Ann. Inst. Statist. Math.
Vol. 45, No. 2, 367–381 (1993)

GENERALIZED MULTIVARIATE HERMITE DISTRIBUTIONS AND RELATED POINT PROCESSES

R. K. MILNE[1] AND M. WESTCOTT[2]

[1] *Department of Mathematics, University of Western Australia, Nedlands 6009, Australia*
[2] *Division of Mathematics and Statistics, CSIRO, P.O. Box 1965, Canberra City 2601, Australia*

(Received July 15, 1990; revised June 19, 1992)

Abstract. This paper is primarily concerned with the problem of characterizing those functions of the form

$$G(\boldsymbol{z}) = \exp\left\{ \sum_{0 \le \boldsymbol{k}'\boldsymbol{1} \le m} a_{\boldsymbol{k}} (\boldsymbol{z}^{\boldsymbol{k}} - 1) \right\},$$

where $\boldsymbol{z} = [z_1, \dots, z_n]'$, which are probability generating functions. The corresponding distributions are called generalized multivariate Hermite distributions. Use is made of results of Cuppens (1975), with particular interest attaching to the possibility of some of the coefficients $a_{\boldsymbol{k}}$ being negative.

The paper goes on to discuss related results for point processes. The point process analogue of the above characterization problem was raised by Milne and Westcott (1972). This problem is not solved but relevant examples are presented. Ammann and Thall (1977) and Waymire and Gupta (1983) have established a related characterization result for certain infinitely divisible point processes. Their results are considered from a probabilistic viewpoint.

Key words and phrases: Hermite distribution, generalized multivariate Hermite distribution, point process, probability generating function.

1. Introduction

Consider a probability generating function (p.g.fn) of the form

$$G(\boldsymbol{z}) = \exp\{P(\boldsymbol{z}) - P(\boldsymbol{1})\} \tag{1.1}$$

where $P(\boldsymbol{z})$ is a polynomial in $\boldsymbol{z} \in \boldsymbol{R}^n$, and $\boldsymbol{1} = [1, \dots, 1]'$. When all the coefficients of $P(\boldsymbol{z})$ are non-negative, $G(\boldsymbol{z})$ is a p.g.fn; in fact it is the p.g.fn of a compound Poisson distribution. (We employ here the terminology of Feller ((1968),

p. 288 ff.); other authors, e.g. Kemp and Kemp (1965), call such distributions generalized Poisson. Both terminologies are widespread in the literature.) As such, it is infinitely divisible i.e. $[G(\boldsymbol{z})]^{1/r}$ is a p.g.fn for each positive integer r.

It is however possible for $G(\boldsymbol{z})$ to be a p.g.fn, though not an infinitely divisible p.g.fn, even when some of the coefficients in $P(\boldsymbol{z})$ are negative. Lévy ((1937), p. 236), quoted by Lukacs ((1970), p. 251) amongst others, has given an example in which $n = 1$ and P is a quartic polynomial with negative coefficient for the quadratic term. Cuppens ((1975), Appendix B), generalizing work of Lévy (1937) for the case $n = 1$, gave necessary and sufficient conditions under which $G(\boldsymbol{z})$ will be a p.g.fn when

$$P(\boldsymbol{z}) = \sum_{0 \le \boldsymbol{k} \le \boldsymbol{p}} a_{\boldsymbol{k}} \boldsymbol{z}^{\boldsymbol{k}},$$

where $\boldsymbol{z} \in \boldsymbol{R}^n$, $\boldsymbol{k}, \boldsymbol{p} \in \boldsymbol{N}^n$ with $\boldsymbol{N} = \{0, 1, 2, \ldots\}$, n is a positive integer, and $\boldsymbol{z}^{\boldsymbol{k}} = z_1^{k_1} z_2^{k_2} \cdots z_n^{k_n}$. (Usually our notation will follow, or be close to, that of Cuppens.)

A primary concern of this paper is with p.g.fns of the form (1.1) when

$$P(\boldsymbol{z}) = \sum_{0 \le \boldsymbol{k}'\boldsymbol{1} \le m} a_{\boldsymbol{k}} \boldsymbol{z}^{\boldsymbol{k}}, \tag{1.2}$$

with $m \in \boldsymbol{N}$ i.e. with p.g.fns of the form

$$G(\boldsymbol{z}) = \exp\left\{ \sum_{0 \le \boldsymbol{k}'\boldsymbol{1} \le m} a_{\boldsymbol{k}} (\boldsymbol{z}^{\boldsymbol{k}} - 1) \right\}. \tag{1.3}$$

Specifically, we are concerned with conditions under which some of the coefficients $a_{\boldsymbol{k}}$ may be negative yet $G(\boldsymbol{z})$, given by (1.3), remains a p.g.fn. In this situation we aim to reduce Cuppens' conditions to a more concrete form.

A related problem, outside the scope of the present work, is to find conditions on the coefficients $a_{\boldsymbol{k}}$ which ensure that $G(\boldsymbol{z})$ given by (1.3) is a p.g.fn in the case $m = \infty$. This problem is analogous to the problem ostensibly tackled by Waymire and Gupta (1983) for functionals of the form (2.2); see Section 2 and Section 6 of the present paper. In fact, Waymire and Gupta in effect assumed infinite divisibility. Correspondingly, for functions of the form (1.3) with $m = \infty$ to be infinitely divisible p.g.fns it is necessary and sufficient that the coefficients $a_{\boldsymbol{k}}$ be non-negative. For some univariate results when negative a's are permitted, see the remarks following Theorem 3.1. We shall not attempt to study the multivariate case further in this paper.

