Rendiconti di Matematica, Serie VII
Volume 28, Roma (2008), 63–81

Exchangeability and semigroups

PAUL RESSEL

Abstract: *Exchangeability of a "random object" is a strong symmetry condition, leading in general to a convex set of distributions not too far from a "simplex" - a set easily described by its extreme points, in this case distributions with very special properties as for example iid coin tossing sequences in de Finetti's original result. Although in most cases of interest the symmetry is defined via a non–commutative group acting on the underlying space, it very often can be described by a suitable factorization involving an abelian semigroup. The factorizing function typically turns out to be positive definite, and results from Harmonic Analysis on semigroups become applicable. In this way many known theorems on exchangeability can be given an alternative proof, more analytic/algebraic in a sense, but also new results become available.*

1 – Introduction

A sequence $X = (X_1, X_2, \dots)$ of random variables is called *exchangeable* if for any permutation π of $\mathbb{N}$ the sequence $(X_{\pi(1)}, X_{\pi(2)}, \dots)$ has the same distribution as X; of course it is enough to require this property for *finite* permutations π (in the sense that $\{i \in \mathbb{N} \mid \pi(i) \neq i\}$ is a finite set). This holds obviously for an iid–sequence and so also for a mixture (in distribution) of iid's, since exchangeable distributions form a convex set. As is well known, in 1930 Bruno de Finetti published the pathbreaking result that for $\{0,1\}$–valued random variables the converse holds, too: exchangeable sequences are precisely the mixtures of iid Bernoulli sequences. A few years later de Finetti generalized this to real–valued random variables, and in 1955 Hewitt and Savage proved the

Key Words and Phrases: *Exchangeability – Semigroups – Positive definite functions – de Finetti's theorem – Random partition.*
A.M.S. Classification: 60E05, 43A35

corresponding result for arbitrary compact Hausdorff spaces, from which it is immediately seen to be true also for Borel subsets of compact spaces, for example for locally compact spaces. In ([6], Theorem 4) a further generalization to completely regular Hausdorff spaces was shown.

In contrast to this "global" point of view (i.e. considering *all* exchangeable distributions) a different kind of question seems natural: is it possible to characterize mixtures of iid–sequences of a particular type, say normal or Poisson distributed, or with a special form of their Fourier or Laplace transform, or even (for non–negative random variables) of their multivariate survival function?

An attempt for fairly general answers was given in [6] and subsequent papers. Here we present a certain overview, and as "Main Theorem" a new result which in a way is a de Finetti theorem for positive definite functions on abelian semigroups, which appear (perhaps surprising) as a natural tool in this connection.

After explaining some basic notions concerning positive definite and related functions on semigroups in Section 3, the main result will be presented in Section 4. The extended de Finetti–type theorem in Section 5 is given a new proof, based on the main theorem, and followed by a few typical examples. The above mentioned theorem of Hewitt and Savage is shown in Section 6 to be another "almost straightforward" consequence of the main theorem. Finally, the closing Section 7 presents a different point of view to the main theorem, followed by an application to exchangeable random partitions.

2 – Why semigroups?

They enter the scene naturally, as can be seen already in de Finetti's original result.

Let P be an exchangeable probability measure on the space of all (infinite) $0-1$ sequences, abbreviated $P \in M_+^{1,e}(\{0,1\}^\infty)$, the "$e$" referring to exchangeability. Then $P(x_1, \ldots, x_n)$ depends only on $x_1 + \ldots + x_n$, i.e.

$$P(x_1, \ldots, x_n) = \varphi_n\left(\sum_{i=1}^n x_i\right) = \varphi\left(\sum_{i=1}^n x_i, n\right) = \\ = \varphi\left(\sum_{i=1}^n (x_i, 1)\right)$$

with φ defined on the set

$$S := \big\{(k,n) \in \mathbb{N}_0^2 \mid k \leq n\big\}$$

which is a (sub-) semigroup inside $\mathbb{N}_0^2$. The crucial point will be that φ turns out to be a socalled *positive definite* function, therefore a (unique) mixture of socalled *characters*, taking here the form

$$\sigma : (k,n) \longmapsto \sigma(k,n) = p^k q^{n-k}, \quad p,q \in \mathbb{R}$$

and it is easy to see that only characters with $p,q \geq 0$ and $p+q=1$ play a rôle. Inserting this we get (slightly abusing the letter μ as a measure on the characters resp. on $[0,1]$)

$$\begin{aligned} P(x_1,\dots,x_n) &= \varphi\left(\sum_{i=1}^n (x_i,1)\right) = \\ &= \int \sigma\left(\sum_{i=1}^n (x_i,1)\right) d\mu(\sigma) = \\ &= \int \prod_{i=1}^n \sigma(x_i,1) d\mu(\sigma) = \\ &= \int_0^1 \prod_{i=1}^n p^{x_i}(1-p)^{1-x_i} d\mu(p) \end{aligned}$$

for some (unique) $\mu \in M_+^1([0,1])$, which is de Finetti's result, proved in 1930, cf. [3].

3 – Basic definitions and notations

We will make use in a crucial way of some notions and results about positive definite and related functions on semigroups, an introduction to which can be found in [1], Chapter 4.

Let S denote an abelian semigroup, written additively, with neutral element 0, and possibly with an involution, i.e. a mapping $s \longmapsto s^-$, with $(s+t)^- = s^- + t^-, 0^- = 0$ and $(s^-)^- = s$ which in many cases is just the identity.

