \ (x_{(0)} \equiv x_{(n)}), \text{ whenever } n \text{ is even}, \\ \{x_{(i)}\}, & i = 1, \ldots, n, \text{ whenever } n \text{ is odd}, \end{cases}$$

$$E_{i,n} = \{\boldsymbol{x}^{(n)} \in U_n \colon s \neq t \Rightarrow |x_{(s)} - x_{(t)}| > 0 \text{ and } \neq \pi,\ A_n(\boldsymbol{x}^{(n)}) = C_{i,n}(\boldsymbol{x}^{(n)})\},$$
$$1 \leqslant i \leqslant n, \quad n \geqslant 1,$$

$$E'_{i,n}=\left\{x^{(n)}\in U_n:\ s\neq t\Rightarrow |x_{(s)}-x_{(t)}|>0 \text{ and } \neq\pi,\right.$$
$$\left.\sum_{j=1}^{n} d(x_{(i)},x_{(j)})\neq\sum_{j=1}^{n} d(x_{(p)},x_{(j)}) \text{ for all } p\in N_i(n)\right\},\quad 1\leqslant i\leqslant n,\ n\geqslant 2,$$

$$E'_n=\bigcap_{i=1}^{n} E'_{i,n},$$

$$E_n=\bigcup_{i=1}^{n} E_{i,n}.$$

Note in view of Lemma 4.2 that

$$E'_n\subseteq E_n. \tag{4.5}$$

Definition 4.1. Suppose $x_1,\ldots,x_n\in[0,2\pi)$ $(n>2)$. Denote the x_i's, arranged in increasing order by $x_{(1)},\ldots,x_{(n)}$, and write $x^{(n)}=(x_{(1)},\ldots,x_{(n)})$. The c-median of $x_1,\ldots,x_n$, denoted m_n, is defined as follows:

$$\text{(a)}\ m_n=\begin{cases}\dfrac{x_{(1)}+2\pi+x_{(n)}}{2}\ (\text{mod}\ 2\pi) & \text{if } x^{(n)}\in E_{1,n},\\ \dfrac{x_{(i-1)}+x_{(i)}}{2} & \text{if } x^{(n)}\in E_{i,n},\ i=2,\ldots,n,\\ x_{(j)} & \text{if } x^{(n)}\notin E_n, \text{ and } j=\min\{t:\ x_{(t)}\in A_n(x^{(n)})\},\end{cases}$$

whenever n is even;

$$\text{(b)}=\begin{cases}x_{(i)} & \text{if } x^{(n)}\in E_{i,n},\ i=1,\ldots,n,\\ x_{(j)} & \text{if } x^{(n)}\notin E_n, \text{ and } j=\min\{t:\ x_{(t)}\in A_n(x^{(n)})\},\end{cases}$$

whenever n is odd.

Remark 4.2. The definition of sample c-median for $n=1$ or 2 is obvious.We do not state them formally.

Remark 4.3. The idea behind the definition of m_n, when n is even, is the following: recall that $[0,2\pi)$ is a representation of the circle. Thus, for $x^{(n)}\in E_n$, $C_{i,n}(x^{(n)})$ is indeed an 'interval' for every $i=1,\ldots,n$. The c-median m_n is then defined as the mid-point of this interval.

Remark 4.4. Even though our definition of sample c-median appears to be somewhat cumbersome, a close scrutiny of the proofs of Lemmas 4.1 and 4.2 indeed reveals that in order to implement the algorithm described in the definition we need only to compute $\sum_{i=1}^{n}\{\pi-|\pi-|x_{(t)}-x_{(i)}||\}$ for $t=1,\ldots,n$.

The final result of this section indeed states that sample c-median as defined above enjoys a natural property. However, we need some more notation before stating it. We denote by M_n, the sample c-median of $X_1, \dots, X_n$, a set of *i.i.d.* observations from a circular distribution F. Moreover, we denote by G_n, the function obtained from (2.1) by replacing F with F_n, the empirical distribution function corresponding to F, and by $D_n(y)$, the sample analogue of $D(y)$ defined in (3.1). We assume that M_n is measurable.