Notice that instead of considering $G(\boldsymbol{z})$ given by (1.3) we can without loss of generality focus on

$$F(\boldsymbol{z}) = \exp\{P(\boldsymbol{z})\}, \tag{1.4}$$

with $P(\boldsymbol{z})$ given by (1.2), and seek conditions for non-negativity of all coefficients in the power series expansion of $F(\boldsymbol{z})$ about $\boldsymbol{z} = \boldsymbol{0}$. Observe that in either setting we can take

$$P(\boldsymbol{z}) = \sum_{1 \le \boldsymbol{k}'\boldsymbol{1} \le m} a_{\boldsymbol{k}} \boldsymbol{z}^{\boldsymbol{k}}. \tag{1.5}$$

We call the distribution corresponding to a p.g.fn $G(\boldsymbol{z})$ of the form (1.3) a *generalized multivariate Hermite distribution* or, more particularly, an *m-th order n-variate Hermite distribution*. Any m-th order univariate Hermite distribution will be referred to as a *generalized Hermite distribution*.

Example 1. *First order univariate*: (univariate) Poisson distributions.

Example 2. *First order n-variate*: infinitely divisible n-variate Poisson distributions cf. Teicher (1954).

Example 3. *Second order univariate*: (univariate) Hermite distributions cf. Kemp and Kemp (1965).

Example 4. *Second order bivariate*: these are the distributions H_5 of Kemp and Papageorgiou (1982). Here we write

$$G(\boldsymbol{z}) = \exp\{a_{10}(z_1 - 1) + a_{20}(z_1^2 - 1) \\ + a_{01}(z_2 - 1) + a_{02}(z_1^2 - 1) + a_{11}(z_1 z_2 - 1)\}.$$

Example 5. *Second order n-variate*: finite-dimensional distributions of Gauss-Poisson processes cf. Milne and Westcott (1972) and Section 2 below.

Example 6. *Fourth order bivariate*: the distributions H_8 of Kemp and Papageorgiou (1982) are special cases of this form.

Example 7. *m-th order univariate*: Gupta and Jain (1974) have considered a special type of m-th order univariate Hermite distribution (only a_1 and a_m, in the form (1.5), are non-zero). They used the term 'generalized Hermite distribution' for this type.

Several authors have discussed the formal derivation of the Hermite distribution as a mixture, that is as the distribution of X where the conditional distribution of X given λ is Poisson with parameter λ, and λ is Normally distributed with mean μ and variance σ^2 where $\mu \gg \sigma^2$. Multivariate analogues of this result have been given for second order n-variate Hermite distributions; see for example Steyn (1976), and Kemp and Papageorgiou (1982). However, no higher order n-variate Hermite distribution can be derived by mixture starting from independent Poisson distributions. This is an easy consequence of a theorem of Marcinkiewicz (cf. Lukacs (1970), p. 213).

Our interest in generalized (multivariate) Hermite distributions began in a point process context, with Milne and Westcott (1972). Some background, related especially to this context, is presented in the following section. In Section 3 we address the characterization problem for p.g.fns of the form (1.3). Section 4 presents two examples of generalized multivariate Hermite distributions which exhibit certain extremes of possible behaviour and which were considered originally for the light they throw on the corresponding point process characterization problem. This problem, as originally posed by Milne and Westcott (1972), is not solved but relevant examples are given in Section 5. Finally, in Section 6, we discuss, from a probabilistic viewpoint, related characterization results obtained by Ammann and Thall (1977) and Waymire and Gupta (1983) for certain infinitely divisible point processes. The probabilistic significance of the conditions obtained by these authors appeared to be obscured by their largely analytic treatment.

2. Background

Milne and Westcott ((1972), p. 169) posed the problem of characterizing the possibly signed measures $H_k(\cdot)$ on $\boldsymbol{R}^k$, $k = 1, 2, \ldots$, such that the functional

$$G[\xi] = \exp\left\{\sum_{k=1}^{m} \int_{\boldsymbol{R}^k} \prod_{i=1}^{k} [\xi(t_i) - 1] H_k(dt_1 \cdots dt_k)\right\} \tag{2.1}$$

is the probability generating functional (p.g.fl) of a point process on $\boldsymbol{R}$. (For unfamiliar concepts and notation refer to the cited paper. We continue to deal, as there, with point processes on $\boldsymbol{R}$ although it is clear that the results hold for point processes on more general phase spaces.)

Milne and Westcott had already solved this problem in the case $m = 2$. The more general problem was ostensibly tackled by Ammann and Thall (1977, 1978) and by Waymire and Gupta (1983); they considered the characterization problem for functionals of the forms

$$G[\xi] = \exp\left\{\sum_{k=1}^{\infty} \int_{\boldsymbol{R}^k} \prod_{i=1}^{k} [\xi(t_i) - 1] H_k(dt_1 \cdots dt_k)\right\} \tag{2.2}$$

and

$$G[\xi] = \exp\left\{\sum_{k=1}^{\infty} \int_{\boldsymbol{R}^k} \left[\prod_{i=1}^{k} \xi(t_i) - 1\right] M_k(dt_1 \cdots dt_k)\right\}. \tag{2.3}$$

However, as we shall indicate in Section 6, they considered in effect only the problem of when $G[\xi]$, given by (2.2) or (2.3), is an *infinitely divisible* p.g.fl. (A p.g.fl $G[\xi]$ is infinitely divisible if $(G[\xi])^{1/r}$ is a p.g.fl for each positive integer r.) The main point made in Waymire and Gupta was that Corollary 3.2 of Ammann and Thall (1977) was incorrect and that a p.g.fl in the form (2.2) cannot necessarily be re-expressed in the form (2.3) because this involves a not necessarily permissible inversion of an infinite series.

The case $m = 2$ yields the so-called Gauss-Poisson processes (cf. Newman (1970)). For such a process, the joint p.g.fn $G_1(\boldsymbol{z})$, of the counts in pairwise disjoint bounded Borel sets $A_1, \ldots, A_n$, has the form

$$G_1(\boldsymbol{z}) = \exp\left\{\sum_{i=1}^{n} H_1(A_i)(z_i - 1) + \sum_{i=1}^{n}\sum_{j=1}^{n} H_2(A_i \times A_j)(z_i - 1)(z_j - 1)\right\}. \tag{2.4}$$

In this case the necessary and sufficient conditions for (2.1) to be a p.g.fl (cf. Theorem 1 of Milne and Westcott (1972)) imply that the resultant p.g.fl is infinitely divisible, or equivalently that each p.g.fn $G_1(\boldsymbol{z})$ as at (2.4) is infinitely divisible.