$\sigma : S \longrightarrow \mathbb{C}$ is a *character* iff

$$\sigma(s+t) = \sigma(s)\cdot\sigma(t), \quad \sigma(s^-) = \overline{\sigma(s)}, \quad \sigma(0) = 1$$

$\varphi : S \longrightarrow \mathbb{C}$ is *positive definite* (abbrev. "p.d.") iff

$$\sum_{j,k=1}^n c_j \overline{c_k} \varphi(s_j + s_k^-) \geq 0 \quad \forall\, n \in \mathbb{N}, c_j \in \mathbb{C}, s_j \in S$$

$\varphi : S \longrightarrow \mathbb{C}$ is *completely positive definite* ("c.p.d.") iff $s \longmapsto \varphi(s+a)$ is positive definite $\forall a \in S$;

$\alpha : S \longrightarrow \mathbb{R}_+$ is an *absolute value* iff

$$\alpha(s+t) \leq \alpha(s)\cdot\alpha(t), \quad \alpha(s^-) = \alpha(s), \quad \alpha(0) = 1$$

$f : S \longrightarrow \mathbb{C}$ is α–bounded (with α a fixed absolute value) iff

$$|f(s)| \leq C\cdot\alpha(s) \quad \forall\, s \in S, \quad \text{for some} \quad C \geq 0, \quad \text{briefly:} \quad |f| \leq C\alpha$$

(if furthermore $f(0) = 1$ and f is p.d., then one can take $C = 1$);
f is *exponentially bounded* iff it is α–bounded with respect to some absolute value α;
$S^* :=$ set of all characters of S;
$\mathcal{P}(S) :=$ set of all positive definite functions on S;
$S^\alpha := \{\sigma \in S^* \mid \sigma \text{ is } \alpha\text{–bounded}\}$ then $S^\alpha = \{\sigma \in S^* \mid |\sigma| \leq \alpha\}$;
$\mathcal{P}^\alpha(S) := \{\varphi \in \mathcal{P}(S) \mid \varphi \text{ is } \alpha\text{–bounded}\}$;
$\hat{S} :=$ all bounded characters on S, then $\hat{S} = \{\sigma \in S^* \mid |\sigma| \leq 1\}$;
$\mathcal{P}^b(S) :=$ all bounded positive definite functions on S.

For any set B of complex functions on S, the symbols B_+ and B_1 denote respectively $B \cap \{f \mid f(x) \geq 0 \quad \forall s \in S\}$ and $B \cap \{f \mid f(0) = 1\}$.

It is easily seen that

$S^* \subseteq \mathcal{P}_1(S) := \{\varphi \in \mathcal{P}(S) \mid \varphi(0) = 1\}$;
$S^\alpha \subseteq \mathcal{P}_1^\alpha(S), \quad \varphi \in \mathcal{P}_1^\alpha(S) \Longrightarrow |\varphi| \leq \alpha$;
$\hat{S} \subseteq \mathcal{P}_1^b(S), \quad \varphi \in \mathcal{P}_1^b(S) \Longrightarrow |\varphi| \leq 1$

and each $\sigma \in S_+^*$ is even c.p.d.

4 – The main result

If K is a non–empty compact convex subset of some locally convex vector space, then K is by Krein-Milman's theorem the closed convex hull of $ex(K)$, the extreme points of K. If $ex(K)$ is closed, and if furthermore the representation of points in K as barycenters of (Radon) measures on $ex(K)$ is unique, K is called a *Bauer simplex*. In all subsequent applications K will be a subset of $\mathbb{C}^S$, the set of all complex–valued functions on S, equipped with the topology of pointwise convergence. Note that in this case for any given $\gamma : S \longrightarrow \mathbb{R}_+$ the set $K := \{f \in \mathbb{C}^S \mid |f| \leq \gamma\}$ is compact (and convex).

Let us recall first the basic result concerning exponentially bounded positive definite functions. We shall only consider abelian semigroups with a neutral element.

Theorem (Berg/Maserick, cf. [2] or [1], 4.2.6 and 4.2.7). *For a semigroup S and an absolute value α on S the set $\mathcal{P}_1^\alpha(S)$ is a Bauer simplex with S^α as its* *set of extreme points. In other words, for any $\varphi \in \mathcal{P}_1^\alpha(S)$ there exists a unique Radon probability measure μ on S^α such that*

$$\varphi(s) = \int \sigma(s) d\mu(\sigma) \quad \forall\, s \in S\,.$$

We'll also make use of the following

Corollary (cf. [6], Proposition 1). *If $\varphi \in \mathcal{P}^\alpha(S)$ is completely positive definite, then the unique measure representing φ is concentrated on S_+^α.*

From now on we will typically deal with two semigroups R, S, and a mapping $t : R \longrightarrow S$ with the properties $t(0) = 0$, $t(r^-) = (t(r))^-$ and such that $t(R)$ generates S as a semigroup. Furthermore, a function $\beta : R \longrightarrow \mathbb{C}\backslash\{0\}$ is given with $\beta(0) = 1$ and $\beta(r^-) = \overline{\beta(r)}$ for all r; in most of the examples we'll have $\beta \equiv 1$. The direct product

$$R^{(\infty)} := \{(r_1, r_2, \dots) \in R^\infty \mid r_i = 0 \ \text{ finally}\}$$

of countably many copies of R will play a particular rôle.

The following result is new in this generality.

Main Theorem. *Let R and S be semigroups, and $t : R \longrightarrow S$, $\beta : R \longrightarrow \mathbb{C} \setminus \{0\}$ be functions as just described*:

(i) *if $\Phi(r_1, r_2, \dots) := \prod \beta(r_i) \cdot \varphi(\sum t(r_i))$ for some function $\varphi : S \longrightarrow \mathbb{C}$, and Φ is positive definite then so is φ;*

(ii) *if furthermore $|\Phi(r_1, r_2, \dots)| \le C \cdot \prod \gamma(r_i)$ for some function $\gamma : R \longrightarrow \mathbb{R}_+, \gamma(0) = 1$, and some $C > 0$, then*

$$\alpha(s) := \inf \left\{ \prod \frac{\gamma(r_i)}{|\beta(r_i)|} \mid \sum t(r_i) = s \right\}$$

is an absolute value on S, φ is α–bounded, and the measure μ representing φ is concentrated on

$$W := \{\sigma \in S^\alpha \mid \beta \cdot (\sigma \circ t) \quad \textit{is positive definite on} \quad R\}$$

(iii) *conversely, for $\mu \in M_+(W)$ and $\varphi(s) := \int \sigma(s) d\mu(\sigma)$ the function Φ as defined in* (i) *is positive definite and fulfills* (ii) *for some $C > 0$ and some function γ;*

(iv) *a corresponding result holds for completely positive definite functions, the measure in* (ii) *being then concentrated on W_+.*

For the proof the following lemma is crucial. Since the statement here differs slightly from earlier presentations, we present it with the (short) proof. Recall that for a non–empty set M a function $\psi : M \times M \longrightarrow \mathbb{C}$ is a *positive semidefinite kernel* iff for any finite subset $\{x_1, \ldots, x_k\} \subseteq M$ the matrix $(\psi(x_i, x_j))_{i,j \le k}$ is positive semidefinite.

