STRUCTURE OF EXCHANGEABLE INFINITELY DIVISIBLE SEQUENCES OF POISSON RANDOM VECTORS

Robert C. GRIFFITHS
Department of Mathematics, Monash University, Clayton, 3168, Australia

Robin K. MILNE
Department of Mathematics, University of Western Australia, Nedlands, 6009, Australia

Received 14 October 1985

De Finetti's Theorem reveals a simple explicit structure for an infinite exchangeable sequence of zero-one random variables. Although more general results are known, simple explicit results might be expected in particular settings. In this paper such results are obtained for exchangeable sequences of infinitely divisible Poisson random variables and random vectors. The methods employed are elementary, except in that they involve appeal to moment theorems.

exchangeability * Poisson random variables * Poisson random vector * moment theorems

1. Introduction

A well-known result due to de Finetti exhibits a simple explicit structure for an infinite exchangeable sequence of random variables that can take only the values zero or one. Although very general results about exchangeable sequences are known (see Feller [6, Section VII.4], Kingman [10], and Olshen [11]), it is still of interest to derive simpler explicit results in particular cases.

This paper is concerned with exchangeable sequences of Poisson random vectors. To achieve explicit results it is necessary to start from some other structural assumptions about the sequences. The assumption that they are infinitely divisible leads to a structure that is satisfyingly explicit, yet surprisingly rich. Apart from requisite probability theory, the methods to be used are relatively elementary. The deepest results employed are moment theorems, and these are collected together in the appendix.

To facilitate understanding in the simplest setting we have felt it appropriate to treat initially the random variable case, starting first with finite sequences and then extending to infinite sequences. This is done in Section 2. Then, Section 3 deals with the general random vector case which requires appeal to a multivariate moment result concerning measures on an $r$-dimensional simplex. This moment result, Theorem A3, appears in itself to be new.

In a sequel [7] we discuss analogous structure for infinite exchangeable infinitely divisible sequences of Poisson point processes.
2. Exchangeable infinitely divisible sequences of Poisson random variables

Let \( \{X_1, \ldots, X_k\} \) be a sequence of \( k \) Poisson random variables that are jointly infinitely divisible. From results of Dwass and Teicher [5] it follows that the logarithm of \( P(s), s = (s_1, \ldots, s_k) \in [0, 1]^k \), the joint probability generating function (pgfn) of \( X_1, \ldots, X_k \), must be a multi-linear function of \( s_1, \ldots, s_k \) or equivalently of \( s_1 - 1, \ldots, s_k - 1 \). Under the further condition that the sequence \( \{X_1, \ldots, X_k\} \) be exchangeable, it is clear that the pgfn \( P(s) \) can be written as

\[
P(s) = \exp \left\{ \sum_{i=1}^{k} \alpha_i e_i(s - 1) \right\},
\]

where \( \alpha_1, \ldots, \alpha_k \) are constants, not necessarily non-negative, \( 1 \) denotes the \( k \)-vector of ones, and \( e_i(s) \), \( i = 1, \ldots, k \), are the elementary symmetric functions in \( s_1, \ldots, s_k \).

Recall that \( e_1(s), \ldots, e_k(s) \) can be defined by the generating function

\[
\prod_{i=1}^{k} (1 + u s_i) = 1 + \sum_{m=1}^{k} u^m e_m(s).
\]

The constant \( \alpha_i \) is simply the \( i \)th order joint factorial cumulant of the random variables \( X_1, \ldots, X_k \).

For suitable constants \( \beta_1, \ldots, \beta_k \), (2.1) can be expressed in the form

\[
P(s) = \exp \left\{ \sum_{m=1}^{k} \beta_m [e_m(s) - \gamma_m] \right\},
\]

where \( \gamma_m = \binom{k}{m} \), \( m = 1, \ldots, k \). A well-known result (cf. Horn and Steutel [8, Theorem 2.1]), gives non-negativity of \( \beta_1, \ldots, \beta_k \) as necessary and sufficient conditions for a function of the form (2.3) to be a pgfn. The corresponding conditions for a function of the form (2.1) to be a pgfn are not so straightforward, but will be obtained shortly.

The precise connection between the forms (2.1) and (2.3) is provided by the reciprocal relationships

\[
\alpha_i = \sum_{j=1}^{k} \binom{k-i}{j-i} \beta_j, \quad i = 1, \ldots, k,
\]

and

\[
\beta_m = \sum_{n=m}^{k} \binom{k-m}{n-m} (-1)^{n-m} \alpha_n, \quad m = 1, \ldots, k,
\]

which are easily deduced using the identities

\[
1 + \sum_{i=1}^{k} u^i e_i(s - 1) = \prod_{i=1}^{k} [(1 - u) - u s_i] = (1 - u)^k + \sum_{m=1}^{k} (1 - u)^{k-m} u^m e_m(s).
\]

In view of the preceding discussion we have the following result.
Lemma 1. Necessary and sufficient conditions for (2.1) to be a pgfn are that
\[
\sum_{n=m}^{k} \alpha_n \binom{k-m}{n-m} (-1)^{n-m} \geq 0, \quad m = 1, \ldots, k.
\] (2.5)

This lemma completes the structural picture for finite exchangeable sequences of Poisson random variables that are jointly infinitely divisible. For infinite sequences we can refine this structure.

