Z. Wahrscheinlichkeitstheorie verw. Gebiete
57, 407–417 (1981)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

A Probabilistic Characterisation of Negative Definite and Completely Alternating Functions

Jens Peter Reus Christensen[1] and Paul Ressel[2]

[1] Københavns Universitets Matematiske Institut, Universitetsparken, DK-2100 København, Denmark
[2] Institut für Mathematische Statistik der Westfälischen Wilhelms-Universität, Roxeler Straße 64, D-4400 Münster, Federal Republic of Germany

0. Introduction

In 1956 Hoeffding proved the following result ([5], Theorem 3): if $X_1, \dots, X_n$ are independent Bernoulli random variables with parameters $p_1, \dots, p_n$, and if g: $\{0, 1, \dots, n\} \to \mathbb{R}$ is strictly concave, then

$$E_{\bar{p}}\left[g\left(\sum_{i=1}^n X_i\right)\right] \leqq E_{p_1, \dots, p_n}\left[g\left(\sum_{i=1}^n X_i\right)\right]$$

where $\bar{p} = \frac{1}{n} \sum_{i=1}^n p_i$ and on the left hand side it is assumed that $P(X_i = 1) = \bar{p}$ for all $i = 1, \dots, n$.

Recently, Bickel and van Zwet found a considerable extension of this result: let g: $\mathbb{R}^m \to \mathbb{R}$ and $k \in \mathbb{N}$ be given and let $X_1, \dots, X_n$ be independent real valued random variables with distributions $\mu_1, \dots, \mu_n$ and the extra condition that there is a finite subset $A \subseteq \mathbb{R}^m$ with cardinality k such that $\mu_j(A) = 1$ for all $j = 1, \dots, n$. Then the following three conditions are equivalent: (1) the inequality

$$(*) \qquad E_{\bar{\mu}}\left[g\left(\sum_{i=1}^n X_i\right)\right] \leqq E_{\mu_1, \dots, \mu_n}\left[g\left(\sum_{i=1}^n X_i\right)\right]$$

holds for all n and all such $\mu_1, \dots, \mu_n$; (2) the inequality $(*)$ holds for $n = 2$ and all such μ_1, μ_2; (3) the $k \times k$-matrix $(g(x_i + x_j))_{i,j=1,\dots,k}$ is negative definite $\left(\text{i.e.}\ \sum_{i,j=1}^k c_i c_j g(x_i + x_j) \leqq 0 \text{ for all } (c_1, \dots, c_k) \in \mathbb{R}^k \text{ with } \sum_{i=1}^k c_i = 0\right)$ for every choice of $x_1, \dots, x_k \in \mathbb{R}^m$; see [3].

Condition (3) is closely related to the notion of "negative definite functions" on abelian semigroups which has been investigated in [2], and it was tempting to find out the relationship between this class of functions on arbitrary abelian

0044-3719/81/0057/0407/$02.20

semigroups and inequalities "of type (*)". Following this idea we have obtained a complete characterization of negative definite functions (Theorem 1 below). For the important subclass of completely alternating functions a similar question arises, however this problem turned out to be much more complicated and remained open for about one year. Quite recently however we found out that our original guess was correct, providing for completely alternating functions, too, a new characterisation of probabilistic nature (Theorem 2 below).

1. The Discrete Case

Let $(S,+,0)$ be an abelian semigroup with neutral element 0, let $(\Omega,\mathscr{F},P)$ be a probability space and denote by $X_1, X_2, \ldots$ some S-valued random variables on $(\Omega,\mathscr{F},P)$. In this first part we do not impose some measurable or topological structure on S; we therefore say that $X\colon \Omega\to S$ is a (simple) random variable iff $\{X=s\}\in\mathscr{F}\ \forall s\in S$ and $\{X=s\}=\emptyset$ for all but finitely many $s\in S$. Any partial sum $S_n=X_1+\ldots+X_n$ then is also a random variable. The distribution μ_j of X_j is of course a finite convex combination of one-point measures on S. If g is an arbitrary real valued function on S and X is a (simple) S-valued random variable then $g\circ X$ again is a simple real valued random variable. Let us recall the following

Definition (cf. [2]). A function $g\colon S\to\mathbb{R}$ is called

(i) *positive definite* iff $\sup_{s\in S}|g(s)|<\infty$ and for every $n\geqq 1$ and every n-tuple $(s_1,\ldots,s_n)\in S^n$ the $n\times n$-matrix $(g(s_i+s_j))_{i,j=1,\ldots,n}$ is positive semidefinite,