Approximation lemma. *Let $p \ge 2$ be an integer, M a non-empty set, $\psi : M \times M \longrightarrow \mathbb{C}$ a positive semidefinite kernel, $(a_{ij}) \in \mathbb{C}^{p \times p}$ a given $p \times p$–matrix. Suppose that for each $n \in \mathbb{N}$ there exist $\{x_{jm}^n \mid j = 1, \ldots, p; m = 1, \ldots, n\} \subseteq M$ such that*

$$\psi\left(x_{ik}^n, x_{jm}^n\right) = a_{ij} \quad \forall\, (i,k) \ne (j,m)$$

and

$$\sup_{j,m,n} \psi\left(x_{jm}^n, x_{jm}^n\right) < \infty\,.$$

Then (a_{ij}) is positive semidefinite.

Proof. Let $c_1, \ldots, c_p \in \mathbb{C}$ be given; with $\{x_{jm}^n\}$ as indicated put $d_{jm} := c_j/n$. Then

$$\begin{aligned} 0 &\le \sum_{i,j=1}^{p} \sum_{k,m=1}^{n} d_{ik} \bar{d}_{jm} \psi\left(x_{ik}^n, x_{jm}^n\right) = \\ &= \sum_{\substack{i,j=1 \\ i \ne j}}^{p} c_i \bar{c}_j a_{ij} + \frac{n^2 - n}{n^2} \sum_{j=1}^{p} |c_j|^2 a_{jj} + \frac{1}{n^2} \sum_{j=1}^{p} \sum_{m=1}^{n} |c_j|^2 \psi(x_{jm}^n, x_{jm}^n) = \\ &= \sum_{i,j=1}^{n} c_i \bar{c}_j a_{ij} + R_n \end{aligned}$$

where $R_n := -\frac{1}{n} \sum_{j=1}^{p} |c_j|^2 a_{jj} + \frac{1}{n^2} \sum_{j=1}^{p} \sum_{m=1}^{n} |c_j|^2 \psi(x_{jm}^n, x_{jm}^n)$ and so $R_n \longrightarrow 0$ for $n \longrightarrow \infty$, showing positive semidefiniteness of (a_{ij}). □

Proof of the Main Theorem.

(i) Let $s_1, \ldots, s_p \in S$ and $n \in \mathbb{N}$ be given. By assumption

$$s_j = \sum_{\ell=1}^{q_j} t(r_{j\ell}) \quad \text{for suitable} \quad r_{j\ell} \in R\,.$$

Let $\{N_{jm} \mid j = 1, \ldots, p; m = 1, \ldots, n\}$ be disjoint subsets of $\mathbb{N}$ with cardinalities $|N_{jm}| = q_j \;\forall\, j, m$, and define $x_{jm} \in R^{(\infty)}$ (for $N_{jm} = \{\nu_1, \ldots, \nu_{q_j}\}$) by

$$x_{jm}(\nu_\ell) := r_{j\ell}, \quad x_{jm}(i) := 0 \quad \text{for} \quad i \notin N_{jm}\,.$$

Put $\xi_j := \prod_{\ell=1}^{q_j} \beta(r_{j\ell})$, $j = 1, \dots, p$. Then for $(i,k) \neq (j,m)$

$$\Phi(x_{ik} + x_{jm}^-) = \xi_i \bar{\xi}_j \varphi(s_i + s_j^-)\,,$$

and $\Phi(x_{jm} + x_{jm}^-) = \prod_{\ell=1}^{q_j} \beta(r_{j\ell} + r_{j\ell}^-)\varphi(\sum_{\ell=1}^{q_j} t(r_{j\ell} + r_{j\ell}^-))$, independent of m and n. Hence by the Approximation lemma $(\xi_i \bar{\xi}_j \varphi(s_i + s_j^-))_{i,j \leq p}$ is positive definite, and so is also $(\varphi(s_i + s_j^-))_{i,j \leq p}$.
Suppose now Φ to be c.p.d., and let an additional element $a \in S$ be given, $a = t(r_1) + \ldots + t(r_v)$. Choose in the preceding argument the $N_{jm} \subseteq \mathbb{N} \setminus \{1, \dots, v\}$, and define $y \in R^{(\infty)}$ by $y(1) := r_1, \dots, y(v) := r_v, y(i) := 0$ else. Then for $(i,k) \neq (j,m)$

$$\Phi(y + x_{ik} + x_{jm}^-) = \xi_i \bar{\xi}_j \varphi(a + s_i + s_j^-) \cdot \prod_{\ell=1}^{v} \beta(r_\ell)$$

and if the positive semidefinite matrix on the RHS is not identically zero, $\prod_{\ell=1}^{v} \beta(r_\ell) > 0$ and then $(\varphi(a + s_i + s_j^-))$ is positive semidefinite, i.e. φ is completely positive definite.

(ii) If $s = \sum t(r_j)$ we get from

$$\Phi(r_1, r_2, \dots) = \prod \beta(r_j) \varphi(\sum t(r_j))$$

that

$$|\varphi(s)| \leq C \cdot \prod \frac{\gamma(r_j)}{|\beta(r_j)|}\,,$$

hence

$$\frac{1}{C}|\varphi(s)| \leq \alpha(s) := \inf \left\{ \prod \frac{\gamma(r_j)}{|\beta(r_j)|} \mid \sum t(r_j) = s \right\}$$

and α is immediately seen to be an absolute value. The function φ being positive definite and α–bounded, has a unique representing measure μ supported by the compact set S^α in view of the Berg/Maserick theorem. Define $f : S^* \longrightarrow \mathbb{C}$ by

$$f(\sigma) := \sum_{u,v=1}^{w} c_u \bar{c}_v \beta(a_u + a_v^-) \sigma(t(a_u + a_v^-)) \tag{$*$}$$

for given $a_1, \dots, a_w \in R$ and $c_1, \dots, c_w \in \mathbb{C}$. Then f is continuous, and on the compact subset S^α the function f is bounded. We want to show that f is μ–a.e. nonnegative, or equivalently that the measure $\nu := f \cdot \mu$ is nonnegative. This will be shown if $\hat{\nu}(s) := \int \sigma(s) d\nu(\sigma)$ turns out to be positive definite, by the Berg/Maserick theorem.

Let again $s_1, \dots, s_p \in S$, $d_1, \dots, d_p \in \mathbb{C}$ and $n \in \mathbb{N}$ be given, with

$$s_j = \sum_{\ell=1}^{q_j} t(r_{j\ell})$$

as in the proof of (i).