Theorem 1. In order that (2.1) for \( k = 1, 2, \ldots \) be the sequence of finite-dimensional pgfns of some infinite sequence \( \mathcal{X} = \{X_1, X_2, \ldots \} \) of non-degenerate random variables, it is necessary and sufficient that \( \alpha_n = \int_{[0,1]} u^{n-1} \rho(du), \quad n = 1, 2, \ldots \) (2.6)

where \( \rho \) is a finite measure on \([0,1]\) such that \( \rho([0,1]) > 0 \); equivalently, that \( \alpha_n = \gamma \mathbb{E}(U^{n-1}), \quad n = 1, 2, \ldots \) (2.7)

where \( \gamma \in (0, \infty) \) and \( U \) is a random variable taking values in \([0,1]\). Further, when either of (2.6) or (2.7) is satisfied the sequence \( \mathcal{X} \) is an exchangeable sequence of Poisson random variables that are jointly infinitely divisible.

Remark 1. One could cover also the case where the random variables \( X_1, X_2, \ldots \) are all degenerate at zero by allowing \( \rho \) to be a measure on \([0,1]\) such that \( \rho([0,1]) = 0 \).

Proof. Observe that provided (2.1) is a pgfn for \( k = 1, 2, \ldots \), these pgfns are consistent, and hence are the finite-dimensional pgfns of some infinite sequence of random variables.

Now define a linear functional \( \mathcal{L}_\alpha \) as at (A1) of the appendix, but only for \( m = 1, 2, \ldots \). Then, using the outer expressions in (2.4), we have for \( k = 1, 2, \ldots \) and \( u \in [0,1] \)
\[
\sum_{i=1}^{k} \alpha_i e_i(s-1) = \mathcal{L}_\alpha[(1-u)^k - 1] + \sum_{m=1}^{k} \mathcal{L}_\alpha[(1-u)^{k-m}u^m]e_m(s).
\]

Recalling the necessary and sufficient conditions for a function of the form (2.3) to be a pgfn, it is clear that (2.1) is a pgfn for \( k = 1, 2, \ldots \) iff
\[
\mathcal{L}_\alpha[(1-u)^{k-m}u^m] \geq 0, \quad m = 1, \ldots, k, \quad k = 1, 2, \ldots
\]

By Corollary 2 to Theorem A1 of the appendix, these conditions are necessary and sufficient for \( \alpha_1, \alpha_2, \ldots \) to be representable in the form (2.6). Setting \( \gamma = \rho([0,1]) \) the alternative representation (2.7) is immediate from (2.6).

Finally, in view of the remarks preceding (2.1), it is clear that the remaining assertions of this theorem hold. \( \square \)
**Corollary 1.** When the conditions of the theorem are satisfied, the random variables of the sequence $\mathcal{X}$ are

(i) mutually independent iff $U = 0$ almost surely, or
(ii) completely dependent (i.e. almost surely equal) iff $U = 1$ almost surely.

**Corollary 2.** When the conditions of the theorem are satisfied, (2.1) can be written as

$$P(s) = \exp\left\{ \int_{[0,1]} \left( \prod_{i=1}^{k} [1 + u(s_i - 1)] - 1 \right)^{-1} u^{-1} \rho(du) \right\}. \quad (2.8)$$

**Proof.** Use (2.6) and the first identity of (2.4), noting that in the bracket preceding $u^{-1}$ in (2.8) only terms involving $u^n$, $n = 1, \ldots, k$, appear. □

It is convenient to have a further representation for $\alpha_1, \alpha_2, \ldots$.

**Theorem 2.** In order that (2.1) for $k = 1, 2, \ldots$ be the sequence of finite-dimensional pgfns of some infinite sequence $\mathcal{X} = \{X_1, X_2, \ldots\}$ of non-degenerate random variables, it is necessary and sufficient that

$$\alpha_n = \lambda \delta_{1n} + \int_{[0,1]} u^n A(du), \quad n = 1, 2, \ldots, \quad (2.9)$$

where $\delta_{1n}$ is a Kronecker delta, $\lambda \in [0, \infty)$, and $A$ is a not necessarily finite measure on $(0, 1)$ such that $\gamma$, given now by $\gamma = \lambda + \int_{(0,1)} u A(du)$, satisfies $0 < \gamma < \infty$. Further, when (2.9) is satisfied the sequence $\mathcal{X}$ is an exchangeable sequence of Poisson random variables that are jointly infinitely divisible.