Although (2.4) exhibits $G_1(\boldsymbol{z})$ as the exponential of a polynomial in $\boldsymbol{z}-\mathbf{1}$, we can clearly (cf. Kemp and Papageorgiou (1982), in the case $m = n = 2$) express $G_1(\boldsymbol{z})$ as the exponential of a polynomial of the form (1.2) in $\boldsymbol{z}$. One advantage of the expression in terms of $\boldsymbol{z} - \mathbf{1}$ is that $\ln G_1(\boldsymbol{z})$ is then precisely the factorial cumulant generating function. For establishing conditions under which $G_1(\boldsymbol{z})$ is a p.g.fn in other than second order cases it appears to be more straightforward to work with the '$\boldsymbol{z}$ form'. Similar comments can be made about corresponding p.g.fl forms (2.1), (2.2) and (2.3); see Section 6.

A priori it does not seem obvious just what structure, if any, should be considered as 'natural' for the p.g.fn of a multivariate distribution all of whose univariate marginals are generalized Hermite distributions. For instance, in the bivariate case Kemp and Papageorgiou (1982) considered two models they called H_5 and H_8. We choose to regard (1.3) as the basic or 'natural' form for such a p.g.fn. This seems to us appropriate since the p.g.fns (2.4) for the counts of a Gauss-Poisson process in *pairwise disjoint* bounded Borel sets $A_1, \ldots, A_n$, as well as the p.g.fns that would arise correspondingly from (2.1) when this is a p.g.fl, have the structure of (1.3). Our view is further supported in the second order case ($m = 2$) by the observation that the corresponding distribution can be specified (uniquely) by its means, variances and covariances.

In the second order bivariate case, our chosen form reduces to the model H_5 of Kemp and Papageorgiou (1982). These authors had already noted that model H_5 is 'easier to handle and simpler to interpret' than model H_8, and also that model H_5 is special because of the property that if $Z = X + Y$, where the joint distribution of X and Y is that of model H_5, Z has a univariate Hermite distribution. An analogous property differentiates between corresponding multivariate versions of models H_5 and H_8. The essence of this difference can be described, in the context of point processes with p.g.fls of the form (2.1), as the difference between finite-dimensional distributions for counts in pairwise disjoint bounded Borel sets and those for counts in arbitrary bounded Borel sets. While these two sets of finite-dimensional distributions are mathematically equivalent, the former are arguably more basic since the latter are easily constructed from them.

3. Characterizing generalized multivariate Hermite distributions

For the present we restrict attention to the case where all coefficients in $P(\boldsymbol{z})$ given by (1.5) are non-zero. When $\boldsymbol{k} = [k_1, \ldots, k_n]'$ we will often write $a_{\boldsymbol{k}}$ as $a_{k_1 k_2 \cdots k_n}$ cf. Example 4. By direct methods we can establish the following result.

LEMMA 3.1. *In order for $G(\boldsymbol{z})$ given by (1.3) to be a p.g.fn, it is necessary that*

$$a_{10\cdots 0} > 0, \ldots, a_{0\cdots 01} > 0, \tag{3.1}$$

$$a_{\overline{m-1}0\cdots 0} > 0, \ldots, a_{0\cdots 0\overline{m-1}} > 0, \tag{3.2}$$

$$a_{m0\cdots 0} > 0, \ldots, a_{0\cdots 0m} > 0 \tag{3.3}$$

and

$$a_{\overline{m-1}10\cdots 0} > 0, \ldots, a_{0\cdots 01\overline{m-1}} > 0. \tag{3.4}$$

PROOF. To establish (3.1), (3.3), and (3.4) we can use an approach similar that of Kemp and Kemp (1965) and Milne and Westcott (1972). Here $G(\boldsymbol{z})$ is a positive, entire function with Taylor series about $\boldsymbol{z} = \boldsymbol{0}$ convergent for all $\boldsymbol{z}$. Since the coefficients in the expansion of $G(\boldsymbol{z})$ are all non-negative, then both $G(\boldsymbol{z})$ and $\ln G(\boldsymbol{z})$ are non-decreasing functions of each z_i when these are non-negative. Thus for $i = 1, 2, \ldots, n$ we must have

$$\frac{\partial}{\partial z_i} \ln G(\boldsymbol{z}) \geq 0 \qquad (\boldsymbol{z} \geq \boldsymbol{0}). \tag{3.5$_i$}$$

To establish $a_{10\cdots 0} > 0$ set all components of $\boldsymbol{z}$ except the first to zero in (3.5_1). Now, again using (3.5_1), let the first component of $\boldsymbol{z}$ tend to infinity while keeping the others fixed and deduce $a_{m0\cdots 0} > 0$. To obtain $a_{\overline{m-1}10\cdots 0} > 0$ use (3.5_2) with the first component of $\boldsymbol{z}$ tending to infinity and the remaining components zero. The other inequalities of (3.1), (3.3), and (3.4) follow similarly.