Let now $\{N_{ujm} \mid u = 1, \dots, w, j = 1, \dots, p; m = 1, \dots, n\}$ be disjoint subsets of $\mathbb{N} \setminus \{1\}$ with $|N_{ujm}| = q_j \ \forall\, u, j, m$, say $N_{ujm} = \{\nu_1, \dots, \nu_{q_j}\}$, and define $x_{ujm} \in R^{(\infty)}$ by

$$\begin{aligned} x_{ujm}(1) &:= a_u \\ x_{ujm}(\nu_\ell) &:= r_{j\ell} \\ x_{ujm}(i) &:= 0 \ \text{ else.} \end{aligned}$$

Put $\xi_j := \prod_{\ell=1}^{q_j} \beta(r_{j\ell})$. Then for $(u, i, k)) \neq (v, j, \ell)$

$$\Phi(x_{uik} + x_{vj\ell}^-) = \xi_i \bar{\xi}_j \beta(a_u + a_v^-)\varphi(t(a_u + a_v^-) + s_i + s_j^-)$$

and $\Phi(x_{uik} + x_{uik}^-)$ is again bounded uniformly in u, i, k, n. So again the matrix (with index set $A := \{1, \dots, w\} \times \{1, \dots, p\}$)

$$(\beta(a_u + a_v^-) \cdot \varphi(t(a_u + a_v^-) + s_i + s_j^-))_{(u,i),(v,j)\in A}$$

is positive semidefinite, leading to

$$\begin{aligned} \sum_{i,j=1}^{p} d_i \bar{d}_j \hat{\nu}(s_i + s_j^-) = &\sum_{i,j=1}^{p} \sum_{u,v=1}^{w} c_u d_i \bar{c}_v \bar{d}_j \times \\ &\times \int \beta(a_u + a_v^-)\sigma(t(a_u + a_v^-) + s_i + s_j^-) d\mu(\sigma) = \\ = &\sum_{i,j} \sum_{u,v} c_u d_i \bar{c}_v \bar{d}_j \beta(a_u + a_v^-) \cdot \varphi(t(a_u + a_v^-) + s_i + s_j^-) \geq 0. \end{aligned}$$

We have shown $f \cdot \mu$ to be a positive measure , i.e.

$$\mu(f^{-1}(\mathbb{C} \setminus \mathbb{R}_+)) = 0\,.$$

Now $f^{-1}(\mathbb{C} \setminus \mathbb{R}_+)$ is open, and μ is a Radon measure, hence

$$\mu\left(\bigcup f^{-1}(\mathbb{C} \setminus \mathbb{R}_+)\right) = 0$$

the union being taken over all functions f of the form $(*)$. Hence μ–almost surely $r \longmapsto \beta(r) \cdot \sigma(t(r))$ is positive definite on R.

(iii) W is obviously closed, hence compact, so that $C := \mu(W) < \infty$. For $\sigma \in W$ the function

$$\Phi_\sigma(r_1, r_2, \dots) := \prod \beta(r_i) \cdot \sigma(t(r_i))$$

is positive definite as a (tensor) product of such functions. Also

$$|\beta(r)\sigma(t(r))| \le |\beta(r)| \cdot \alpha(t(r)) =: \gamma(r)\,.$$

Now $\Phi = \int \Phi_\sigma d\mu(\sigma)$ is positive definite as a mixture of positive definite functions, and

$$|\Phi(r_1, r_2, \dots)| \le C \cdot \prod \gamma(r_i) \quad \forall\, r_1, r_2, \dots \in R\,.$$

(iv) See the end of the proof of (i) . □

One of the most direct corollaries is the following result, characterizing *spherically exchangeable* (or *symmetric*) sequences, i.e. sequences of real random variables whose finite dimensional distributions are invariant under rotations.

Theorem (Schoenberg, 1938). *Every infinite spherically exchangeable random sequence is a unique variance mixture of centered iid normal sequences. Or formally:*

$$\begin{aligned} &P \in M_+^1(\mathbb{R}^\infty) \quad \textit{is spherically symmetric} \\ &\iff P = \int_0^\infty N(0,c)^\infty d\mu(c) \quad \exists!\mu \in M_+^1(\mathbb{R}_+)\,. \end{aligned}$$

Proof. Only one direction needs a proof. Given a spherically exchangeable P we let Φ be its characteristic function, i.e.

$$\begin{aligned} \Phi(r_1, r_2, \dots) &:= E\left[\exp\left(i \sum r_j X_j\right)\right], \quad (r_1, r_2, \dots) \in \mathbb{R}^{(\infty)} = \\ &= \varphi\left(\sum r_j^2\right) \end{aligned}$$

for some function $\varphi : \mathbb{R}_+ \longrightarrow \mathbb{C}$, by assumption. With $t(r) = r^2, \beta \equiv \gamma \equiv 1$, we get from the Main Theorem that φ is a bounded positive definite function with $\varphi(0) = 1$. Then, for example applying the Berg/Maserick theorem, φ has the unique integral representation

$$\varphi(s) = \int e^{-\lambda s} d\mu(\lambda), \quad \mu \in M_+^1([0,\infty])$$

(with $e^{-\lambda\infty} = 1_{\{0\}}(\lambda), \lambda \in \mathbb{R}_+$). Now φ is obviously continuous, leading to $\mu(\{\infty\}) = 0$, and then from

$$\Phi(r_1, r_2, \dots) = \int_0^\infty e^{-\lambda \sum r_j^2} d\mu(\lambda)$$

we read off the wanted result. □

Schoenberg (cf. [8]) proved this result in the totally different connection of the imbedding problem for quasi-metric spaces into a Hilbert space.

With only slightly more effort we get the following characterization of

Mixtures of the full 2-parameter normal family. Let $X = (X_1, X_2, \dots)$ be any real random sequence with characteristic function Φ. Then

$$\Phi(r_1, r_2, \dots) = \varphi\left(\sum r_j, \sum r_j^2\right) \quad \text{for some} \quad \varphi$$

iff

$$P^X = \int_{\mathbb{R}\times\mathbb{R}_+} N(a,c)^\infty d\mu(a,c) \quad \text{for some} \quad \mu \in M_+^1(\mathbb{R}\times\mathbb{R}_+)\,.$$

For a proof, see [6], Example 6.