Theorem 4.1. *Assume that F is continuous. Then, we have*

(a) $G_n(M_n) = \frac{1}{2}$ *a.s. whenever n is even, and*

(b) $G_n(M_n) = \frac{1}{2} - \frac{1}{2n}$ *a.s. whenever n is odd.*

Proof. We shall provide the proof for part (a) only. The proof for part (b) is similar.

We begin with the observation that $P(X^{(n)} \in E'_{i,n}) = 1$ for every $i = 1, \dots, n$, which is easy to verify. In view of (4.5), it therefore suffices to show that $X^{(n)} \in E_n$ implies that $G_n(M_n) = \frac{1}{2}$.

Choose $\omega \in \Omega$ such that $X^{(n)}(\omega) \equiv (X_{(1)}(\omega), \dots, X_{(n)}(\omega)) \in E_{i,n}$ for some i with $1 \leqslant i \leqslant n$, say i_0. This implies $D_n(y)$ restricted to $C_{i_0,n}(X^{(n)}(\omega))$ is a constant.

Observe now that for $x \in [0, 2\pi) - \{X_{(1)}(\omega), \dots, X_{(n)}(\omega), X^*_{(1)}(\omega), \dots, X^*_{(n)}(\omega)\}$,

$$D_n(x) = n[1 - 2G_n(x)]x - \sum X_{(i)}(\omega) + \sum X_{(i)}(\omega) + \sum (2\pi - x_{(i)}(\omega)) \tag{4.6}$$

if $0 \leqslant x < \pi$, where the first sum in (4.6) is taken over all those i for which $0 \leqslant X_{(i)}(\omega) < x$, the second one over all those i for which $x < X_{(i)}(\omega) < x + \pi$, and the last one over all those i for which $x + \pi < X_{(i)}(\omega) < 2\pi$. A similar relation holds if $\pi \leqslant x < 2\pi$. Let us recall now from Lemma 4.2, part (a) and Definition 4.1 that $M_n \in C_{i_0,n}(X^{(n)}(\omega))$. Taken with the last sentence of the preceding paragraph and (4.6), this implies $G_n(M_n) = \frac{1}{2}$, completing the proof. □

Remark 4.5. In view of Fact 2.1, Theorems 3.1 and 4.1, it is clear that arc distance median is indeed a meaningful version of Mardia's (1972) median. See also some of the discussions of Liu and Singh (1992) in this context.

Remark 4.6. In Purkayastha (1991b), some more properties enjoyed by sample c-median M_n can be found.

5. Strong consistency of sample circular median

We begin by recalling that $([0, 2\pi), d)$ is a compact metric space. In view of this fact, Theorem 3.1, and our definition of sample c-median, we can now adapt suitably Theorem (ii) of Sverdrup-Thygeson (1981) to establish the following result. For the result to hold, we need the same assumptions as the ones we did for Theorem 3.1 to hold. We omit the details.

Theorem 5.1. $\lim_{n\to\infty} d(M_n, M)=0$.

6. Main representation theorem

We need a few lemmas to prove the main representation theorem. The first result is true for any circular distribution F.

Lemma 6.1. *There exists a positive constant C (not depending on F) such that*

$$P\left(\sup_{0\leqslant x<2\pi} |G_n(x)-G(x)|>d\right)\leqslant C\,\mathrm{e}^{-nd^2/2}, \quad d>0,\ n=1,2,\ldots$$

Proof. It is easy to see that

$$\sup\{|G_n(x)-G(x)|: 0\leqslant x<2\pi\}=\sup\{|\{F_n(x^*)-F(x^*)\} - \{F_n(x)-F(x)\}|: 0\leqslant x<2\pi\}.$$

The proof now follows by appealing to the well-known exponential-type probability inequality for the Kolmogorov–Smirnov distance (Dvoretzky *et al.*, 1956). □