Notice that the inequalities of (3.1) and (3.3) could have been derived after passing first to appropriate univariate marginal p.g.fns. To establish the inequalities (3.2) it is simplest to pass directly to such marginals and then to employ a device due to Lévy (1937); see also Cuppens ((1975), p. 224). Consider the marginal distribution of the first component and set $G_1(z) = G(\boldsymbol{z})$, $P_1(z) = P(\boldsymbol{z})$, where $\boldsymbol{z} = [z, 0, \ldots, 0]' \in \boldsymbol{R}^n$. Abbreviate the coefficients $a_{10\cdots 0}, \ldots, a_{m0\cdots 0}$ to $a_1, \ldots, a_m$ and suppose that a_{m-1} is negative. Set $\rho = e^{2\pi i/m}$. Then $\rho^m = 1$ and $\rho^{m-1} \neq 1$. Thus $P_1(z) - P_1(\rho z)$ varies like $a_{m-1}(1 - \rho^{m-1})z^{m-1}$ as $z \to +\infty$ and so $\operatorname{Re}\{P_1(z) - P_1(\rho z)\} \to -\infty$ as $z \to +\infty$. It follows that

$$\left| \frac{G(\rho z)}{G(z)} \right| = |e^{P_1(\rho z) - P_1(z)}| \to +\infty$$

as $z \to +\infty$. This establishes a contradiction since $|G(\rho z)| \leq |G(z)|$ for z non-negative. Thus a_{m-1} i.e. $a_{\overline{m-1}0\cdots 0}$ must be positive. The remaining inequalities of (3.2) can be established similarly. □

The inequalities (3.1)–(3.4) are by no means the only necessary conditions that can be derived cf. Remark 4 following the main theorem. Nevertheless, as we shall show, these conditions are also sufficient subject only to the additional restriction that any negative coefficient be suitably small. To establish this result we refine a result due to Cuppens (1975).

For fixed $\boldsymbol{p} = [p_1, \ldots, p_n]' \in \boldsymbol{N}^n$ set $Q(z) = P(z^{p_1}, \ldots, z^{p_n})$, $z \in \boldsymbol{R}$ and write $Q(z) = \sum_l b_l z^l$, where l runs through the possible values of $\boldsymbol{k}'\boldsymbol{p}$ for the given $\boldsymbol{p}$. Since so far as the definition of Q is concerned the order of the arguments of P is immaterial, we can assume that $\boldsymbol{p}$ satisfies $p_1 \geq p_2 \geq \cdots \geq p_n$. Define $\mathcal{A}'_j = \{l : b_l > 0, l > j\}$ and $\mathcal{B}'_j = \{l : b_l > 0, l < j\}$. Cuppens' key conditions are:

CONDITION A. for any $b_j < 0$, $j \in (\boldsymbol{Z})\mathcal{A}'_j$ i.e. is a linear combination of elements of $\mathcal{A}'_j$ with integer coefficients;

CONDITION B. for any $b_j < 0$, $j \in (\boldsymbol{N})\mathcal{B}'_j$ i.e. is a linear combination of the elements of $\mathcal{B}'_j$ with non-negative integer coefficients.

Cuppens gives in addition some conditions equivalent to these. The following result is a restatement of a special case of Theorem B.3.2 of Cuppens (1975).

PROPOSITION 3.1. *In order for $G(\boldsymbol{z})$ given by (1.3) to be a p.g.fn it is necessary and sufficient that the coefficients, b_l, of $Q(z)$ satisfy Conditions A and B for all $\boldsymbol{p} \in \boldsymbol{N}^n$ and that the moduli of any negative $a_{\boldsymbol{k}}$ are sufficiently small.*

We now seek information on the effect of Conditions A and B in our situation. Consider those $a_{\boldsymbol{k}}$ with $\boldsymbol{k}'\boldsymbol{1} > 1$ i.e. $a_{\boldsymbol{k}}$ other than $a_{10\cdots0}, \dots, a_{0\cdots01}$. For given $\boldsymbol{p} \in \boldsymbol{N}^n$, form $\boldsymbol{k}'\boldsymbol{p}$. Now let $K = \{i \in \{1, 2, \dots, n\} : k_i > 0\}$ and define $S_{\boldsymbol{k}}$ as the subset of those

$$[1, 0, \dots, 0]\boldsymbol{p}, \dots, [0, \dots, 0, 1]\boldsymbol{p}$$

having, for $i \in K$, a 1 as the i-th element of their first (row) vector. Then $\boldsymbol{k}'\boldsymbol{p}$ is a linear combination of elements of $S_{\boldsymbol{k}}$ with positive integer coefficients, and each member of $S_{\boldsymbol{k}}$ is less than $\boldsymbol{k}'\boldsymbol{p}$. Further, any b_l with $l \in S_{\boldsymbol{k}}$ will be positive if (3.1) holds, provided only that any negative $a_{\boldsymbol{k}}$ are sufficiently small in modulus. Thus we have $\boldsymbol{k}'\boldsymbol{p} \in (\boldsymbol{N})\mathcal{B}'_{\boldsymbol{k}'\boldsymbol{p}}$, and this happens for any $\boldsymbol{p} \in \boldsymbol{N}^n$. So we have proved the following result.

LEMMA 3.2. *If (3.1) holds, Condition B places no constraints on the signs of any other coefficients $a_{\boldsymbol{k}}$, provided that the moduli of any negative $a_{\boldsymbol{k}}$ are suitably small.*

We can draw a very similar conclusion regarding Condition A; this is shown in the next lemma.

LEMMA 3.3. *If (3.2)–(3.4) hold, Condition A places no constraints on the signs of any other coefficients $a_{\boldsymbol{k}}$, provided that the moduli of any negative $a_{\boldsymbol{k}}$ are suitably small.*

PROOF. Consider any $a_{\boldsymbol{k}}$ with $\max_{1\le i\le n} k_i \le m-2$. For any $\boldsymbol{p} \in \boldsymbol{N}^n$, set $j = \boldsymbol{k}'\boldsymbol{p}$; recall we assume $p_1 \ge p_2 \ge \cdots \ge p_n$.