Different transforms may of course be used. An example with Laplace transforms is this:

Let $X = (X_1, X_2, \dots)$ *be non–negative random variables. Then*

$$E\left[\exp\left(-\sum r_j X_j\right)\right] = \varphi\left(\prod(1+r_j)\right)\,.$$

For some $\varphi : [1,\infty[\longrightarrow \mathbb{R}$ *iff*

$$P^X = \int_0^\infty \gamma_\lambda^\infty d\mu(\lambda)$$

where γ_λ *denotes the Gamma* $(\lambda, 1)$ *distribution, with* $\gamma_1 = e_1$*, the exponential distribution with parameter* 1; cf. Example 8 in [6].

The natural question if mixtures of exponential iid sequences can be characterized similarly, can be answered immediately:

For a non–negative sequence X *we have*

$$P^X = \int_0^\infty e_\lambda^\infty d\mu(\lambda)$$

iff

$$P(X_1 \geq a_1, X_2 \geq a_2, \dots) = \varphi\left(\sum a_j\right), \quad a \in \mathbb{R}_+^{(\infty)}$$

for some $\varphi : \mathbb{R}_+ \longrightarrow \mathbb{R}$; cf. Example 11 in [6].

5 – De Finetti's theorem in extended form

Theorem. *Let $\mathcal{X}$ be a finite or countable set, S a semigroup, $t : \mathcal{X} \longrightarrow S$ such that $t(\mathcal{X})$ generates $S \setminus \{0\}$, $\beta : \mathcal{X} \longrightarrow]0,\infty[$, $\varphi : S \longrightarrow \mathbb{R}_+$. Then $P \in M_+^1(\mathcal{X}^\infty)$ fulfills*

$$P(x_1,\dots,x_n) = \prod_{i=1}^n \beta(x_i)\cdot\varphi\left(\sum_{i=1}^n t(x_i)\right) \quad \forall\, n, x_i$$

iff

$$P = \int \kappa_\sigma^\infty \, d\mu(\sigma)$$

where $\mu \in M_+^1(S_+^*)$ is concentrated on

$$W := \left\{\sigma \in S_+^* \mid \kappa_\sigma := \beta\cdot(\sigma\circ t) \in M_+^1(\mathcal{X})\right\}$$

(*cf. Theorem* 4 *in* [6]*; we'll derive it here as a consequence of the Main Theorem*).

Proof. Let $R := \{1_{\{x\}} \mid x \in \mathcal{X}\} \cup \{0,1\}$ with pointwise multiplication, considered as a subsemigroup of $\mathbb{R}^{\mathcal{X}}$, add an absorbing element ζ to $S, S' := S \cup \{\zeta\}$, and define $t' : R \longrightarrow S'$ by $t'(1_{\{x\}}) := t(x), t'(1) := 0, t'(0) := \zeta$. Put $\beta'(1_{\{x\}}) := \beta(x), \beta'(1) := 1, \beta'(0) := 2$ (or any number > 1), $\varphi(\zeta) := 0$, and let $X_1, X_2, \dots$ be the natural projections $\mathcal{X}^\infty \longrightarrow \mathcal{X}$. Then

$$\begin{aligned}\Phi(r_1,r_2,\dots) :&= E[r_1(X_1)\cdot r_2(X_2)\cdot\ldots] = \\ &= \prod \beta'(r_j)\cdot\varphi\left(\sum t'(r_j)\right)\end{aligned}$$

for all $(r_1, r_2, \dots) \in R^{(\infty)}$.

Denoting the semigroup operation in $R^{(\infty)}$ by "$\oplus$" we get for $r^{(1)},\dots,r^{(n)} \in R^{(\infty)}$ and $c_1,\dots,c_n \in \mathbb{R}$

$$\sum_{i,j=1}^n c_i c_j \Phi(r^{(i)} \oplus r^{(j)}) = E\left\{\left[\sum_{i=1}^n c_i r^{(i)}(X)\right]^2\right\} \geq 0$$

where $r(X) := r_1(X_1)\cdot r_2(X_2)\cdot\ldots$ for $r \in R^{(\infty)}$, showing Φ to be positive definite.

By the Main Theorem (i) φ is positive definite, and (ii) being fulfilled with $C = 1, \gamma \equiv 1, \varphi$ is α–bounded with

$$\alpha(s) = \inf\left\{\left(\prod\beta(x_i)\right)^{-1} \mid \sum t(x_i) = s\right\} \quad \text{for} \quad s \in S$$

and

$$\alpha(\zeta) = 0 \quad (\text{since } \ \beta'(0) > 1)\,.$$

Also, the measure μ' representing φ (on S') concentrates on $\{\sigma' \in (S')^\alpha \mid \beta' \cdot (\sigma' \circ t')$ is positive definite on $R\} =: V'$ and each $\sigma' \in V'$ is non-negative since R is idempotent, so $\sigma := \sigma' \,|_S \geq 0$. Let μ be the image of μ' under $\sigma' \longmapsto \sigma' \,|_S$, and $V := \{\sigma' \,|_S \mid \sigma' \in V'\}$. Then

$$\varphi(s) = \int_{V'} \sigma'(s) d\mu'(\sigma') = \int_V \sigma(s) d\mu(\sigma)$$

for $s \in S$, and

$$\begin{aligned} P(x_1, \dots, x_n) &= \Phi\left(1_{\{x_1\}}, 1_{\{x_2\}}, \dots, 1_{\{x_n\}}, 1, 1, \dots\right) = \\ &= \prod_{i=1}^n \beta(x_i) \cdot \varphi\left(\sum_{i=1}^n t(x_i)\right) = \\ &= \int_V \prod_{i=1}^n \beta(x_i)\sigma(t(x_i)) d\mu(\sigma)\,, \end{aligned}$$

leading to

$$1 = \sum_{x_1, \dots, x_n \in \mathcal{X}} P(x_1, \dots, x_n) = \int \left[\sum_{x \in \mathcal{X}} \beta(x)\sigma(t(x))\right]^n d\mu(\sigma)$$

for all $n \in \mathbb{N}$, which shows that μ is in fact concentrated on W. □

The technicalities in the above proof were perhaps slightly more complicated than expected, but then calculations with (Fourier, Laplace) transforms are often easier than those with the distributions themselves ... The following examples will show, however, that the result is easy to apply.