(ii) *negative definite* iff $g\geqq 0$ and for every $n\geqq 2$, every n-tuple $(s_1,\ldots,s_n)\in S^n$ and every n-tuple $(c_1,\ldots,c_n)\in\mathbb{R}$ such that $c_1+\ldots+c_n=0$ we have

$$\sum_{i,j=1}^{n} c_i c_j g(s_i+s_j)\geqq 0,$$

(iii) *completely monotone* iff $g\geqq 0$ and for every $n\geqq 1$, every n-tuple $(a_1,\ldots,a_n)\in S^n$ and every $s\in S$ we have

$$\nabla_n g(s;a_1,\ldots,a_n)\geqq 0,$$

where

$$\nabla_1 g(s;a_1):=g(s)-g(s+a_1)$$

and

$$\nabla_k g(s;a_1,\ldots,a_k)=\nabla_{k-1} g(s;a_1,\ldots,a_{k-1}) - \nabla_{k-1} g(s+a_k;a_1,\ldots,a_{k-1}) \quad \text{for } k\geqq 2,$$

(iv) *completely alternating* iff $g\geqq 0$ and for every $n\geqq 1$, every n-tuple $(a_1,\ldots,a_n)\in S^n$ and every $s\in S$ we have

$$\nabla_n g(s;a_1,\ldots,a_n)\leqq 0.$$

The set of all functions fulfilling (i)–(iv) is denoted resp. $\mathscr{P}(S)$, $\mathscr{N}(S)$, $\mathscr{M}(S)$, $\mathscr{A}(S)$. These four sets of functions are all closed convex cones and in each case there exists a uniquely determined integral representation (see [2]) which will be important for the proofs of our results.

Let us agree in saying that a given function $g\colon S\to\mathbb{R}_+$ fulfills *Hoeffding's inequality of order n* iff for every sequence $X_1,\dots,X_n$ of n independent simple S-valued random variables the inequality

$$E_{\mu_1,\dots,\mu_n}[g(X_1+\dots+X_n)]\geqq E_{\bar\mu}[g(X_1+\dots+X_n)]$$

holds, where on the lefthandside X_j has distribution μ_j for all $j=1,\dots,n$, while on the righthandside each X_j has the same (average) distribution

$$\bar\mu=\frac{1}{n}\sum_{j=1}^{n}\mu_j.$$

Remark. The condition $g\geqq 0$ can of course be replaced by $\inf_{s\in S} g(s)>-\infty$. Without that assumption we could not hope for good results, because there exist unbounded positive definite functions on certain semigroups which are not moment functions, cf. [1].

The first connection between the different classes of functions defined above is given by

Theorem 1. *Let the function $g\colon S\to\mathbb{R}_+$ be given. Then the following conditions are equivalent:*

(i) *g is negative definite*

(ii) *g fulfills Hoeffding's inequality of order* 2.

Proof. (i)⇒(ii): Let g be negative definite. Then g has the following Lévy-type representation (cf. [2], Theorem 3.7):

$$g(s)=c+h(s)+\int_{\hat S\setminus\{1\}}(1-\rho(s))\,d\nu(\rho)\qquad \forall s\in S$$

where $c\in[0,\infty[$, $h\colon S\to[0,\infty[$ is additive and ν is a nonnegative Radon measure on the locally compact space $\hat S\setminus\{1\}$. Here $\hat S$ denotes the "dual" semigroup of S consisting of all multiplicative functions $\rho\colon S\to[-1,1]$ normalized such that $\rho(0)=1$, endowed with the topology of pointwise convergence; in this way $\hat S$ is always a compact abelian topological semigroup with neutral element $\rho\equiv 1$. From the above representation of g it is immediately seen to be sufficient to verify the inequality in (ii) separately for the two cases $g=h$ and $g(s)=1-\rho(s)$ for some $\rho\in\hat S$.