Case $\boldsymbol{k}'\boldsymbol{1} = m$. Define $M = \max\{i \in \{1, 2, \dots, n\} : k_i > 0\}$, and assume for the moment that $p_1 > p_M$. Then

$$[m, 0, \dots, 0]\boldsymbol{p}, [m-1, 1, \dots, 0]\boldsymbol{p}, \dots, [m-1, 0, \dots, 0, 1, 0, \dots, 0]\boldsymbol{p},$$

where the 1 in the last row vector is in the M-th position, all belong to $\mathcal{A}'_j$, since $p_1 > p_M$ and (3.3)–(3.4) hold. So $\mathcal{A}'_j$ contains $mp_1, (m-1)p_1+p_2, \dots, (m-1)p_1+p_M$, at least. Now form

$$k_2[(m-1)p_1+p_2] + \cdots + k_M[(m-1)p_1+p_M] - (m-k_1-1)[mp_1], \tag{3.6}$$

a linear combination of elements of $\mathcal{A}'_j$ with integer coefficients. Using $\boldsymbol{k}'\mathbf{1} = m$, it is easy to show that this combination equals $k_1 p_1 + \cdots + k_M p_M$, and hence j. So $j \in (\boldsymbol{Z})\mathcal{A}'_j$ for any such j, and hence for any such j with $b_j < 0$, provided $p_1 > p_M$. If $p_1 = p_M$, $\mathcal{A}'_j$ is empty since there is no b_l with $l > j$. But this does not matter, since b_j is positive if any negative $a_{\boldsymbol{k}}$ are sufficiently small. Thus the conclusion holds for all $\boldsymbol{p} \in \boldsymbol{N}^n$.

Case $\boldsymbol{k}'\mathbf{1} < m$. We now have $[m-1, 0, \ldots, 0]\boldsymbol{p} = (m-1)p_1 \in \mathcal{A}'_j$ also. Suppose $\boldsymbol{k}'\mathbf{1} = m - d$, $d > 0$. Add the term $d[(m-1)p_1]$ to the linear combination (3.6). Again, elementary algebra shows that the combination equals j, and hence the same conclusion as above holds for all $\boldsymbol{p} \in \boldsymbol{N}^n$.

Collecting together the results of the proposition and the three lemmas, we obtain our main result.

THEOREM 3.1. *In order for $G(\boldsymbol{z})$, given by (1.3) with $a_{\boldsymbol{k}} \neq \mathbf{0}$ for those $\boldsymbol{k}$ satisfying $\boldsymbol{k}'\mathbf{1} = 1$ or $\max_{1 \le i \le n} k_i \geq m-1$, to be a p.g.fn it is necessary and sufficient that these $a_{\boldsymbol{k}}$ be positive, and that the moduli of any negative coefficients among the remaining $a_{\boldsymbol{k}}$ be suitably small.*

Remark 1. In the general univariate case with

$$G(z) = \exp\{a_1(z-1) + a_2(z^2-1) + \cdots + a_{m-1}(z^{m-1}-1) + a_m(z^m-1)\}$$

where at least a_1, a_{m-1}, and a_m are non-zero, it follows from Lemma 3.1 that $a_1 > 0$, $a_{m-1} > 0$, and $a_m > 0$ are necessary conditions for G to be a p.g.fn. Potentially, in view of Theorem 3.1, any of the remaining coefficients could be negative provided only that they are suitably small (cf. Lévy (1937), p. 263).

Remark 2. To treat other cases where some of the a's are zero is much more complicated. For example, suppose $a_t \neq 0$ and that $a_{t+1} = \cdots = a_{m-1} = 0$, where $t < m-1$. The approach used by Cuppens ((1975), p. 224) to prove necessity in Theorem B.1.1, cf. our proof of (3.2) in Lemma 3.1, can be used to deduce that $a_t > 0$ provided t and m are relatively prime; this makes use of Euclid's algorithm. Notice that Cuppens treats the (multivariate) case without the restriction that the a's be non-zero.

Remark 3. Lévy ((1937), p. 263) pointed out that the fourth order (univariate) case is the simplest admitting a negative coefficient: for the function

$$G(z) = \exp\{a(z-1) - b(z^2-1) + c(z^3-1) + d(z^4-1)\} \tag{3.7}$$

with a, b, c, d non-zero, to be a p.g.fn it is necessary that $a > 0$, $c > 0$, and $d > 0$ and sufficient that these inequalities hold and b be suitably small. In this case one can obtain four inequalities that constrain b in terms of the other coefficients (cf. van Harn (1978), p. 84). In particular $G(z)$ given by (3.7) is a p.g.fn when

$a = c = d = 1$ and $0 < b \leq 1/4$ (cf. Lukacs (1970), p. 251). If $c = 0$ then b is necessarily positive.

Remark 4. It is of course possible to write down many other necessary conditions. To see this in the univariate case consider $G(z)$ given in the form

$$G(z) = \exp\{\alpha_1(z-1) + \alpha_2(z-1)^2 + \cdots + \alpha_{m-1}(z-1)^{m-1} + \alpha_m(z-1)^m\},$$

where $\alpha_m \neq 0$. The necessary conditions $a_1 > 0$, $a_{m-1} > 0$, and $a_m > 0$ mentioned in Remark 1 above are equivalent to

$$\text{(3.8)} \qquad \alpha_1 - 2\alpha_2 + \cdots + (-1)^{m-1} m\alpha_m > 0, \qquad \alpha_{m-1} - m\alpha_m > 0,$$

and $\alpha_m > 0$. However, α_1 must be positive, because it is the mean of a non-degenerate distribution on $\boldsymbol{N}$; this particular condition does not follow directly from (3.8).

Remark 5. Kemp and Papageorgiou ((1982), p. 273) ask whether their model H_8 can have $a_5 < 0$ (in their notation) or $a_{11} < 0$ (in our notation) and still be a p.g.fn. We assert that this is possible; it can be established by methods similar to those illustrated in Example 9 of the following section.

Remark 6. For the univariate case with $m = \infty$ but possibly negative a's, we note the following results.

(i) It is possible that $\liminf_{n\to\infty} a_n < 0$, so that infinitely many negative a's are allowed. For example, if

$$P(z) = 1 + z + \cdots + z^{N-1} - bz^N + z^{N+1} + \cdots + z^{2N-1} - bz^{2N} + z^{2N+1} + \cdots,$$

it can be shown, as in Lukacs ((1970) p. 251), that, if $N \geq 3$, there is a range of positive b such that $G(z)$ is a p.g.fn.

