Available online at www.sciencedirect.com
SCIENCE DIRECT®

ELSEVIER

Expo. Math. 23 (2005) 281 – 287

EXPOSITIONES MATHEMATICAE

www.elsevier.de/exmath

Spherical distributions: Schoenberg (1938) revisited

A.G.M. Steerneman, F. van Perlo-ten Kleij*

Department of Econometrics, University of Groningen, Landleven 5, P.O. Box 800, 9700 AV Groningen, Netherlands

Received 24 May 2004; received in revised form 26 August 2004

Abstract

An m-dimensional random vector X is said to have a spherical distribution if and only if its characteristic function is of the form $\phi(\|t\|)$, where $t \in \mathbb{R}^m$, $\|.\|$ denotes the usual Euclidean norm, and ϕ is a characteristic function on $\mathbb{R}$. A more intuitive description is that the probability density function of X is constant on spheres. The class Φ_m of these characteristic functions ϕ is fundamental in the theory of spherical distributions on $\mathbb{R}^m$. An important result, which was originally proved by Schoenberg (Ann. Math. 39(4) (1938) 811–841), is that the underlying characteristic function ϕ of a spherically distributed random m-vector X belongs to Φ_∞ if and only if the distribution of X is a scale mixture of normal distributions. A proof in the context of exchangeability has been given by Kingman (Biometrika 59 (1972) 492–494). Using probabilistic tools, we will give an alternative proof in the spirit of Schoenberg we think is more elegant and less complicated.

MSC 2000: primary 60E05; 60E10; 62E10; 62H05; secondary 42A61; 60F05; 60F15

Keywords: Spherical distribution; Characteristic function; Normal distribution

1. Introduction

The class of spherical probability distributions can be defined in a number of equivalent ways (see, e.g., [4,6,7]). We use a definition in terms of characteristic functions, that is, the

* Corresponding author. Tel.: +31 50 3633671.

E-mail addresses: a.g.m.steerneman@eco.rug.nl (A.G.M. Steerneman), f.van.perlo-ten.kleij@eco.rug.nl (F. van Perlo-ten Kleij).

0723-0869/$ - see front matter © 2005 Elsevier GmbH. All rights reserved.
doi:10.1016/j.exmath.2005.01.002

random m-vector X is said to have a spherical distribution if and only if the corresponding characteristic function $\psi(t)$, $t \in \mathbb{R}^m$, is a function of $\|t\|$ only. The class of all characteristic functions ϕ on $\mathbb{R}$ such that $\phi(\|t\|)$ is a characteristic function is denoted by Φ_m. If $\phi \in \Phi_k$ for all k, we say that $\phi \in \Phi_\infty$. Notice that $\Phi_1 \supset \Phi_2 \supset \cdots \supset \Phi_\infty$.

In the theory of spherical distributions, a key role is played by the random vector U that is uniformly distributed on the unit sphere S^{m-1} in $\mathbb{R}^m$. We denote this by $U \sim \mathscr{U}(S^{m-1})$. Let $\mathscr{O}(m)$ denote the group of orthogonal $m \times m$ matrices. Because the distribution of U is invariant under orthogonal transformations, we know that $E\mathrm{e}^{\mathrm{i}t'U} = E\mathrm{e}^{\mathrm{i}t'CU}$ for all $C \in \mathscr{O}(m)$, $t \in \mathbb{R}^m$. Therefore, this characteristic function is invariant under the action of the group of all orthogonal $m \times m$ matrices, and hence it depends on t via the function $\|t\|$, where $\|\cdot\|$ denotes the usual Euclidean norm. So, we can write $E\mathrm{e}^{\mathrm{i}t'U} = \Omega_m(\|t\|)$. Let ω_m denote the area of S^{m-1} and $\mathrm{d}\omega_m$ its surface element, then

$$\begin{aligned}\Omega_m(\|t\|) &= \frac{1}{\omega_m}\int_{S^{m-1}} \mathrm{e}^{\mathrm{i}t'u}\,\mathrm{d}\omega_m(u)\\ &= \frac{1}{\omega_m}\int_{S^{m-1}} \mathrm{e}^{\mathrm{i}\|t\|u_m}\,\mathrm{d}\omega_m(u),\end{aligned}$$

where we have used the fact that $\Omega_m(\|t\|) = \Omega_m(\|C't\|)$ for all $C \in \mathscr{O}(m)$ and in particular for $C = (c_1, c_2, \ldots, c_{m-1}, \|t\|^{-1}t)$, where the columns of this matrix are chosen to be mutually orthogonal and have unit length. Hence, it can be written