If g is additive then

$$\begin{aligned}E_{\mu_1,\mu_2}[g(X_1+X_2)]&=E_{\mu_1}[g(X_1)]+E_{\mu_2}[g(X_2)]=\int g\,d\mu_1+\int g\,d\mu_2\\&=2\int g\,d\bar\mu=E_{\bar\mu}[g(X_1)]+E_{\bar\mu}[g(X_2)]=E_{\bar\mu}[g(X_1+X_2)],\end{aligned}$$

and if $g=1-\rho$ for some $\rho\in\hat{S}$ we obtain

$$\begin{aligned}E_{\bar{\mu}}[g(X_1+X_2)]&=E_{\bar{\mu}}[1-\rho(X_1)\,\rho(X_2)]=1-(\int\rho\,d\bar{\mu})^2\\&\leqq 1-(\int\rho\,d\mu_1)(\int\rho\,d\mu_2)=\iint[1-\rho(s_1+s_2)]\,d\mu_1(s_1)\,d\mu_2(s_2)\\&=E_{\mu_1,\mu_2}[g(X_1+X_2)]\end{aligned}$$

where we used the obvious inequality $ab\leqq\left(\frac{a+b}{2}\right)^2$.

(ii)⇒(i): Let $k\geqq 2$, some k-tuple $(s_1,\dots,s_k)\in S^k$ and $(c_1,\dots,c_k)\in\mathbb{R}^k$ be given such that $c_1+\dots+c_k=0$. Of course we may assume that $c_0=\sum c_j^+=\sum c_j^-$ is strictly positive. Let $p_j=c_j^+/c_0$ and $q_j=c_j^-/c_0$, $j=1,\dots,k$. Then

$$\mu_1=\sum_{j=1}^{k}p_j\varepsilon_{s_j}\quad\text{and}\quad\mu_2=\sum_{j=1}^{k}q_j\varepsilon_{s_j}$$

define two probability distributions on S. By Hoeffding's inequality we obtain

$$\begin{aligned}0&\geqq\sum_{i,j=1}^{k}g(s_i+s_j)\left(\frac{p_i+q_i}{2}\cdot\frac{p_j+q_j}{2}-p_i q_j\right)\\&=\frac{1}{4}\sum_{i,j=1}^{n}g(s_i+s_j)(p_i-q_i)(p_j-q_j)=\frac{1}{4c_0^2}\sum_{i,j=1}^{n}c_i c_j g(s_i+s_j).\end{aligned}$$

Hence g is negative definite. □

Corollary. *For any positive definite function g on S the reverse inequality*

$$E_{\mu_1,\mu_2}[g(X_1+X_2)]\leqq E_{\bar{\mu}}[g(X_1+X_2)]$$

holds. If on the other hand this reverse inequality holds and g is bounded, then g $=c+f$ *for some* $c\in\mathbb{R}$ *and positive definite f.*

The proof follows from Proposition 3.10 in [2].

In [2], Theorem 4.2 it was shown, that the set of all completely alternating functions on S is an extreme subcone of the set of all negative definite functions and that the two cones coincide in case the semigroup is 2-divisible (i.e. for all $s\in S$ there is some $t\in S$ with $s=t+t$). Therefore on a 2-divisible semigroup the completely alternating functions are already characterised by Theorem 1. The situation is however different for non 2-divisible semigroups. If f. ex. $S=\mathbb{N}_0\,(=\mathbb{N}\cup\{0\})$ with usual addition then $g=1_{\{1,3,5,\dots\}}\in\mathcal{N}(S)\setminus\mathcal{A}(S)$.

Hence g fulfills Hoeffding's inequality of order 2, but not that of order 3, as the following example shows.

Let X_1, X_2, X_3 be independent Bernoulli random variables with parameters $p_1=0$, $p_2=\frac{3}{4}$ and $p_3=\frac{3}{4}$. Then $\bar{p}=\frac{1}{2}$ and

$$E_{\bar{p}}[g(X_1+X_2+X_3)]=3\bar{p}(1-\bar{p})^2+\bar{p}^3=\tfrac{1}{2}$$

whereas

$$\begin{aligned}&E_{p_1,p_2,p_3}[g(X_1+X_2+X_3)]\\&\quad=p_1(1-p_2)(1-p_3)+(1-p_1)p_2(1-p_3)\\&\qquad+(1-p_1)(1-p_2)p_3+p_1p_2p_3\\&\quad=\tfrac{3}{8}.\end{aligned}$$

In the proof of Theorem 2 we shall need the following

Lemma. *If g satisfies Hoeffding's inequality of order* 3, *then*

$$\int g\,d(\nu*\mu^2)-2\int g\,d(\nu^2*\mu)+\int g\,d(\nu^3)\leqq 0$$

for all probability measures μ, ν with finite support. (The powers here denote convolution powers.)