Before we prove the next result, we need some more definitions. Assume (A.1)–(A.3) and (A.5). Observe that the function G, defined in (2.1), is differentiable at each $x\in\{z\in[0,2\pi): d(M,z)<\varepsilon\}\cup\{z\in[0,2\pi): d(M^*,z)<\varepsilon\}$ (this ε is the same as the one in (A.3)) with derivative $G'(x)=f(x^*)-f(x)$. We denote $G'(x)$ by $g(x)$. Define a sequence of positive numbers $\{\varepsilon_n: n\geqslant 2\}$ by

$$\varepsilon_n=\frac{8(\log n)^{1/2}}{|g(M)|n^{1/2}}.$$

Define subsets $\{I_n: n\geqslant 2\}$ and $\{J_n: n\geqslant 2\}$ of $[0,2\pi)$ by

$$I_n=\{x\in[0,2\pi): d(x,M)\leqslant\varepsilon_n\}, \qquad J_n=\{x\in[0,2\pi): d(x,M^*)\leqslant\varepsilon_n\}.$$

Lemma 6.2. *Assume* (A.1)–(A.5). *Then with probability* 1,

$$M_n\in I_n \tag{6.1}$$

for all n, sufficiently large.

Proof. We provide the proof corresponding to the case $0<M<\pi$ only. The proofs for the other cases are essentially similar except for a few minor modifications.

Observe that if we assert that with probability 1,

$$M_n \in I_n \cup J_n \tag{6.2}$$

for all n, sufficiently large, then in view of Theorem 4.1 and the definitions of I_n and J_n, (6.1) will follow. It, therefore, suffices to establish (6.2).

Note that in view of the facts about the function G, recorded in Section 2 (in particular Lemma 2.1), we have the following: there exists an integer $N>0$ such that

$$\inf\{|G(x)-\tfrac{1}{2}|:\ x\in(I_n\cup J_n)^c\} = |G(M+\varepsilon_n)-\tfrac{1}{2}| \text{ or } |G(M^*-\varepsilon_n)-\tfrac{1}{2}| \tag{6.3}$$

for all $n\geqslant N$ $[(I_n\cup J_n)^c \equiv [0,2\pi)-(I_n\cup J_n)]$. Again from differentiability of G at both M and M^* it can be seen that there exists an integer $N_1\geqslant N$ such that

$$|G(M+\varepsilon_n)-\tfrac{1}{2}| > \frac{4(\log n)^{1/2}}{n^{1/2}} \quad\text{and}\quad |G(M^*+\varepsilon_n)-\tfrac{1}{2}| > \frac{4(\log n)^{1/2}}{n^{1/2}} \tag{6.4}$$

for all $n\geqslant N_1$. From (6.3) and (6.4) now it follows that

$$\inf\{|G(x)-\tfrac{1}{2}|:\ x\in(I_n\cup J_n)^c\} > \frac{4(\log n)^{1/2}}{n^{1/2}} \tag{6.5}$$

for all $n\geqslant N_1$. Taken with Theorem 4.1, (6.5) now implies that

$$P(M_n\in(I_n\cup J_n)^c) \leqslant P(|G_n(M_n)-G(M_n)| > \frac{2(\log n)^{1/2}}{n^{1/2}} \tag{6.6}$$

for all $n\geqslant N_1$, The assertion (6.2) now follows by employing Lemma 6.1 in (6.6), and then by invoking the Borel–Cantelli lemma. □

Lemma 6.3. *Assume* (A.1)–(A.5). *Then the following expansion for the function* $G(x)$ *holds*:

(a) *if* $0<M<2\pi$,

$$G(x)=G(M)+(x-M)g(M)+\mathrm{O}(|x-M|^{3/2}),\quad |x-M|<\varepsilon,$$

and

(b) *if* $M=0$,

$$G(x)=\begin{cases} G(M)+xg(M)+\mathrm{O}(|x|^{3/2}), & 0\leqslant x\leqslant\varepsilon,\\ G(M)+(x-2\pi)g(M)+\mathrm{O}(|x-2\pi|^{3/2}), & 2\pi-\varepsilon<x<2\pi. \end{cases}$$