**Remark 2.** The case where $X_1, X_2, \ldots$ are all degenerate at zero could be covered by allowing $\gamma = 0$ i.e. by taking $\lambda = 0$ and $A$ as the zero measure.

**Proof.** Given a measure $\rho$ on $[0, 1]$ define $\lambda = \rho(\{0\})$ and a measure $A$ on $(0, 1]$ by

$$A(B) = \int_B u^{-1} \rho(du)$$

for $B$ any Borel set of $(0, 1]$. Then, clearly, $\lambda \geq 0$ and finiteness of $\rho$ implies finiteness of both $\lambda$ and $\int_{[0,1]} u A(du)$. Conversely, given $\lambda \in [0, \infty)$ and a measure $A$ on $(0, 1]$ satisfying $\int_{[0,1]} u A(du) < \infty$, define a finite measure $\rho$ on $[0, 1]$ by

$$\rho(B) = \lambda \delta_0(B) + \int_B u A(du)$$

where $\delta_0$ denotes Dirac measure (i.e. ‘a unit mass at’) zero.

With such identifications, (2.6) and (2.9) are equivalent and the required results follow immediately from Theorem 1. □
Corollary 3. When the conditions of Theorem 2 are satisfied, (2.1) can be written as

\[ P(s) = \exp \left\{ \lambda \sum_{i=1}^{k} (s_i - 1) + \int_{(0,1)} \left[ \prod_{i=1}^{k} (1 - u + u s_i) - 1 \right] \Lambda(du) \right\}. \]  

(2.10)

Proof. Use (2.9), and the first identity of (2.4) or, alternatively, (2.8). □

Remark 3. If \( \lambda = 0 \) in (2.10) then

\[ \mathbb{P}(X_1 = \cdots = X_k = 0) = \exp \left\{ - \int_{(0,1)} [1 - (1 - u)^k] \Lambda(du) \right\}. \]

Hence, since \( A_k = \{X_1 = \cdots = X_k = 0\} \) is a non-increasing sequence of sets,

\[ \mathbb{P}(X_i = 0, i = 1, 2, \ldots) = \lim_{k \to \infty} \mathbb{P}(A_k) = \exp(-\Lambda((0,1])). \]

Thus, when \( \lambda = 0 \), we have

\[ \mathbb{P}(X_i = 0, i = 1, 2, \ldots) > 0 \quad \text{iff} \quad \Lambda((0,1)) < \infty. \] □

The above results yield a simple probabilistic construction for an exchangeable infinitely divisible sequence, \( \mathcal{X} \), of Poisson random variables.

Consider a Poisson process defined on \((0,1]\) and having mean measure \( \Lambda \) satisfying \( \int_{(0,1]} u \Lambda(du) < \infty \). (See Remark 4 below.) Take countably many identical copies of this process, label the copies 1, 2, \ldots and, conditional on the Poisson process, delete points independently from each copy. Within each copy a point at \( u \) is deleted with probability \( 1 - u \), or retained with probability \( u \), independently of other points in that copy. Suppose that \( \mathcal{V} = \{V_1, V_2, \ldots\} \) gives the number of points that remain in each of the copies. Then the joint pgfn of \( V_1, \ldots, V_k \) is

\[ \exp \left\{ \int_{(0,1)} \left[ \prod_{i=1}^{k} (1 - u + u s_i) - 1 \right] \Lambda(du) \right\}. \]  

(2.11)

This is readily seen, for example, from results about Poisson cluster processes or compound Poisson processes (cf. Cox and Isham [3, Chapter 3]; Daley and Vere-Jones [4, Section 5]).

Now suppose that \( \mathcal{U} = \{U_1, U_2, \ldots\} \) is independent of \( \mathcal{V} \) and consists of mutually independent and identically distributed Poisson random variables with mean \( \lambda \). Set \( \mathcal{X} = \mathcal{U} + \mathcal{V} \), where addition of sequences is defined element-wise. Then the joint pgfn of \( X_1, \ldots, X_k \) is of the form (2.10) for \( k = 1, 2, \ldots \) with \( \Lambda \) as in Theorem 2.

Remark 4. In the theory of point processes it is usual to restrict attention to processes which have, almost surely, a finite number of points in each bounded set and, in the case of Poisson processes, to processes having a mean measure which is finite on bounded sets. (See for example Cox and Isham [3, pp. 22-23] and Daley and Vere-Jones [4, Sections 1 and 2].) A Poisson process on \((0,1]\) with mean measure \( \Lambda \) satisfying \( \int_{(0,1]} u \Lambda(du) < \infty \) will not in general meet this requirement, since it may
be that $\Lambda((0, 1]) = \infty$. However, such a process can be considered, using the usual theory, as being obtained from a suitable Poisson process on $[1, \infty)$ by means of the transformation $x \to x^{-1}$ which transforms points in $[1, \infty)$ into points in $(0, 1]$.