Proof. We apply Hoeffding's inequality to the measures $\mu_1=\mu_2=\nu$, $\mu_3=p\nu+(1-p)\mu$, where $0<p<1$. Then

$$\bar{\mu}=\frac{1-p}{3}\mu+\frac{2+p}{3}\nu,$$

and

$$\begin{aligned}&\bar{\mu}^3-\mu_1*\mu_2*\mu_3\\&\quad=\sum_{j=0}^{3}\binom{3}{j}\left(\frac{2+p}{3}\right)^j\left(\frac{1-p}{3}\right)^{3-j}\nu^j*\mu^{3-j}-p\nu^3-(1-p)\nu^2*\mu\\&\quad=\left(\frac{1-p}{3}\right)^3\mu^3+(2+p)\left(\frac{1-p}{3}\right)^2\nu*\mu^2-\frac{1}{9}(p+5)(1-p)^2\nu^2*\mu\\&\qquad+\frac{1}{27}(p+8)(1-p)^2\nu^3.\end{aligned}$$

Dividing out $(1-p)^2$ and letting then p tend to 1 we get the wanted result. □

Theorem 2. *Let a function $g\colon S\to\mathbb{R}_+$ be given. Then the following three conditions are equivalent:*

(i) *g is completely alternating,*

(ii) *g fulfills Hoeffding's inequality of any order,*

(iii) *g fulfills Hoeffding's inequality of order* 3.

Proof. (i) ⇒ (ii): From Theorem 4.4 in [2] we get the representation

$$g(s)=c+h(s)+\int_{\hat{S}_+\setminus\{1\}}(1-\rho(s))\,d\nu(\rho)$$

where $c\geqq 0$, h is additive and ν is some nonnegative Radon measure on $\hat{S}_+\setminus\{1\}$. Again it is sufficient to prove Hoeffding's inequality separately for $g=h$ and $g=1-\rho$ for some $\rho\in\hat{S}_+$. We omit the additive case. For $g=1-\rho$,

$\rho \in \hat{S}_+$, we obtain

$$\begin{aligned} E_{\bar{\mu}}[g(X_1+\ldots+X_n)] &= E_{\bar{\mu}}\left[1-\prod_{j=1}^{n}\rho(X_j)\right] \\ &= 1-(\int \rho\, d\bar{\mu})^n \leqq 1-\prod_{j=1}^{n}\int \rho\, d\mu_j \\ &= \int\int \ldots \int [1-\rho(s_1+\ldots+s_n)]\, d\mu_1(s_1)\ldots d\mu_n(s_n) \\ &= E_{\mu_1,\ldots,\mu_n}[g(X_1+\ldots+X_n)] \end{aligned}$$

where we used the well known inequality between the geometrical and the arithmetical mean of nonnegative numbers.

(iii) $\Rightarrow$ (i): Our proof consists of three steps. First we show that g also fulfills Hoeffding's inequality of order 2. Let μ and ν be two probability measures of finite support and denote by ε_0 the one-point measure in $0 \in S$. Hoeffding's inequality of order 3 is now applied to the measures

$$\mu_1 = p\,\varepsilon_0 + (1-p)\frac{\mu+\nu}{2}, \qquad \mu_2 = p\,\varepsilon_0 + (1-p)\mu,$$

$$\mu_3 = p\,\varepsilon_0 + (1-p)\nu$$

implying $\bar{\mu} = \mu_1$ and therefore

$$\begin{aligned} \bar{\mu}^3 - \mu_1 * \mu_2 * \mu_3 &= \mu_1 * (\mu_1^2 - \mu_2 * \mu_3) \\ &= \left[p\,\varepsilon_0 + (1-p)\frac{\mu+\nu}{2}\right] * \left[(1-p)\frac{\mu-\nu}{2}\right]^2 \end{aligned}$$

so that, again dividing out $(1-p)^2$ and letting p tend to 1, we may conclude

$$\int g\, d[(\mu-\nu)^2] \leqq 0$$

for all μ and ν, and this of course is only another formulation for Hoeffding's inequality of order 2.