$$\Omega_m(\|t\|) = \frac{1}{\omega_m} f_m(\|t\|), \tag{1}$$

where

$$f_m(r) = \int_{S^{m-1}} \mathrm{e}^{\mathrm{i}ru_m}\,\mathrm{d}\omega_m(u)$$

and

$$\omega_m = f_m(0).$$

It is well-known that

$$\mathrm{d}\omega_m = (1-t^2)^{(m-3)/2}\,\mathrm{d}t\,\mathrm{d}\omega_{m-1}$$

and

$$\omega_m = f_m(0) = \frac{2\pi^{m/2}}{\Gamma(m/2)},$$

see, e.g., [10]. It hence follows from (1) that

$$\Omega_m(r) = \frac{\Gamma(m/2)}{\sqrt{\pi}\,\Gamma((m-1)/2)}\int_{-1}^{1} \mathrm{e}^{\mathrm{i}rt}(1-t^2)^{(m-3)/2}\,\mathrm{d}t. \tag{2}$$

Schoenberg [11] gave the following representation for characteristic functions in Φ_m. We give a short proof.

Theorem 1 (*Schoenberg [11]*). *The random m-vector X has a spherical distribution with characteristic function* $\psi(t) = \phi(\|t\|)$ *if and only if*

$$\phi(\|t\|) = \int_0^\infty \Omega_m(\|t\| y)\, \mathrm{d}Q(y) \tag{3}$$

for some probability measure Q on $[0, \infty)$ (*Q is in fact the probability distribution of* $\|X\|$).

Proof. Assume X has a spherical distribution with characteristic function $\psi(t) = \phi(\|t\|)$. Let U be uniformly distributed on the unit sphere S^m, independent of X. First note that

$$\begin{aligned}\phi(\|t\|) = \psi(t) = E\psi(\|t\|U) &= E\, E\mathrm{e}^{\mathrm{i}\|t\|U'X} = E\Omega_m(\|t\|\|X\|)\\ &= \int_0^\infty \Omega_m(\|t\|y)\, \mathrm{d}Q(y),\end{aligned}$$

where we have used the fact that $\psi(t)$ is constant on unit spheres, that is $\psi(t) = \psi(\|t\|u)$ for all u with $\|u\| = 1$.

Conversely, the right-hand side of (3) is the characteristic function of a spherical distribution on $\mathbb{R}^m$. This is easily seen if we define $X = RU$ where $R \geqslant 0$ has probability distribution Q, and $U \sim \mathscr{U}(S^{m-1})$ independent of R, because

$$\int_0^\infty \Omega_m(\|t\|y)\, \mathrm{d}Q(y) = E\Omega_m(\|t\|R) = E\, E\mathrm{e}^{\mathrm{i}Rt'U} = E\mathrm{e}^{\mathrm{i}t'(RU)},$$

which is the characteristic function of $X = RU$. Its distribution is obviously invariant under orthogonal transformations. □

The class Φ_∞ is the topic of the next section. We will give an alternative proof of the well-known result that a characteristic function ϕ belongs to Φ_∞ if and only if it is the characteristic function corresponding to a scale mixture of normal distributions. An important element in this proof is the behaviour of $\Omega_m(r\sqrt{m})$ as $m \to \infty$. Pointwise convergence of $\Omega_m(r\sqrt{m})$ has been pointed out to Schoenberg by J. von Neumann (see [11, footnote 12]):

$$\lim_{m\to\infty} \Omega_m(r\sqrt{m}) = \mathrm{e}^{-r^2/2}; \tag{4}$$

Hartman and Wintner [8] attributed this result to Laplace, see their footnote 13. However, global uniform convergence of $\Omega_m(r\sqrt{m})$ is needed, that is, (4) must hold uniformly for all real values of r. This global uniform convergence is a key element in the proof, yet it is not easy to establish. Schoenberg's proof is quite complex and is organized in the form of three lemmas [11, Lemmas 1–3]. We will prove global uniform convergence of $\Omega_m(r\sqrt{m})$ by using probabilistic tools, which makes the proof, as we think, shorter, more elegant, and less complicated.

2. The class Φ_∞

More formally, we focus on the following theorem.

Theorem 2 (*Schoenberg [11]*). *The elements $\phi : \mathbb{R} \to [-1, 1]$ of Φ_∞ can be represented as*

$$\phi(t) = \int_0^\infty \mathrm{e}^{-t^2 y^2/2} \, \mathrm{d}Q(y), \tag{5}$$

where Q is a probability measure on $[0, \infty)$.

A proof in the context of exchangeability has been given by Kingman [9] and a slightly adapted version of this proof can be found in Fang et al. [7]. The proof we will give uses the basic ideas of Schoenberg [11], however, the crucial step of global uniform convergence of $\Omega_m(r\sqrt{m})$ is proved by applying more modern probabilistic tools. While it took Schoenberg quite some effort to establish global uniform convergence of $\Omega_m(r\sqrt{m})$, our argument is relatively short and, as we think, more transparent.