Proof. The proof follows immediately from the mean-value theorem and (A.4). □

Let us now recall that the Bahadur representation for the linear median gives indeed a representation for

$$|\xi_{1/2,n}-\xi_{1/2}|\operatorname{sgn}(\xi_{1/2,n}-\xi_{1/2}),$$

where $\xi_{1/2}$ is the population median and $\xi_{1/2,n}$ is the sample median based on a random sample of size n. In the present situation, we shall replace $|\xi_{1/2,n}-\xi_{1/2}|$ by $d(M_n, M)$. As regards the sgn part, we observe that there is no natural order on the circle. Therefore, we shall force one such notion into the picture that serves our purpose. The following definition is made towards this end.

Definition 6.1. We define a sequence of random variables $\{Z_n\colon n\geqslant 1\}$ as follows: for $0\leqslant M<\pi$,

$$Z_n=\begin{cases} d(M_n,M) & \text{if } M<M_n\leqslant M^*,\\ -d(M_n,M) & \text{otherwise,}\end{cases}$$

and for $\pi\leqslant M<2\pi$,

$$Z_n=\begin{cases} -d(M_n,M) & \text{if } M^*<M_n\leqslant M,\\ d(M_n,M) & \text{otherwise.}\end{cases}$$

Lemma 6.4. *Assume* (A.1)–(A.5). *Then with probability* 1,

$$G(M_n)-G(M)=Z_n g(M)+\mathrm{O}(\{d(M_n,M)\}^{3/2}),\quad n\to\infty.$$

Proof. The proof follows from Theorem 5.2 and Lemma 6.3. □

The following result is a slightly modified version of Lemma 1 of Bahadur (1966). A careful scrutiny of the proof of this lemma reveals that it does not make use of the assumption that the underlying distribution function (H, say) is twice differentiable at the quantile under consideration (ξ, say) in its full force, rather it only assumes that H is continuously differentiable on a neighbourhood containing ξ. In view of this observation, the following result follows immediately from this lemma.

Lemma 6.5. *Suppose that H is a distribution function defined over $\mathbb{R}$. Suppose $a\in\mathbb{R}$ is such that H is differentiable on $(a,a+\delta)$ for some $\delta>0$ and differentiable from the right at a. Write $H'(x)=h(x)$, $x\in I=(a,a+\delta)$. Suppose h is continuous on I, and moreover $\lim_{x\to a}h(x)=H'_+(a)$. Let $\{a_n\}$ be a sequence of positive numbers such that*

$$a_n\sim\frac{c_0(\log n)^{1/2}}{n^{1/2}},\quad n\to\infty.$$

Put

$$K_n=\sup\{|[H_n(a+x)-H_n(a)]-[H(a+x)-H(a)]|\colon 0\leqslant x\leqslant a_n\},$$

where H_n is the empirical distribution function corresponding to H. Then with probability 1,

$$K_n=\mathrm{O}(n^{-3/4}(\log n)^{3/4}),\quad n\to\infty.$$

Remark 6.1. If instead of assuming H to be differentiable continuously on $[a, a+\delta)$ we assume H to be differentiable continuously on $(a-\delta, a]$, the resulting assertion (with supremum taken over $-a_n \leqslant x \leqslant 0$) is also true.

Lemma 6.6. *Assume* (A.1)–(A.5). *Define*

$$T_n = \sup\{|[G_n(x) - G_n(M)] - [G(x) - G(M)]|: x \in I_n\}$$

Then with probability 1,

$$T_n = O(n^{-3/4}(\log n)^{3/4}), \quad n \to \infty.$$

Proof. The proof follows by expressing G_n and G in terms of F_n and F, respectively, and then by adapting suitably Lemma 6.5 and Remark 6.1 to this situation. □

Now we prove the main representation theorem of this paper.