In the second step we shall prove the wanted result for the special semigroup $S = \mathbb{N}_0$ with usual addition. By Theorem 1 we know that g is negative definite and the integral representation of g used already in the proof of Theorem 1 takes the form

$$g(k) = c + \alpha k + \int_{[-1,1[} (1-t^k)\, d\nu(t)$$

where $c \geqq 0$, $\alpha \geqq 0$ and ν is a nonnegative Radon measure on $[-1,1[$ such that $\int (1-t^k)\, d\nu(t) < \infty$ for all $k \in \mathbb{N}$. Here we used the fact that $\hat{\mathbb{N}}_0$ is isomorphic to the multiplicative semigroup $[-1,1]$. Without loss of generality we assume $c = \alpha = 0$.

We define two probability measures σ, τ on $\mathbb{N}_0$ by

$$\sigma = 2^{1-n} \sum_{\substack{i=0 \\ i\,\text{odd}}}^{n} \binom{n}{i} \varepsilon_i = 2^{-n}[(\varepsilon_0+\varepsilon_1)^n - (\varepsilon_0-\varepsilon_1)^n]$$

$$\tau = 2^{1-n} \sum_{\substack{i=0 \\ i\,\text{even}}}^{n} \binom{n}{i} \varepsilon_i = 2^{-n}[(\varepsilon_0+\varepsilon_1)^n + (\varepsilon_0-\varepsilon_1)^n]$$

so that

$$(\tau-\sigma)^2 = 2^{2-2n}(\varepsilon_0-\varepsilon_1)^{2n}$$

and

$$\sigma * (\tau-\sigma)^2 = 2^{2-3n}[(\varepsilon_0-\varepsilon_1)^{2n} * (\varepsilon_0+\varepsilon_1)^n - (\varepsilon_0-\varepsilon_1)^{3n}].$$

Observing now that for any signed measure κ with finite support on $\mathbb{N}_0$ we have

$$\int_{\mathbb{N}_0} g\,d\kappa = \sum_k \int_{[-1,1[} (1-t^k)\,d\nu(t) \cdot \kappa(\{k\}) = \int_{[-1,1[} [\kappa(\mathbb{N}_0) - \hat{\kappa}]\,d\nu$$

where $\hat{\kappa}(t) = \sum_k \kappa(\{k\}) t^k$ is the generating function of κ, we apply the Lemma and get

$$\int_{\mathbb{N}_0} g\,d[(\varepsilon_0-\varepsilon_1)^{2n} * (\varepsilon_0+\varepsilon_1)^n] = \int_{[-1,1[} -(1-t)^{2n}(1+t)^n\,d\nu(t)$$

$$\leqq \int_{[-1,1[} -(1-t)^{3n}\,d\nu(t) = \int_{\mathbb{N}_0} g\,d[(\varepsilon_0-\varepsilon_1)^{3n}]$$

i.e.

$$\int_{[-1,1[} (1-t)^{3n}\,d\nu(t) \leqq \int_{[-1,1[} (1-t)^{2n}(1+t)^n\,d\nu(t).$$

If ν' is the measure having density $1-t$ with respect to ν, then ν' is a finite Radon measure on $[-1,1[$ and

$$\int_{[-1,1[} (1-t)^{3n-1}\,d\nu'(t) \leqq \int_{[-1,1[} (1-t)^{2n-1}(1+t)^n\,d\nu'(t)$$

$$= \int_{[-1,1[} (1-t^2)^n(1-t)^{n-1}\,d\nu'(t) \leqq \int_{[-1,1[} (1-t)^{n-1}\,d\nu'(t).$$

Now we use the well known fact that on a finite measure space the L^p-norms converge to the L^∞-norm if p tends to infinity. Let

$$a = \operatorname*{ess\,sup}_{t\in[-1,1[}(1-t) \qquad \text{with respect to } \nu'.$$

Then

$$a = \lim_{n\to\infty} \Big[\int_{[-1,1[} (1-t)^{3n-1}\,d\nu'(t)\Big]^{\frac{1}{3n-1}}$$

$$\leqq \lim_{n\to\infty} \Big[\int_{[-1,1[} (1-t)^{n-1}\,d\nu'(t)\Big]^{\frac{1}{n-1}\cdot\frac{n-1}{3n-1}}$$

$$= \sqrt[3]{a}$$

so that $a \leqq 1$ and therefore $\nu'([-1,0[)=0$ and finally also $\nu([-1,0[)=0$. By Theorem 4.4 of [2] g is completely alternating.