(ii) Clearly, all a's negative from some point on is not permissible.

(iii) Condition B is still necessary.

(iv) Some of the equivalences for Condition A referred to just before the proposition break down if m is infinite.

(v) If all the a's from some point on are nonnegative, then for $G(z)$ to be a p.g.fn it is sufficient that, for some finite m_0, Conditions A and B are satisfied for the finite polynomial $P_{m_0}(z) = \sum_{k=1}^{m_0} a_k z^k$. The necessity of Condition B follows from (iii).

4. Further examples

Theorem 3.1 shows that a generalized multivariate Hermite distribution could have a substantial proportion of its defining coefficients a_k negative, provided of course that these are small enough in modulus. Of particular interest for the point process problem which motivated our enquiry is whether this constraint becomes more constraining as n, the number of variates, increases; that is, does the possibility of negative a_k disappear as $n \to \infty$? The next example shows that this behaviour is possible. A further example shows that this behaviour need not always occur.

Example 8. Consider an m-th order $(2n)$-variate Hermite distribution with $m, n \in \boldsymbol{N}$, $m \geq 3$ and $a_k \equiv 1$, except that

$$a_{110\cdots 0} = a_{1010\cdots 0} = \cdots = a_{0\cdots 011} = -b \qquad (b > 0),$$
$$a_{\boldsymbol{k}} = 0 \quad \text{for all} \quad \boldsymbol{k} \quad \text{with} \quad \max_{1\leq i\leq 2n} k_i = 1, \quad \text{and} \quad \boldsymbol{k}'\boldsymbol{1} \geq 3.$$

Theorem 3.1 guarantees that this specifies a p.g.fn provided b is sufficiently small.

Now consider the coefficient, $\psi \equiv \psi_n(b)$ say, of $z_1 z_2 \cdots z_{2n}$ in $F(\boldsymbol{z})$ given by (1.4). Clearly, the powers of $P(\boldsymbol{z})$ which contribute to ψ are those between n and $2n$, and a straightforward combinatorial argument shows that the coefficient of $z_1 z_2 \cdots z_{2n}$ in $[P(\boldsymbol{z})]^{2n-j}$ $(j = 0, 1, \ldots, n)$ is

$$\left(-\frac{1}{2}b\right)^j (2n)! \binom{2n-j}{j},$$

so

$$\psi_n(b) = \sum_{j=0}^{n} \left(-\frac{1}{2}b\right)^j (2n)! \binom{2n-j}{j} \Big/ (2n-2j)!,$$

and

$$(2b)^{2n}\psi_n\left(\frac{1}{2b^2}\right) = \sum_{j=0}^{n}(-1)^j (2b)^{2n-2j} \frac{(2n)!}{j!(2n-2j)!} \equiv H_{2n}(b), \tag{4.1}$$

where $H_{2n}(b)$ denotes the $(2n)$-th Hermite polynomial (Szegö (1939), p. 101).

Since $\psi_n(0) = 1$ and $\psi_n(b)$ is continuous, we confirm the result, known from Theorem 3.1, that $\psi_n(b)$ is positive for sufficiently small b. Now the zeros of $H_{2n}(b)$ are all real (Szegö (1939), p. 43); hence so are those of $\psi_n(b)$, by (4.1), and they are clearly positive. Thus, for fixed n, $\psi_n(b) \geq 0$ if $b \in [0, \beta_n]$ where β_n is the smallest zero of $\psi_n(b)$. Because the product of the zeros of $\psi_n(b)$ is

$$(-1)^n \frac{\text{coefficient of } b^0}{\text{coefficient of } b^n} = [1 \cdot 3 \cdots (2n-1)]^{-1}$$

we find

$$\beta_n < [1 \cdot 3 \cdots (2n-1)]^{-1/n}. \tag{4.2}$$

By Stirling's formula the right-hand side of (4.2) varies like $e/2n$ as $n \to \infty$.

Thus the range of b for which $\psi_n(b) \geq 0$ becomes degenerate as $n \to \infty$ and so, ultimately all the negative coefficients in $P(\boldsymbol{z})$ become arbitrarily small.

Note that, provided the relevant $a_{\boldsymbol{k}}$ with $\max_{1\leq i\leq n} k_i \geq 2$ are positive, the actual magnitudes of these coefficients do not affect the example.

Example 9. Consider a 3rd order n-variate Hermite distribution with $n \geq 3$ and $a_{\boldsymbol{k}} \equiv 1$ except that $a_{110\cdots0} = -b$ where $b > 0$. Again, Theorem 3.1 says that for b sufficiently small this specifies a p.g.fn. However, as opposed to the previous example, the permissible range of b does not vary with n.

To show this we use the approach of Lukacs ((1970), p. 251). It is routine to check that all coefficients in $[P(\boldsymbol{z})]^2$ are positive provided $b < 1$. With thought it is clear that in $[P(\boldsymbol{z})]^3$ there are really only a fixed number of constraints on b which are completely determined by the case $n = 3$; higher values of n just reproduce the $n = 3$ constraints because we have chosen the $a_{\boldsymbol{k}}$'s equal if $\boldsymbol{k} \neq [1, 1, 0, \ldots, 0]'$. Thus there is a range of positive values of b such that $[P(\boldsymbol{z})]^3$ has all coefficients positive for any $n \geq 3$. Some tedious algebra shows that the permissible range for b is again the interval $(0, 1)$. Hence, for b in this range, $[P(\boldsymbol{z})]^j$ has positive coefficients for all $j \geq 2$ and all $n \geq 3$. It only remains to check the coefficient of $z_1 z_2$ in $G(\boldsymbol{z})$. This is easily seen to be $2 - b$ and is positive when $b < 1$.

Thus provided $0 < b < 1$, we have obtained a $P(\boldsymbol{z})$ with one negative coefficient and such that $G(\boldsymbol{z})$ given by (1.3) is a p.g.fn for all $n \geq 3$.