Example 4.1. The original De Finetti theorem: here $\mathcal{X} = \{0, 1\}, S = \{(k, n) \in \mathbb{N}_0^2 \mid k \leq n\}$ (cf. Section 1), $t(x) = (x, 1), \beta \equiv 1$. A general non–negative character on S has the form $\sigma(k, n) = p^k q^{n-k}$ with $p, q \geq 0$. The condition $\sigma \circ t \in M_+^1(\mathcal{X})$ translates into

$$\sigma(t(0)) + \sigma(t(1)) = \sigma(0, 1) + \sigma(1, 1) = q + p = 1$$

which gives the result.

Example 4.2. A slight extension of 4.1. We consider $\mathcal{X} = \{0, 1, 2, \dots, k\}$, where $k \in \mathbb{N}$. Let again $P \in M_+^1(\mathcal{X}^\infty)$ fulfill

$$P(x_1, \dots, x_n) = \varphi_n\left(\sum_{i=1}^n x_i\right) = \varphi\left(\sum_{i=1}^n (x_i, 1)\right)$$

as before. Then

$$P = \int_0^1 \kappa_p^\infty \, d\mu(p)$$

with

$$\kappa_p(\{j\}) = p^j q^{k-j}, \quad q = q(p) \quad \text{from} \quad p^k + p^{k-1}q + \ldots + pq^{k-1} + q^k = 1\,.$$

Example 4.3. A further "extension": $\mathcal{X} = \mathbb{N}_0, P$ as before. Then

$$P = \int_{]0,1]} \gamma_a^\infty \, d\mu(a)$$

γ_a denoting the geometric distribution with parameter a, i.e. $\gamma_a(\{k\}) = a(1-a)^k$.

Example 4.4. $\mathcal{X} = \mathbb{N}_0$ as before, $P \in M_+^1(\mathcal{X}^\infty)$. Then

$$P(x_1, \ldots, x_n) = \frac{1}{\prod_{i=1}^n x_i!} \cdot \varphi_n\left(\sum_{i=1}^n x_i\right)$$

iff

$$P = \int_0^\infty \pi_\lambda^\infty \, d\mu(\lambda)$$

where π_λ denotes the Poisson distribution with parameter λ. Here we have for the first time the non–trivial function $\beta(x) = 1/x!$. The choice $\beta(x) = 1/(x+1)$ leads instead to mixtures of

$$\kappa_u(\{x\}) := \frac{1}{-\log(1-u)} \cdot \frac{u^{x+1}}{x+1} \quad (0 < u < 1) \quad \text{and} \quad \kappa_0 = \varepsilon_0\,,$$

and $\beta(x) = \binom{x+r-1}{r-1}$ would lead to negative binomials.

6 – Abstract results

De Finetti's original result dealt with $\{0,1\}$–valued random variables, and was generalized by him a few years later to the real-valued case. In 1955 a considerable further extension to arbitrary compact Hausdorff spaces was presented:

Theorem (Hewitt–Savage). *Let $\mathcal{X}$ be a compact Hausdorff space. Then $P \in M_+^1(\mathcal{X}^\infty)$ is exchangeable iff*

$$P = \int \kappa^\infty \, d\mu(\kappa)$$

for some $\mu \in M_+^1(M_+^1(\mathcal{X}))$.

[Here $M_+^1(\mathcal{X})$ is by definition the set of all Radon probability measures on $\mathcal{X}$; and $M_+^1(\mathcal{X})$ is given the usual weak topology in which it is again compact.]

In the following proof we'll use the notion of the *free abelian semigroup* (without involution) over a set A, denoted $\mathbb{N}_0^{(A)}$, and consisting of all functions $s : A \longrightarrow \mathbb{N}_0$ such that $\{s > 0\}$ is finite, with the usual addition.

A character σ on $\mathbb{N}_0^{(A)}$ can be identified with a function $\tau : A \longrightarrow \mathbb{R}$ via $\tau(a) = \sigma(\delta_a)$, where $\delta_a := 1_{\{a\}} \in \mathbb{N}_0^{(A)}$.

Proof. Let $A := \{f : \mathcal{X} \longrightarrow [0,1] \mid f \text{ is continuous}\}$. Then for any finite collection of $f_1, f_2, \ldots \in A$

$$E\left[\prod f_j(X_j)\right] = \varphi\left(\sum \delta_{f_j}\right)$$

with φ being defined on $\mathbb{N}_0^{(A)}$. By the Main Theorem, φ is c.p.d. (and bounded), so

$$\varphi\left(\sum \delta_{f_j}\right) = \int \prod \tau(f_j) d\mu(\tau)$$

for some $\mu \in M_+^1([0,1]^A)$.

An easy argument shows μ to be concentrated on

$$T := \{\tau : A \longrightarrow [0,1] \mid \tau(1) = 1, \tau \quad \text{finitely additive}\}$$

(cf. [6], Theorem 2) , and each $\tau \in T$ extends uniquely to a positive linear functional on $C(\mathcal{X})$, i.e. τ can be identified with a Radon probability measure on $\mathcal{X}$. Inserting this above gives the desired result. □

Remark 1. If $\mathcal{X}$ was just a measurable space then with $A := \{f : \mathcal{X} \longrightarrow [0,1] \mid f \text{ measurable}\}$ one obtains

$$E\left[\prod f_j(X_j)\right] = \int \prod \tau(f_j) d\mu(\tau),\ f_j \in A\,,$$

with $\mu \in M_+^1(T)$ and

$$T := \{\tau : A \longrightarrow [0,1] \mid \tau(1) = 1, \tau \quad \text{additive}\}$$

which is a "weak" form of a general De Finetti type result.

Remark 2. As noted above, the Berg/Maserick theorem is an essential ingredient in the proof of the Main Theorem. It can however also be deduced from it: if $\varphi : S \longrightarrow \mathbb{C}$ is p.d. and α–bounded then $\Phi(s_1, s_2, \ldots) := \varphi(\sum s_j)$ is p.d., and

$$\mid \Phi(s_1, s_2, \ldots) \mid \leq C \cdot \prod \alpha(s_j)\,.$$

With $R = S$, $t = id_S$ and $\beta \equiv 1$ the set W in the Main Theorem reduces to S^α.