In the last step let again S be arbitrary. For a fixed $s \in S$ consider g_s: $\mathbb{N}_0 \to \mathbb{R}_+$ defined by $g_s(n)=g(ns)$. Then g_s fulfills Hoeffding's inequality of order 3 and is therefore completely alternating. The function g is negative definite by Theorem 1 and has the representation

$$g(s)=c+h(s)+\int_{\hat{S}\setminus\{1\}} [1-\rho(s)]\,d\nu(\rho)$$

already mentioned above where again we may and do assume $c=0$ and $h=0$. Now

$$\begin{aligned} g_s(n)=g(ns) &= \int_{\hat{S}\setminus\{1\}} [1-\rho(ns)]\,d\nu(\rho) \\ &= \int_{\hat{S}\setminus\{1\}} [1-(\rho(s))^n]\,d\nu(\rho) = \int_{[-1,1]} (1-t^n)\,d\nu^{\varphi_s}(t) \end{aligned}$$

where φ_s: $\hat{S} \to [-1,1]$ is the continuous function $\varphi_s(\rho)=\rho(s)$, and where ν^{φ_s} is the image of ν under φ_s. Of course ν^{φ_s} need not be a Radon measure, but the restriction of ν^{φ_s} to $[-1,1[$ certainly is. Hence

$$0=\nu^{\varphi_s}([-1,0[)=\nu(\{\rho \in \hat{S}: \rho(s)<0\})$$

and finally, ν being a Radon measure, we get

$$\nu(\hat{S}\setminus\hat{S}_+)=\nu\Big(\bigcup_{s\in S}\{\rho: \rho(s)<0\}\Big)=0$$

thus finishing the proof of Theorem 2. □

Corollary. *For any completely monotone function g on S the reverse Hoeffding inequality*

$$E_{\mu_1,\ldots,\mu_n}[g(X_1+\ldots+X_n)] \leqq E_{\bar{\mu}}[g(X_1+\ldots+X_n)]$$

holds for every order n. If on the other hand this reverse inequality holds for $n=3$ and g is bounded, then $g=c+f$ for some $c \in \mathbb{R}$ and completely monotone f.

2. Extension to Non-Discrete Semigroups

Whereas for the characterisation of negative definite and completely alternating functions it is sufficient to consider only probability measures of finite support, one would of course like to have Hoeffding's inequality also in a non-discrete situation.

Let $(S,+,0)$ denote an abelian topological semigroup with neutral element 0, and let $M_\tau^1(S)$ be the set of all τ-smooth probability measures on S. From [7], Theorem 1 it follows easily that the convolution of two elements of $M_\tau^1(S)$ is a well defined new element of $M_\tau^1(S)$, and furthermore that $(M_\tau^1(S), *, \varepsilon_0)$ again is an abelian topological semigroup (in the usual topology of weak convergence) with neutral element ε_0.

Theorem 3. *Let S be an abelian topological semigroup and let g denote a continuous, negative definite function on S. Then g fulfills Hoeffding's inequality of order 2, and if g is furthermore completely alternating, it satisfies Hoeffding's inequality of any order (with respect to τ-smooth probability distributions).*

Proof. Let $F_n=\{g\leqq n\}$ and let two τ-smooth probability measures μ_1, μ_2 on S be given. First we assume that there is some $n\in\mathbb{N}$ such that μ_1 and μ_2 both are concentrated on F_n. There exist two nets $\{\mu_{1,\alpha}\}$ and $\{\mu_{2,\alpha}\}$ of probability measures with finite support contained in F_n such that

$$\mu_{1,\alpha}(B)\to\mu_1(B) \quad \text{and} \quad \mu_{2,\alpha}(B)\to\mu_2(B)$$

for every Borel subset $B\subseteq S$; then certainly $\mu_{1,\alpha}\to\mu_1$ and $\mu_{2,\alpha}\to\mu_2$ weakly implying $\bar{\mu}_\alpha=\frac{1}{2}(\mu_{1,\alpha}+\mu_{2,\alpha})\to\bar{\mu}=\frac{1}{2}(\mu_1+\mu_2)$, and furthermore, by continuity of convolution, $\mu_{1,\alpha}*\mu_{2,\alpha}\to\mu_1*\mu_2$ as well as $\bar{\mu}_\alpha*\bar{\mu}_\alpha\to\bar{\mu}*\bar{\mu}$.