Though they exhibit two interesting extremes of possible behaviour, these examples do not of themselves allow us to draw any simple conclusions regarding the point process problem. This is considered, again through examples, in the following section.

5. Point process examples

Suppose first that we allow point processes on an arbitrary (complete separable metric) phase space. Take the phase space to be any single point set. The p.g.fl of such a point process is simply an ordinary (univariate) p.g.fn. Thus the Lévy example, of Remark 3 in Section 3 above, provides an example of a point process, though admittedly a rather trivial one, for which not all the measures $H_k(\cdot)$ of (2.1) need be nonnegative. The Lévy example viewed as a special point process was referred to by Matthes *et al.* ((1978), p. 79) in connection with a problem about factorization of distributions.

A more satisfying point process example can be constructed as follows. Consider a Poisson process on $\boldsymbol{R}$ with mean measure $\Lambda\mu(\cdot)$, where $\mu(\cdot)$ is a totally finite measure with $\mu(\boldsymbol{R}) \equiv \omega < \infty$ and Λ is a discrete random variable whose p.g.fn is the Lévy example with $P(z)$ as in Remark 3 of Section 3. The p.g.fl of this process is easily seen to be $G[\xi]$, where

$$(5.1) \qquad \begin{aligned} \log G[\xi] &= P\left(\exp\left\{\int_{\boldsymbol{R}} [\xi(t) - 1]\mu(dt)\right\}\right) - P(1) \\ &= \sum_{k=1}^{\infty} \frac{1}{k!}(a - 2^k b + 3^k c + 4^k d) \\ &\quad \times \int_{\boldsymbol{R}} \cdots \int_{\boldsymbol{R}} \prod_{i=1}^{k} [\xi(t_i) - 1]\mu(dt_1) \cdots \mu(dt_k) \end{aligned}$$

$$(5.2) \qquad = \sum_{k=1}^{\infty} \frac{1}{k!}(ae^{-\omega} - 2^k be^{-2\omega} + 3^k ce^{-3\omega} + 4^k de^{-4\omega})$$
$$\times \int_R \cdots \int_R \left[\prod_{i=1}^{k} \xi(t_i) - 1\right] \mu(dt_1)\cdots\mu(dt_k).$$

Now choose a, b, c, d so that $a - 2^k b + 3^k c + 4^k d$ is non-negative for every $k \in \{1, 2, \ldots\}$. Notice that the previously mentioned choice $a = c = d = 1$, $b = 1/4$ is adequate. Then comparison of (5.1) with (2.2) reveals that in this case all the $H_k(\cdot)$ are non-negative measures.

On the other hand, for this same choice of a, b, c, d and for $k = 2\omega$, we find that the terms involving ω in (5.2) become

$$e^{-4\omega}\left[(e^3)^\omega - \frac{1}{4}(4e^2)^\omega + (9e)^\omega + 16^\omega\right].$$

For sufficiently large ω this expression is negative since $4e^2$ is the largest number being raised to the power ω. Hence, it is possible to choose ω so that, for at least one k, the corresponding measure $M_k(\cdot)$ in (5.2) is always negative.

It is pleasing that Lévy's example, the first and simplest instance of the general phenomenon we are studying, can be used to generate an interesting point process example. A rather similar use of the Lévy example is in Kallenberg ((1975), Exercise 8.6); he used it to exhibit a mixed Poisson distribution which is infinitely divisible even though the mixing distribution is not. More complicated mixing distributions leading to the same result can be found in Shanbhag and Westcott (1977).

In one respect the mixed Poisson process example considered above is not entirely satisfactory, as any point process so constructed will be a.s. finite. As an obvious extension we might try the same construction using a σ-finite $\mu(\cdot)$ such as Lebesgue measure. In this case the construction does work and the log p.g.fl can be expressed in the form (5.1). However, rearrangement into the form (5.2) is *not* possible. This provides a nice contrast to an example in Waymire and Gupta (1983) where it is the form (2.3), corresponding to (5.2) above, that cannot always be rearranged in the form (2.2) corresponding to (5.1) above.

6. Remarks on point process characterizations

Recall that it was the problem of characterizing p.g.fls of the form (2.1) which motivated this research. In Section 2, we noted the attempts on this problem, *via* (2.2) and (2.3), by Ammann and Thall (1977, 1978) and Waymire and Gupta (1983). Because they concentrated on the rather special subclass of regular infinitely divisible point processes, there is a simple direct probabilistic route to their results.

The key observation is that any regular infinitely divisible point process is a Poisson cluster process with almost surely finite clusters. This result was first proved in the stationary case by Kerstan and Matthes (1964); see also Daley and Vere-Jones ((1988), Proposition 8.4. VIII(ii)). The p.g.fl of such a process therefore has a well-known form, dating back at least to Moyal (1958) (see also Moyal (1962)), namely

$$\log G[\xi] = \int_{\boldsymbol{R}} \{G_s[\xi \mid x] - 1\}\mu(dx). \tag{6.1}$$

Here, $\mu(\cdot)$ is a Borel measure and $G_s[\xi \mid x]$ is the p.g.fl of the (almost surely finite) subsidiary process generated by the cluster centre at x. It follows readily that $\log G[\xi]$ can be expressed in the form (2.3) where each $M_k(\cdot)$ is non-negative; see Daley and Vere-Jones ((1988), Proposition 8.3. III).