Remark 3. The Main Theorem can be looked at as a result on *exchangeable p.d. functions* (here for simplicity we assume S without involution): let $\Phi : R^{(\infty)} \longrightarrow \mathbb{R}$ be p.d. and exchangeable. This leads to a factorization of the form

$$\Phi(r_1, r_2, \dots) = \varphi\left(\sum \delta_{r_j}\right)$$

with $\varphi : \mathbb{N}_0^{(R\setminus\{0\})} \longrightarrow \mathbb{R}$ (and $\delta_0 := 0$). Then φ is p.d., and if $|\Phi(r_1, r_2, \dots)| \leq C \cdot \prod \gamma(r_j)$ for some function $\gamma : R \longrightarrow \mathbb{R}_+, \gamma(0) = 1$, and some $C > 0$, the function φ is α-bounded with $\alpha(\sum \delta_{r_j}) := \prod \gamma(r_j)$ – so α is even a character. We get

$$\varphi\left(\sum \delta_{r_j}\right) = \int \sigma\left(\sum \delta_{r_j}\right) d\mu(\sigma)$$

where μ is a Radon measure on all characters σ of $\mathbb{N}_0^{(R\setminus\{0\})}$ with $|\sigma| \leq \alpha$. Since such a σ can be identified with the function τ on $R \setminus \{0\}$ given by $\tau(r) := \sigma(\delta_r)$, completed by $\tau(0) := 1$, we see that μ can be considered as a measure on

$$W := \{\tau \in \mathcal{P}_1(R) \mid |\tau| \leq \gamma\},$$

leading to

$$\Phi(r_1, r_2, \dots) = \int_W \prod \tau(r_j) d\mu(\tau),$$

a mixture of tensor powers of p.d. functions on R.

Note that in the special case where $\Phi(r_1, r_2, \dots) = \varphi(\sum r_j)$ depends on the sum of the entries, the function $\varphi : R \longrightarrow \mathbb{R}$ is automatically p.d., so that if furthermore φ is a *moment function* (i.e. a mixture of characters) Φ would be the corresponding mixture of infinite tensor powers of characters.

7 – A different point of view

Let's take another look at the Main Theorem (with $\beta \equiv 1$) :

$$\Phi(r_1, r_2, \dots) = \varphi\left(\sum t(r_j)\right)$$

and the conclusion Φ p.d. $\Longrightarrow \varphi$ p.d.

Put $U := R^{(\infty)}, \psi(r_1, r_2, \dots) := \sum t(r_j)$, then $\psi : U \longrightarrow S$ is onto and the theorem says: $\boxed{\varphi \circ \psi \text{ p.d.} \implies \varphi \text{ p.d.}}$

What is the crucial property of ψ enabling this conclusion?

The answer looks complicated:

$$\left|\begin{array}{l} \forall \text{ finite subsets } \{s_1, \dots, s_n\} \subseteq S \text{ and } \{u_1, \dots, u_m\} \subseteq U \text{ and} \\ \forall\, N \in \mathbb{N} \quad \exists\, \{u_{jp\alpha} \mid j \leq n, p \leq m, \alpha \leq N\} \subseteq U \text{ such that} \\ \psi(u_{jp\alpha} + u^-_{kq\beta}) = s_j + s^-_k + \psi(u_p + u^-_q) \text{ for } (j,p,\alpha) \neq (k,q,\beta) \end{array}\right.$$

If this is fulfilled, and $\psi(0) = 0$, we call ψ *strongly almost additive.*

This holds for example if ψ is a homomorphism and onto, but this case is not too interesting.

In this more general framework we shall deal only with bounded functions, being no restriction for the applications we have in mind.

THEOREM. *Let U, S be two semigroups, $\psi : U \longrightarrow S$ be strongly almost additive, and $\varphi : S \longrightarrow \mathbb{C}$ bounded. Then*

$$\varphi \circ \psi \ p.d. \implies \varphi \ p.d.$$

and φ is in fact a mixture of characters in

$$\hat{S}_\psi := \{\sigma \in \hat{S} \mid \sigma \circ \psi \ p.d.\}$$

(n.b.: a compact subsemigroup of $\hat{S}$).

Furthermore:

$$\{\varphi : S \longrightarrow \mathbb{C} \mid \varphi \text{ bounded}, \varphi(0) = 1, \varphi \circ \psi \ p.d.\}$$

is a Bauer simplex with $\hat{S}_\psi$ as extreme points.

Cf. [7], Theorem 1, where a slightly more general result is shown.

We shall apply this theorem to socalled exchangeable random partitions of the positive integers.

$V = \{v_1, v_2, \dots\}$ is a *partition* of $\mathbb{N} :\iff v_j \neq \emptyset, v_j \cap v_k = \emptyset$ for $j \neq k$, and $\bigcup_j v_j = \mathbb{N}$.

Examples are $\{\{i\} \mid i \in \mathbb{N}\}$ or $\{\mathbb{N}\}$, the two "extreme" partitions of $\mathbb{N}$.

Let $\mathcal{P}$ denote the set of all partitions of $\mathbb{N}$. Any $V \in \mathcal{P}$ can be identified with the equivalence relation $E(V) := \bigcup_{v \in V} v \times v \subseteq \mathbb{N}^2$ or with $1_{E(V)} \in \{0,1\}^{\mathbb{N}^2}$, this last identification defining the (natural) topology on $\mathcal{P}$, turning it into a compact metric space.

For $A \subseteq \mathbb{N}$ and $V \in \mathcal{P}$ we write

$$A \sqsubseteq V :\iff \exists\, v \in V \quad \text{with} \quad A \subseteq v$$

(that is: A is not separated by the classes of V).

For $U, V \in \mathcal{P}$ we define

$$U \leq V :\iff u \sqsubseteq V \quad \forall\, u \in U \quad [\iff E(U) \subseteq E(V)] \ .$$

Every subset of $\mathcal{P}$ has a unique minimal element w.r. to "$\leq$", and for a family $\mathcal{A}$ of subsets of $\mathbb{N}$ there is a smallest $W \in \mathcal{P}$ such that $A \sqsubseteq W$ for each $A \in \mathcal{A}$. In the particular case of $\mathcal{A} = U \cup V$ for $U, V \in \mathcal{P}$ we write $U \vee V$ for this minimum, and call it (of course) their *maximum.*

The order intervals $P_U := \{W \in \mathcal{P} \mid U \leq W\}$ fulfill $P_U \cap P_V = P_{U\vee V}$. For $U \in \mathcal{P}$ the classes $u \in U$ with $|u| \geq 2$ are called *non-trivial*, their union $\langle U\rangle$ is called the *support* of U. Obviously $\langle U \vee V\rangle \subseteq \langle U\rangle \cup \langle V\rangle$, so that

$$\mathcal{U} := \{U \in \mathcal{P} \mid \langle U\rangle \text{ is finite}\}$$

is a subsemigroup w.r. to "$\vee$ ", with neutral element $U_0 = \{\{j\} \mid j \in \mathbb{N}\}$. The order intervals P_U for $U \in \mathcal{U}$ are clopen and generate the Borel sets of $\mathcal{P}$. Probability measures on $\mathcal{P}$ will be called *random partitions.*

Theorem. *$\varphi : \mathcal{U} \longrightarrow \mathbb{R}$ is p.d. and normalized (i.e. $\varphi(U_0) = 1$) $\iff \exists$ (unique) random partition $\mu \in M_+^1(\mathcal{P})$ with*

$$\varphi(U) = \mu(P_U) \quad \forall\, U \in \mathcal{U}$$

[*cf.* [4], *Theorem* 1].

