Statistics and Probability Letters 97 (2015) 116–119

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

# A generalization of an integral arising in the theory of distance correlation

Johannes Dueck [a], Dominic Edelmann [a], Donald Richards [b,*]

[a] Institut für Angewandte Mathematik, Universität Heidelberg, Im Neuenheimer Feld 294, 69120 Heidelberg, Germany

[b] Department of Statistics, Pennsylvania State University, University Park, PA 16802, USA

ARTICLE INFO

*Article history:*
Received 5 November 2014
Received in revised form 13 November 2014
Accepted 13 November 2014
Available online 21 November 2014

*MSC:*
primary 62H20
secondary 32A55
30E20

*Keywords:*
Analytic continuation
Distance correlation
Generalized functions
Measures of dependence
Meromorphic function
Singular integral

ABSTRACT

We generalize an integral which arises in several areas in probability and statistics and which is at the core of the field of distance correlation, a concept developed by Székely et al. (2007) to measure dependence between random variables.



## 1. Introduction

Székely et al. (2007) and Székely and Rizzo (2009) introduced the powerful concepts of distance covariance and distance correlation as measures of dependence between collections of random variables. In later papers, Rizzo and Székely (2010, 2011) and Székely and Rizzo (2012, 2013, 2014) gave applications of the distance correlation concept to several problems in mathematical statistics. In recent years, there have appeared an enormous number of papers in which the distance correlation coefficient has been applied to many fields. In particular, the concept of distance covariance has been extended to abstract metric spaces (Lyons, 2013) and has been related to machine learning (Sejdinovic et al., 2013); and there have been applications to the analysis of wind data (Dueck et al., 2014); to detecting associations in large astrophysical databases (Martínez-Gómez et al., 2014) and to interpreting those associations (Richards et al., 2014); to measuring nonlinear dependence in time series data (Zhou, 2012); and to numerous other fields.

Székely and Rizzo (2005), in developing the foundations of distance correlation, derived an intriguing multidimensional singular integral. It is this integral which is the subject of the present paper; and in stating the result, we denote for $t, x \in \mathbb{R}^d$ the standard inner product and Euclidean norm by $\langle t, x \rangle$ and $\|t\|$, respectively.

* Corresponding author.
*E-mail address:* richards@stat.psu.edu (D. Richards).

http://dx.doi.org/10.1016/j.spl.2014.11.012
0167-7152/

Suppose that $\alpha \in \mathbb{C}$ satisfies $0 < \Re(\alpha) < 2$. Székely and Rizzo (2005) proved that, for all $x \in \mathbb{R}^d$,

$$\int_{\mathbb{R}^d} \frac{1-\cos(\langle t, x\rangle)}{\|t\|^{d+\alpha}}\, \mathrm{d}t = C(d,\alpha)\, \|x\|^{\alpha}, \tag{1.1}$$

where

$$C(d,\alpha) = \frac{2\pi^{d/2}\, \Gamma(1-\alpha/2)}{\alpha\, 2^{\alpha}\, \Gamma\big((d+\alpha)/2\big)}. \tag{1.2}$$

Székely and Rizzo defined the integral (1.1) by means of a regularization procedure, where the integrals at 0 and at $\infty$ are in a principal value sense: $\lim_{\epsilon\to 0}\int_{\mathbb{R}^d\setminus\{\epsilon B+\epsilon^{-1}B^c\}}$, where $B$ is the unit ball centered at the origin in $\mathbb{R}^d$ and $B^c$ is the complement of $B$.

In this paper, we generalize the integral (1.1) by inserting into the integrand a truncated Maclaurin expansion of the function $\cos(\langle t, x\rangle)$. We show that the generalization is valid for all $\alpha \in \mathbb{C}$ such that $2(m-1) < \Re(\alpha) < 2m$, where $m$ is any positive integer. Moreover, we prove that the generalization converges absolutely under the stated condition on $\alpha$; as a consequence, we deduce that (1.1) converges without the need for regularization.

We note that the integral (1.1) arises in other areas of probability and statistics. Indeed, in the area of generalized random fields, (1.1) provides the spectral measure of a power law generalized covariance function, which corresponds to fractional Brownian motion; see Reed et al. (1995) or Chilès and Delfiner (2012, p. 266, Section 4.5.6). In mathematical analysis, a related integral is treated by Gelfand and Shilov (1964, pp. 192–195), and a similar singular integral arises in Fourier analysis in the derivation of the norms of integral operators between certain Sobolev spaces of functions (Stein, 1970, pp. 140 and 263).

We remark that the extension of (1.1) to more general values of $\alpha$ raises the intriguing possibility that a general theory of distance correlation can be developed for values of $\alpha$ outside the range $(0, 2)$.

## 2. The main result

Let $m \in \mathbb{N}$, the set of positive integers. Also, for $v \in \mathbb{R}$, define

$$\cos_m(v) := \sum_{j=0}^{m-1} (-1)^j \frac{v^{2j}}{(2j)!} \tag{2.1}$$

to be the truncated Maclaurin expansion of the cosine function, where the expansion is halted at the $m$th summand.

The following result generalizes (1.1) to arbitrary $m \in \mathbb{N}$.

**Theorem.** *Let $m \in \mathbb{N}$ and $x \in \mathbb{R}^d$. For $\alpha \in \mathbb{C}$,*

$$\int_{\mathbb{R}^d} \frac{\cos_m(\langle t, x\rangle) - \cos(\langle t, x\rangle)}{\|t\|^{d+\alpha}}\, \mathrm{d}t = C(d,\alpha)\, \|x\|^{\alpha}, \tag{2.2}$$

*with absolute convergence if and only if $2(m-1) < \Re(\alpha) < 2m$, where $C(d,\alpha)$ is given in (1.2).*

**Proof.** We shall establish the proof by induction on $m$.

Throughout the proof, we let $B_a = \{x \in \mathbb{R}^d : \|x\| < a\}$ denote the ball which is centered at the origin and which is of radius $a$.

Consider the case in which $m = 1$. In this case, observe that for $t \in B_a$ where $a$ is sufficiently small, the function

$$t \mapsto \cos_1(\langle t, x\rangle) - \cos(\langle t, x\rangle) \equiv 1 - \cos(\langle t, x\rangle)$$

is asymptotic to $\|t\|^2$. Then the integrand in (2.2), when restricted to $B_a$, is asymptotic to $\|t\|^{-d-\alpha+2}$. By a transformation to spherical coordinates to compute the integral over the unit ball $B$ we deduce that the integrand is integrable over $B_a$, and hence integrable over any compact neighborhood of the origin, if and only if $\Re(\alpha) < 2$.

For $\|t\| \to \infty$, we apply the bound $|1 - \cos(\langle t, x\rangle)| \le 2$ to deduce that the integrand in (2.2) (with $m = 1$) is integrable over $\mathbb{R}\setminus B_a$ if and only if $\Re(\alpha) > 0$. Consequently, for $m = 1$, the integral converges for all $x \in \mathbb{R}^d$ if and only if $0 < \Re(\alpha) < 2$.

To conclude the proof for the case in which $m = 1$, we proceed precisely as did Székely et al. (2007, p. 2771) to obtain the right-hand side of (2.2).

Next, we assume by inductive hypothesis that the assertion holds for a given positive integer $m$. Note that the right-hand side of (2.2), as a function of $\alpha \in \mathbb{C}$, is meromorphic with a pole at each nonnegative integral $\alpha$.

By (2.1),

$$\cos_{m+1}(v) = \cos_m(v) + (-1)^m \frac{v^{2m}}{(2m)!}.$$