We must of course ensure that the cluster process is well-defined, in the sense of having an almost surely finite number of points in any bounded set. This is equivalent (Matthes (1963); see also Westcott (1971)) to

$$\int_{\boldsymbol{R}} W(x; A)\mu(dx) < \infty \tag{6.2}$$

for all bounded Borel sets A, where $W(x; A)$ is the probability of obtaining at least one subsidiary point in A from the cluster centred at x. In terms of the $M_k(\cdot)$, (6.2) can be expressed as either

$$\sum_{k=1}^{\infty}\sum_{j=1}^{k} \binom{k}{j} M_k(A^j \times (\boldsymbol{R}\backslash A)^{k-j}) < \infty \tag{6.3}$$

or

$$\sum_{k=1}^{\infty}\sum_{j=1}^{k} (-1)^{j-1} \binom{k}{j} M_k(A^j \times \boldsymbol{R}^{k-j}) < \infty \tag{6.4}$$

for all bounded Borel sets A; (6.3) comes from expressing the event 'at least one point in A' in terms of the mutually exclusive events 'k points in the cluster and j of them in A' ($k, j \geq 1$), while (6.4) comes from using inclusion-exclusion formulae on the same event. Condition (6.3) is (4.13) of Waymire and Gupta (1983) (cf. (8.3.24) of Daley and Vere-Jones (1988)) and (6.4) is, essentially, (C_2^*) of Ammann and Thall (1977).

We believe this approach has the advantage that the analytical conditions $M_k(\cdot) > 0$ and (6.3), (6.4) emerge naturally from the probability structure. It also emphasizes that (2.3) rather than (2.2) is the more natural form to consider when the restriction of infinite divisibility is imposed. This was pointed out by Waymire and Gupta (1983). However, for finite sum exponents, such as in (2.1), we prefer (2.1) to the equivalent of (2.3) because of the interpretation of the $H_k(\cdot)$ as factorial cumulant measures.

Beyond the class of regular infinitely divisible point processes, the original characterization problem for (2.1), with $m > 2$, is still open. The main example of Section 5 does not really elucidate the more subtle aspects of the problem. Even the next simplest case, $m = 3$, seems intractable. Is (2.1) then always infinitely divisible, as is true for $m = 2$, or are signed measures allowed, as suggested by Example 9?

Acknowledgements

This work was initiated whilst one author (MW) was a Visiting Lecturer in the Department of Mathematics, University of Western Australia, and completed during the tenure by the other author (RKM) of a Mathematical Sciences Research Centre Visiting Fellowship in the Department of Statistics (I. A. S.), the Australian National University. Support from both these sources is gratefully acknowledged.

It is a pleasure to thank Dr. D. J. Daley for comments on an earlier draft, and also the referee for his careful reading and suggestions.

References

Ammann, L. P. and Thall, P. F. (1977). On the structure of regular infinitely divisible point processes, *Stochastic Process. Appl.*, **6**, 87–94.

Ammann, L. P. and Thall, P. F. (1978). Count distributions, orderliness and invariance of Poisson cluster processes, *J. Appl. Probab.*, **16**, 261–273.

Cuppens, R. (1975). *Decomposition of Multivariate Probabilities*, Academic Press, New York.

Daley, D. J. and Vere-Jones, D. (1988). *An Introduction to the Theory of Point Processes*, Springer, Berlin.

Feller, W. (1968). *An Introduction to Probability Theory and Its Applications*, Volume 1, 3rd ed., Wiley, New York.

Gupta, R. P. and Jain, G. C. (1974). A generalized Hermite distribution and its properties, *SIAM J. Appl. Math.*, **27**, 359–363.

Kallenberg, O. (1975). *Random Measures*, Akademie-Verlag, Berlin (Also (1976) Academic Press, London).

Kemp, A. W. and Kemp, C. D. (1965). Some properties of the 'Hermite' distribution, *Biometrika*, **52**, 381–394.

Kemp, C. D. and Papageorgiou, H. (1982). Bivariate Hermite distributions, *Sankhyā Ser. A*, **44**, 269–280.

Kerstan, J. and Matthes, K. (1964). Stationäre zufällige Punktfolgen II, *Jahresbericht der Deutschen Mathematiker-Vereinigung*, **66**, 106–118.

Lévy, P. (1937). Sur les exponentielles des polynômes et sur l'arithmétique des produits des lois de Poisson, *Ann. Sci. École Norm. Sup.*, **54**, 231–292.

Lukacs, E. (1970). *Characteristic Functions*, 2nd ed., Griffin, London.

Matthes, K. (1963). Unbeschränkt teilbare Verteilungsgesetze stationärer Punktfolgen, *Wissenschaftliche Zeitschrift der Hochschule für Elektrotechnik Ilmenau*, **9**, 235–238.

Matthes, K., Kerstan, J. and Mecke, J. (1978). *Infinitely Divisible Point Processes*, Wiley, Chichester.

Milne, R. K. and Westcott, M. (1972). Further results for Gauss-Poisson processes, *Adv. in Appl. Probab.*, **4**, 151–176.

Moyal, J. E. (1958). Discussion on the paper of Neyman and Scott (1958), 36–37.

Moyal, J. E. (1962). The general theory of stochastic population processes, *Acta Math.*, **108**, 1–31.

Newman, D. S. (1970). A new family of point processes which are characterized by their second moment properties, *J. Appl. Probab.*, **7**, 338–358.

Neyman, J. and Scott, E. L. (1958). A statistical approach to problems of cosmology (with discussion), *J. Roy. Statist. Soc. Ser. B*, **20**, 1–43.

Shanbhag, D. N. and Westcott, M. (1977). A note on infinitely divisible point processes, *J. Roy. Statist. Soc. Ser. B*, **39**, 331–332.

Steyn, H. S. (1976). On the multivariate Poisson Normal distribution, *J. Amer. Statist. Assoc.*, **71**, 233–236.

Szegö, G. (1939). *Orthogonal Polynomials*, American Mathematical Society Colloquium Publications, Vol. XXIII, American Mathematical Society, New York.

Teicher, H. (1954). On the multivariate Poisson distribution, *Skand. Aktuarietidskr.*, **37**, 1–9.
Van Harn, K. (1978). Classifying infinitely divisible distributions by functional equations, Mathematical Centre Tract 103, Mathematisch Centrum, Amsterdam.
Waymire, E. and Gupta, V. K. (1983). An analysis of the Polya point process, *Adv. Appl. Probab.*, **15**, 39–53.
Westcott, M. (1971). On existence and mixing results for cluster point processes, *J. Roy. Statist. Soc. Ser. B*, **33**, 290–300.