The subadditivity of negative definite functions – cf. [2], Proposition 3.5 – shows that g is bounded on F_n+F_n, too, and hence bounded on the support of $\mu_1*\mu_2$. Now we can apply Theorem 1 to obtain

$$\begin{aligned}\int g\,d(\bar{\mu}*\bar{\mu})&=\lim_\alpha\int g\,d(\bar{\mu}_\alpha*\bar{\mu}_\alpha)\\&\leqq\lim_\alpha\int g\,d(\mu_{1,\alpha}*\mu_{2,\alpha})=\int g\,d(\mu_1*\mu_2).\end{aligned}$$

The general result will now be deduced by using the conditional probability measures

$$\mu_{1,n}(B)=\mu_1(B\cap F_n)/\mu_1(F_n) \quad \text{and} \quad \mu_{2,n}(B)=\mu_2(B\cap F_n)/\mu_2(F_n)$$

which are certainly well defined if n is large enough. It is almost immediate that the measures $\mu_{1,n}$ and $\mu_{2,n}$ are also τ-smooth; consequently we have

$$\int g\,d(\bar{\mu}_n*\bar{\mu}_n)\leqq\int g\,d(\mu_{1,n}*\mu_{2,n}) \qquad \text{for all } n,$$

where of course $\bar{\mu}_n=\frac{1}{2}(\mu_{1,n}+\mu_{2,n})$. Letting $h(s,t)=g(s+t)$ we observe that

$$\begin{aligned}\int g\,d(\bar{\mu}_n*\bar{\mu}_n)&=\tfrac{1}{4}\Big[\int h\,d(\mu_{1,n}\otimes\mu_{1,n})+2\int h\,d(\mu_{1,n}\otimes\mu_{2,n})\\&\quad+\int h\,d(\mu_{2,n}\otimes\mu_{2,n})\Big]=\tfrac{1}{4}\Bigg[\frac{1}{(\mu_1(F_n))^2}\int_{F_n\times F_n}h\,d(\mu_1\otimes\mu_1)\\&\quad+\frac{2}{\mu_1(F_n)\,\mu_2(F_n)}\int_{F_n\times F_n}h\,d(\mu_1\otimes\mu_2)+\frac{1}{(\mu_2(F_n))^2}\int_{F_n\times F_n}h\,d(\mu_2\otimes\mu_2)\Bigg]\\&\to\tfrac{1}{4}\Big[\int h\,d(\mu_1\otimes\mu_1)+2\int h\,d(\mu_1\otimes\mu_2)+\int h\,d(\mu_2\otimes\mu_2)\Big]\\&=\int h\,d(\bar{\mu}\otimes\bar{\mu})=\int g\,d(\bar{\mu}*\bar{\mu})\end{aligned}$$

and

$$\begin{aligned}\int g\,d(\mu_{1,n}*\mu_{2,n})&=\frac{1}{\mu_1(F_n)\,\mu_2(F_n)}\int_{F_n\times F_n}h\,d(\mu_1\otimes\mu_2)\\&\to\int h\,d(\mu_1\otimes\mu_2)=\int g\,d(\mu_1*\mu_2).\end{aligned}$$

This proves the first assertion and we omit the similar argument for completely alternating functions. □

Example 1. Let S be the "classical" additive semigroup $\mathbb{R}_+^k$. Every probability measure on S is τ-smooth (even a Radon measure), S being a separable metric space. S is 2-divisible so that negative definite and completely alternating actually mean the same. If now $g \in \mathcal{N}(\mathbb{R}_+^k)$ and $X_1, \ldots, X_n$ are independent $\mathbb{R}_+^k$-valued random variables with distributions $\mu_1, \ldots, \mu_n$, then under the restriction of a given average distribution $\bar{\mu} = \mu_0$, the expectation

$$E_{\mu_1, \ldots, \mu_n}[g(X_1 + \ldots + X_n)]$$

is minimized for $\mu_1 = \mu_2 = \ldots = \mu_n \,(= \mu_0)$.

In dimension $k=1$, standard examples of negative definite functions are

$$g(s) = 1 - e^{-s},$$
$$g(s) = \log(1+s),$$
$$g(s) = s^\alpha, \quad \text{where} \quad 0 < \alpha \leqq 1,$$
$$g(s) = ar\cosh(e^s) = s + \log(1 + \sqrt{1 - e^{-2s}}).$$

The fact, that in this special situation the composition of two negative definite functions has again this property, allows to derive a lot more examples.