Theorem 6.1. *Assume* (A.1)–(A.5). *Then the random variables* Z_n *admit the following Bahadur-type representation*:

$$Z_n = \frac{\frac{1}{2} - G_n(M)}{g(M)} + R_n,$$

where with probability 1,

$$R_n = O(n^{-3/4}(\log n)^{3/4}), \quad n \to \infty,$$

Proof. In view of Lemmas 6.2, 6.4 and 6.6, we have the following: with probability 1,

$$G_n(M_n) - G_n(M) = Z_n g(M) + O(n^{-3/4}(\log n)^{3/4}), \quad n \to \infty. \tag{6.6}$$

The proof now follows from (6.6) and Theorem 4.1. □

The following corollary is an immediate consequence of the above theorem.

Corollary 6.1. *As n tends to* ∞, $\sqrt{n} Z_b$ *converges weakly to a normal distribution with mean* 0 *and variance* $1/4g^2(M)$.

Remark 6.2. It is easy to see that the Euclidean metric on $\mathbb{R}^2$, restricted to S^1, induces the metric d_1 on $[0, 2\pi)$ given by $d_1(\theta_1, \theta_2) = 2\sin(d(\theta_1, \theta_2)/2)$, where d is the metric defined in Section 3. It now seems natural to ask whether we still have assertions like Theorems 5.1, 6.1 and Corollary 6.1, if we redefine Z_n by replacing d by d_1. It is not difficult to see from Theorems 5.1, 6.1, Corollary 6.1 and the fact $|\sin x - x| = O(|x|^3)$, $x \to 0$, that the answer is in the affirmative.

Remark 6.3. This remark is made to discuss applicability of Theorem 5.1 and Corollary 6.1 for a particular distribution.

It turns out that most of the standard circular distributions are symmetric and unimodal around a fixed point on the circle. This guarantees existence of a unique population c-median (see Mardia, 1972, pp. 46, 47). Thus assumptions (A.1)–(A.3) and (A.5) are satisfied by most of the circular distributions. As regards assumption (A.4), it really states that f is Lipschitz continuous of order $\frac{1}{2}$ at both θ_0 and θ_0^* (with Lipschitz continuity being defined suitably when $\theta_0 = 0$). A sufficient condition that implies (A.4) is the following:

(A.4)′ f is differentiable from both left and right at θ_0 and θ_0^* as well if $0 < \theta_0 < \pi$. If, however, $\theta_0 = \pi$, f is differentiable from right at 0, from left at 2π, and from both left and right at π.

Note that it is easier to check (A.4)′ mathematically than to check (A.4).

Remark 6.4. In Ducharme and Milasevic (1987) (henceforth abbreviated to DM), a result similar to Corollary 6.1 can be found. However, it appears that there are some mistakes and gaps in their approach. In what follows, we present a brief discussion of this modestly.

DM restrict their attention to symmetric and unimodal circular distributions, and assume without loss of generality that the population c-median M (μ in the notation of DM) lies in $[0, \pi)$. However, it is a mistake to take the sample c-median M_n (μ_n, in the notation of DM) also in $[0, \pi)$ since such a strategy does not take care of the relation $f(M) > f(M^*)$ and uses prior knowledge about the population c-median. The difficulty becomes apparent when $M = 0$. In this situation, with high probability one would get samples with M_n close to 2π; whereas an appeal to the definition of sample c-median of DM would indeed imply that M_n is close to π for such samples. Consequently, the claim of DM that strong consistency of sample c-median can be established along the lines of Pollard (1984, p. 7) is not true. A similar difficulty is encountered also in Section 2 of DM.

Remark 6.5. In view of Ghosh's (1971) simpler proof of a weaker version of Bahadur's result, it will be of some interest to see whether it is possible to obtain a similar weaker version of Theorem 6.1.