Donoghue [5, pp. 201–206] already presented a simplified proof of the required global convergence, but it is still rather complicated and technical. Chapter 5 of Berg et al. [2] is dedicated to an abstract form of Schoenberg's theorem and generalizations. They used the concept of a Schoenberg triple. Their approach also leads to a simplification.

Proof. First, we will show that $\phi \in \Phi_m$ for all m. Let Y be a random variable that is distributed according to the probability measure Q, and let $X_1, \ldots, X_m$ be random variables that are conditionally independent given Y, each with conditional distribution given Y that is $\mathcal{N}(0, Y^2)$, $i = 1, \ldots, m$. Define $X = (X_1, \ldots, X_m)'$, then X is spherically distributed on $\mathbb{R}^m$ and for its characteristic function we obtain, letting $t \in \mathbb{R}^m$,

$$\begin{aligned} E\mathrm{e}^{\mathrm{i}t'X} &= E\, E(\mathrm{e}^{\mathrm{i}t'X} | Y) \\ &= E\mathrm{e}^{-\|t\|^2 Y^2/2} = \int_0^\infty \mathrm{e}^{-\|t\|^2 y^2/2} \, \mathrm{d}Q(y) \\ &= \phi(\|t\|). \end{aligned}$$

Hence, $\phi \in \Phi_m$ for all m. In terms of integrals the derivation above reads as follows:

$$E\mathrm{e}^{\mathrm{i}t'X} \equiv \int_{-\infty}^\infty \cdots \int_{-\infty}^\infty \exp\left(\mathrm{i} \sum_{j=1}^m t_j X_j\right) g(x_1, \ldots, x_m) \, \mathrm{d}x_1 \cdots \mathrm{d}x_m.$$

The probability density function g is given by

$$g(x_1, \ldots, x_m) = \int_0^\infty \prod_{j=1}^m h(x_j | y) \, \mathrm{d}Q(y),$$

where

$$h(x|y) = (2\pi y^2)^{-\frac{1}{2}} \exp\left(-\frac{x^2}{2y^2}\right).$$

Now we use

$$\int_{-\infty}^\infty \exp\left(\mathrm{i}t_j x_j\right) h(x_j | y) \, \mathrm{d}x_j = \exp\left(-\frac{1}{2} t_j^2 y^2\right)$$

for $j = 1, \ldots, m$. By Fubini's theorem we now obtain

$$Ee^{it'X} = \int_0^\infty \exp\left(-\frac{1}{2}\|t\|^2 y^2\right) \mathrm{d}Q(y).$$

Second, we have to prove that we can find such a representation for any $\phi \in \Phi_\infty$. Suppose $\phi \in \Phi_m$ for all m, then we can write

$$\phi(t) = \int_0^\infty \Omega_m(ty\sqrt{m})\,\mathrm{d}F_m(y) \tag{6}$$

for some probability distribution function F_m on $[0, \infty)$. If we let $m \to \infty$ in (6), it is tempting to exploit (4) to arrive at the representation (5). However, to apply Helly's well-known convergence theorem for distribution functions (see, e.g., [1, Theorem 8.2.1]), pointwise convergence of $\Omega_m(r\sqrt{m})$ is not sufficient. Because the interval of integration in (6) is infinite, we need global uniform convergence of $\Omega_m(r\sqrt{m})$. As we already stated in section 1, Schoenberg [11] proved that $\Omega_m(r\sqrt{m}) \to e^{-r^2/2}$ uniformly on $\mathbb{R}$ as $m \to \infty$. We shall give an alternative proof for this fact.

By making the transformation $t = x/\sqrt{m}$ in (2), it is easy to see that $\Omega_m(r\sqrt{m})$, as a function of r, is the characteristic function corresponding to the probability density function

$$q_m(x) = \frac{\Gamma(m/2)}{\sqrt{\pi m}\,\Gamma((m-1)/2)}\left(1 - \frac{x^2}{m}\right)^{(m-3)/2} I_{(-\sqrt{m},\sqrt{m})}(x).$$

On account of Stirling's formula we know that

$$\lim_{m\to\infty} \frac{\Gamma(m/2)}{\Gamma((m-1)/2)\sqrt{m/2}} = 1.$$

Moreover,

$$\lim_{m\to\infty}\left(1 - \frac{x^2}{m}\right)^{(m-3)/2} = e^{-x^2/2} \quad \text{for all } x.$$

Therefore, $q_m(x) \to q_\infty(x) = (2\pi)^{-1/2}e^{-x^2/2}$ for all x, as $m \to \infty$, where we immediately recognize the standard normal distribution. We now want to apply Scheffé's lemma.