Remark. It is easily seen that the dual semigroup of $\mathbb{R}_+^k$ is topologically isomorphic to $[0, \infty]^k$, also considered as an additive semigroup. Hence there exist discontinuous negative definite functions on $\mathbb{R}_+^k$, but (contrary to the case of the group $\mathbb{R}^k$) every negative definite function is Borel measurable. We have a proof, based on Fubinis Theorem, that on the semigroup $\mathbb{R}_+^k$ the continuity of g actually is not needed for the validity of Hoeffding's inequalities. The proof works if $\hat{S} \setminus \{1\}$ is σ-compact and $(\rho, s) \mapsto \rho(s)$ is Borel measurable on $\hat{S} \times S$.

Example 2. Let S be $\mathbb{N}_0$ under usual addition. Here $\hat{S}$ can be identified with the multiplicative topological semigroup $[-1, 1]$ and continuity problems do not arise. $\mathbb{N}_0$ is not 2-divisible and we already mentioned that $g = 1_{\{1,3,5,\ldots\}}$ belongs to $\mathcal{N}(\mathbb{N}_0)$ but not to $\mathcal{A}(\mathbb{N}_0)$. Let X and Y be independent $\mathbb{N}_0$-valued random variables with distributions μ resp. ν. Then we have from Theorem 3 the rather curious result that $P(X+Y$ is odd) among all μ, ν with a given average, gets minimal for $\mu = \nu$.

If the semigroup S is "finite dimensional", there is a more or less "natural" topology, under which S becomes a topological semigroup. In the infinite dimensional case, however, it may be reasonable to consider different topologies on S, as the following examples shows.

Example 3. Let X be a completely regular Hausdorff space and let $S = C_+(X)$, the cone of all bounded continuous nonnegative functions on X. With usual addition S is a 2-divisible semigroup. Consider on $C(X)$ the L^1-topology $\mathcal{T}_1$, generated by all seminorms $q_\mu(f) = \int |f| \, d\mu$, μ being any totally finite nonnegative Radon measure on X, and the socalled strict topology $\mathscr{T}_2$, generated by the seminorms

$$p_\alpha(f) = \sup\{|\alpha(x) f(x)| : x \in X\},$$

where α runs through all bounded measurable functions on X vanishing at infinity, i.e. $\{|\alpha| \geqq \varepsilon\}$ is relatively compact for all $\varepsilon > 0$. Of course in both topologies S becomes a topological semigroup. Positive and negative definite functions on S have been studied in [6]. If $g \in \mathcal{N}(S)$ satisfies a rather weak continuity property - the restriction of g to the "unit ball" $\{f \in S: f(x) \leqq 1 \,\forall x \in X\}$ should be $\mathscr{T}_2$-continuous in 0 - then by Theorem 4.2 in [6] the function g is $\mathscr{T}_1$-continuous everywhere and therefore fulfills Hoeffding's inequalities for all probability measures which are τ-smooth with respect to $\mathscr{T}_1$, i.g. a larger class of measures than those being τ-smooth w.r. to $\mathscr{T}_2$.

References

1. Berg, C., Christensen, J.P.R., Jensen, C.U.: A remark on the multidimensional moment problem. Math. Ann. **243**, 163–169 (1979)
2. Berg, C., Christensen, J.P.R., Ressel, P.: Positive definite functions on abelian semigroups. Math. Ann. **223**, 253–274 (1976)
3. Bickel, P.J., van Zwet, W.R.: On a theorem of Hoeffding. In: Asymptotic Theory of Statistical Tests and Estimation (ed. by I.M. Chakravarti), pp. 307–324. New York: Academic Press 1980
4. Choquet, G.: Theory of capacities. Ann. Inst. Fourier **5**, 131–295 (1954)
5. Hoeffding, W.: On the distribution of the number of successes in independent trials. Ann. Math. Statist. **27**, 713–721 (1956)
6. Hoffmann-Jørgensen, J., Ressel, P.: On completely monotone functions on $C_+(X)$. Math. Scand. **40**, 79–93 (1977)
7. Ressel, P.: Some continuity and measurability results on spaces of measures. Math. Scand. **40**, 69–78 (1977)

Received March 17, 1980; in revised form March 18, 1981