Remark 6.6. In view of Mardia's (1972, p. 33) extension of the notion of circular median to quantiles of circular random variables, it will be of some interest to see whether it is possible to prove a result analogous to Theorem 6.1 for such quantiles.

Acknowledgements

The present work is based on part of author's dissertation written under the supervision of Professor J.K. Ghosh at the Indian Statistical Institute. The author wishes to record his deep sense of gratitude to Professor Ghosh for his guidance and encouragement while pursuing this work. Thanks are due also to Professor S.N. Joshi for several helpful discussions. The author is grateful also to the Department of Statistics, University of Toronto, Canada and the Natural Sciences and Engineering Research Council of Canada for providing him the facilities for writing the earlier version of this paper. Finally, the author thanks an anonymous referee for a very helpful report that led to improvement over the earlier version of the paper.

References

Bahadur, R.R. (1966). A note on quantiles in large samples. *Ann. Math. Statist.* **37**, 577–580.

Cabrera, J. and G.S. Watson (1990). On a spherical median related distribution. *Comm. Statist. Theory Methods* **19**, 1973–1986.

de Haan, L. and E. Taconis-Haantjes (1979). On Bahadur's representations of sample quantiles. *Ann. Inst. Statist. Math.* **31**, Pt. A, 299–308.

Ducharme, G. and P. Milasevic (1987). Some asymptotic properties of the circular median. *Comm. Statist. Theory Methods* **16**, 659–664.

Dutta, K. and P.K. Sen (1971). On the Bahadur representation of sample quantiles in some stationary multivariate autoregressive processes. *J. Multivariate Anal.* **1**, 186–198.

Dvoretzky, A., J. Kiefer and J. Wolfowitz (1956). Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *Ann. Math. Statist.* **27**, 642–669.

Fisher, N.I. (1985). Spherical medians. *J. Roy Statist. Soc. Ser. B* **47**, 342–348.

Ghosh, J.K. (1971). A new proof of the Bahadur representation of quantiles and an application. *Ann. Math. Statist.* **42**, 1957–1961.

Kiefer, J. (1967). On Bahadur's representation of sample quantiles. *Ann. Math. Statist.* **38**, 1323–1342.

Ko, D, and P. Guttorp (1988). Robustness of estimators for directional data. *Ann. Statist.* **16**, 609–618.

Lenth, R.V. (1981). Robust measures of location of directional data. *Technometrics* **23**, 77–81.

Liu, R. and K. Singh (1992). Ordering directional data: concepts of data depth on circles and spheres. *Ann. Statist.* **20**, 1468–1484.

Mardia, K.V. (1972). *Statistics of Directional Data.* Academic Press, London.

Pollard, D. (1984). *Convergence of Stochastic Processes.* Springer, New York.

Purkayastha, S. (1991a). A rotationally symmetric directional distribution: obtained through maximum likelihood characterization. *Sankhya A* **53**, 70–83.

Purkayastha, S. (1991b). Some contributions to directional statistics: characterization and asymptotics. Ph.D. thesis, Indian Statistical Institute, Calcutta (unpublished).

Sen, P.K. (1968). Asymptotic normality of sample quantiles for m-dependent processes. *Ann. Math. Statist.* **39**, 1724–1730.

Sen, P.K. (1972). On the Bahadur representation of sample quantiles for sequences of ϕ-mixing random variables. *J. Multivariate Anal.* **2**, 77–95.

Sen, P.K. and M. Ghosh (1971). On bounded length sequential confidence intervals based on one-sample rank-order statistics. *Ann. Math. Statist.* **42**, 189–203.

Small, C. (1990). A survey of multidimensional medians, *Internat. Statist. Rev.* **58**, 263–277.

Sverdrup-Thygeson, H. (1981). Strong law of large numbers for measures of central tendency and dispersion of random variables on compact metric spaces. *Ann. Statist.* **6**, 141–145.

Wehrly, T.E. and E.P. Shine (1981). Influence curves of estimators for directional data. *Biometrika* **68**, 334–335.