Lemma 1 (*Scheffe's lemma*). *Let λ be a measure* (*not necessarily finite*) *on a space* $(\Omega, \mathscr{B})$ *and let p and p_n be probability densities w.r.t. λ. If $p_n \to p$ a.e.* $[\lambda]$, *then*

$$\sup_{E\in\mathscr{B}}\left|\int_E p\,\mathrm{d}\lambda - \int_E p_n\,\mathrm{d}\lambda\right| = \frac{1}{2}\int |p - p_n|\,\mathrm{d}\lambda \to 0.$$

For a proof of this lemma see, for example, Billingsley [3, p. 224]. It follows from Lemma 1 that

$$\lim_{m\to\infty}\int |q_m(x) - q_\infty(x)|\,\mathrm{d}x = 0.$$

If we define $\Omega_\infty(r) = \mathrm{e}^{-r^2/2}$, the characteristic function corresponding to $q_\infty(x)$, then we have

$$\begin{aligned} |\Omega_m(r\sqrt{m}) - \Omega_\infty(r)| &= \left| \int \mathrm{e}^{\mathrm{i}rx} q_m(x)\,\mathrm{d}x - \int \mathrm{e}^{\mathrm{i}rx} q_\infty(x)\,\mathrm{d}x \right| \\ &\leqslant \int |q_m(x) - q_\infty(x)|\;\mathrm{d}x. \end{aligned} \tag{7}$$

It now follows that $\Omega_m(r\sqrt{m}) \to \Omega_\infty(r)$ uniformly in r. Following the same reasoning as Schoenberg [11], we have, according to Helly's theorem, that there exists a subsequence F_{m_k} of F_m such that $F_{m_k} \to F$ weakly, where F is a probability distribution function. Now we write

$$\phi(t) = \int_0^\infty (\Omega_m(ty\sqrt{m}) - \mathrm{e}^{-t^2y^2/2})\,\mathrm{d}F_m(y) + \int_0^\infty \mathrm{e}^{-t^2y^2/2}\,\mathrm{d}F_m(y).$$

As $m \to \infty$, we obtain from (7) that

$$\left| \int_0^\infty (\Omega_m(ty\sqrt{m}) - \mathrm{e}^{-t^2y^2/2})\,\mathrm{d}F_m(y) \right| \leqslant \int_{-\infty}^\infty |q_m(x) - q_\infty(x)|\;\mathrm{d}x \to 0,$$

and, from $F_{m_k} \to F$ weakly, that

$$\int_0^\infty \mathrm{e}^{-t^2y^2/2}\,\mathrm{d}F_{m_k}(y) \to \int_0^\infty \mathrm{e}^{-t^2y^2/2}\,\mathrm{d}F(y).$$

Putting things together, we find

$$\phi(t) = \int_0^\infty \mathrm{e}^{-t^2y^2/2}\,\mathrm{d}F(y). \qquad \square$$

Acknowledgements

The authors would like to thank an anonymous referee for suggestions for improvement of the presentation and for the Refs. [2,5].

References

[1] R.B. Ash, Real Analysis and Probability, Academic Press, New York, 1972.
[2] C. Berg, J.P.R. Christensen, P. Ressel, Harmonic Analysis on Semigroups. Theory of Positive Definite and Related Functions, Graduate Texts in Mathematics, vol. 100, Springer, Berlin, Heidelberg, New York.
[3] P. Billingsley, Convergence of Probability Measures, Wiley, New York, 1968.
[4] S. Cambanis, S. Huang, G. Simons, On the theory of elliptically contoured distributions, J. Multivariate Anal. 11 (1981) 368–385.
[5] W.F. Donoghue Jr., Distributions and Fourier Transforms, Academic Press, New York, 1969.
[6] M.L. Eaton, A characterization of spherical distributions, J. Multivariate Anal. 20 (1986) 272–276.
[7] K.T. Fang, S. Kotz, K.W. Ng, Symmetric Multivariate and Related Distributions, Chapman & Hall, London, 1990.
[8] P.H. Hartman, A. Wintner, On the spherical approach to the normal distribution law, Amer. J. Math. 62 (1940) 759–779.

[9] J.F.C. Kingman, On random sequences with spherical symmetry, Biometrika 59 (1972) 492–494.
[10] C. Müller, Spherical Harmonics, Lecture Notes in Mathematics, vol. 17, Springer, Berlin, Heidelberg, New York, 1966.
[11] I.J. Schoenberg, Metric spaces and completely monotone functions, Ann. Math. 39 (4) (1938) 811–841.