Note that the easy direction "$\Longleftarrow$" follows immediately from

$$\sum_{j,k=1}^{n} c_j c_k \varphi(U_j \vee U_k) = \int \left(\sum_{j=1}^{n} c_j 1_{P_{U_j}}\right)^2 d\mu \;\geq 0\,.$$

A permutation π of $\mathbb{N}$ induces $\overline{\pi} : \mathcal{P} \longrightarrow \mathcal{P}$, $\overline{\pi}(V) := \{\pi(v) \mid v \in V\}$, and $\overline{\pi}$ is continuous. π is *finite* iff $\{i \in \mathbb{N} \mid \pi(i) \neq i\}$ is finite.

Definition. $\mu \in M_+^1(\mathcal{P})$ is *exchangeable* $:\Longleftrightarrow \mu^{\overline{\pi}} = \mu\ \forall$ finite π.

Now $\mu^{\overline{\pi}}(P_U) = \mu(P_{\overline{\pi^{-1}(U)}})$, so μ is exchangeable iff

$$\mu(P_U) = \mu(P_V)$$

$\forall\, U, V \in \mathcal{U}$ with $|\{u \in U \mid |u| = k\}) = |\{v \in V| \, |v| = k\}|$ for $k = 2, 3, \ldots$ iff $\mu(P_U) = \varphi \circ g(U)$ for some φ defined on the semigroup

$$S := \mathbb{N}_0^{(\{2,3,\ldots\})}\,,$$

with $g(U) := \sum_{\substack{u\in U\\ |u|\geq 2}} \delta_{|u|}$.

This function $g : \mathcal{U} \longrightarrow S$ is in fact strongly almost additive, cf. [4], Lemma 5.

Theorem (Kingman). *$M_+^{1,e}(\mathcal{P}) := \{\mu \in M_+^1(\mathcal{P}) | \mu$ exchangeable$\}$ is a Bauer simplex whose extreme points are precisely those μ for which*

$$\mu(P_U) = \sigma(g(U)), U \in \mathcal{U} \quad \textit{with} \quad \sigma \in \hat{S}_+\,.$$

Such a character σ is given by a sequence $(t_2, t_3, \dots)$ in $[0,1]$, and we see that

$$t_n = \mu(P_{\{\{1,\dots,n\},\{n+1\},\{n+2\},\dots\}}), \quad n \geq 2$$

is the μ–probability for $\{1,\dots,n\}$ not getting separated. For general $U \in \mathcal{U}$ the multiplicativity of σ is reflected in a certain pattern of independence:

$$\mu(P_U) = \prod_{\substack{u \in U \\ |u| \geq 2}} t_{|u|}\,.$$

Kingman ([5]) showed that there exists a sequence $x = (x_1, x_2, \dots)$ with $x_i \geq 0, \sum x_i \leq 1$, such that

$$t_n = \sum_{i=1}^{\infty} x_i^n \quad \text{for} \quad n = 2, 3, \dots$$

is the associated sequence of power sums.

Using x, there is in fact a natural way to describe this distribution μ : put $x_0 := 1 - \sum_{i=1}^{\infty} x_i$ and let $X_1, X_2, \dots$ be iid with $P(X_1 = i) = x_i, i \geq 0$. Then

$$G := \{\{j \in \mathbb{N} \mid X_j = c\} \mid c \in \mathbb{N}\} \cup \{\{i\} \mid X_i = 0\} \setminus \{\emptyset\}$$

is $\mathcal{P}$–valued and has μ as its distribution.

Acknowledgements

The author is very grateful to the organizing committee of the *Bruno de Finetti Centenary Conference* (Rome, November 2006), in particular to Professor Fabio Spizzichino, for the invitation to give at this highly interesting congress, taking place in the fascinating "Eternal City", a talk of which the present paper is an elaborate version.

Many thanks are due also to the referee(s) for a particularly thorough revision of my first draft, resulting in many improvements.

REFERENCES

[1] C. Berg – J. P. R. Christensen – P. Ressel: *Harmonic Analysis on Semigroups*, Springer–Verlag, Berlin/Heidelberg/New York, 1984.

[2] C. Berg – P. H. Maserick: *Exponentially bounded positive definite functions*, Illinois J. Math., **28** (1984), 162–179.

[3] B. De Finetti: *Funzione caratteristica di un fenomeno aleatorio*, Memorie della R. Accademia Nazionale dei Lincei, **V** (IV) (1930), 86–133.

[4] U. Hirth – P. Ressel: *Random partitions by semigroup methods*, Semigroup Forum, **59** (1999), 126–140.

[5] J. F. C. Kingman: *The representation of partition structures*, J. London Math. Soc., **18** (1978), 374–380.

[6] P. Ressel: *De Finetti–type theorems: an analytical approach*, Ann. Prob., **13** (1985), 898–922.

[7] P. Ressel: *Integral representations for distributions of symmetric stochastic processes*, Prob. Th. Rel. Fields, **79** (1988), 451–467.

[8] I. J. Schoenberg: *Metric spaces and positive definite functions*, Trans. Amer. Math. Soc., **44** (1938), 522–536.

Lavoro pervenuto alla redazione il 8 ottobre 2007
ed accettato per la pubblicazione il 12 ottobre 2007.
Bozze licenziate il 9 aprile 2008

Indirizzo dell'Autore:

Paul Ressel – Universität Eichstätt – Osten str. 26-28 D 85071 Eichstatt, Germany
E-mail: paul.ressel@ku-eichstaett.de