**Remark 5.** Notice that in the above construction deletions are much more likely for points near the origin than for points near one. Thus, even when $\Lambda((0, 1]) = \infty$, this pattern of deletions balances the high concentration of points near the origin in the Poisson process with mean measure $\Lambda$ satisfying $\int_{(0,1]} u\Lambda(du) < \infty$, to give overall a finite number of points in each of the deleted copies i.e. $\mathbb{E} V_i < \infty, i = 1, 2, \ldots$.

**Corollary 4.** Suppose that $\{X_1, X_2, \ldots\}$ is a sequence of random variables such that for $k = 1, 2, \ldots, X_1, \ldots, X_k$ have joint pgfn (2.10). Then $(X_1 + \cdots + X_k)/k$ converges in distribution as $k \to \infty$ to the (infinitely divisible) distribution with Laplace transform $L(t)$ given by

$$L(t) = \exp \left\{ \int_{(0,1]} \left[ e^{-tu} - 1 \right] \Lambda(du) - \lambda t \right\}, \quad t \geq 0. \quad (2.12)$$

**Proof.** The Laplace transform of $(X_1 + \cdots + X_k)/k$ is clearly

$$\exp \left\{ \int_{(0,1]} \left[ (1 + u(e^{-1/k} - 1))^k - 1 \right] \Lambda(du) + \lambda k(e^{-t/k} - 1) \right\}, \quad t \geq 0,$$

and this function converges pointwise, as $k \to \infty$, to $L(t)$ given by (2.12). \qed

3. Exchangeable infinitely divisible sequences of Poisson random vectors

Consider an exchangeable sequence $\mathcal{X} = \{X_1, X_2, \ldots\}$ of Poisson random $l$-vectors which are jointly infinitely divisible. Vectors in $\mathcal{X}$ are row vectors and $X_i = (X_{i1}, \ldots, X_{il}), i = 1, 2, \ldots$. In the case that the sequences

$\mathcal{X}_j = \{X_{1j}, X_{2j}, \ldots\}, \quad j = 1, 2, \ldots, l,$

are mutually independent, the structure of the sequence $\mathcal{X}$ is completely determined by the structure of results of the preceding section applied separately to each of the sequences $\mathcal{X}_1, \ldots, \mathcal{X}_l$. This special case will be considered further in Remark 9 and in a point process setting in Griffiths and Milne [7]. Here we investigate the general (random vector) case.

Let the joint pgfn of $\mathcal{X}_1, \ldots, \mathcal{X}_k$ be

$$P(s_1, \ldots, s_k) = \mathbb{E} \left\{ \prod_{i=1}^{k} s_i^{X_i} \right\} = \mathbb{E} \left\{ \prod_{i=1}^{k} \prod_{j=1}^{l} s_{ij}^{X_{ij}} \right\},$$

where $s_i = (s_{i1}, \ldots, s_{il}) \in [0, 1]^l, \quad i = 1, \ldots, k$. This function will be written also as $P(S)$, where $S = [s_{ij}]$, a $k \times l$ matrix. Since $X_1, \ldots, X_k$ are infinitely divisible Poisson random vectors, it follows from results of Dwass and Teicher [5] that $\log P(S)$ is
a multi-linear function of the elements of \( S \), or equivalently of the elements of \( S - J \) where \( J \) denotes the \( k \times l \) matrix with all unit entries. Thus

\[
\log P(S) = \sum_{C_1, \ldots, C_k} \phi_{C_1, \ldots, C_k} s_{C_1}^* s_{C_2}^* \cdots s_{C_k}^*. \tag{3.1}
\]

Here the summation is over all \( C_1, \ldots, C_k \in \mathcal{P}\{1, \ldots, l\} \) the power set of \( \{1, 2, \ldots, l\} \),

\[
s_{iC}^* = \begin{cases} 
\prod_{j \in C} (s_{ij} - 1), & C \neq \emptyset, \\
1, & C = \emptyset.
\end{cases}
\]

\( i, \ldots, k \), and \( \{\phi_{C_1, \ldots, C_k}\} \) is a set of not necessarily non-negative constants such that \( \phi_{\emptyset, \ldots, \emptyset} = 0 \) (since \( P(J) = 1 \)). When the sequence \( \{X_1, \ldots, X_k\} \) is exchangeable, it follows that \( \phi_{C_{\pi(1)}, \ldots, C_{\pi(k)}} \) is invariant under any permutation \( \pi \) on \( 1, \ldots, k \).

Now let \( \mathcal{P} = \mathcal{P}\{1, \ldots, l\}\setminus\{\emptyset\} \) and suppose that \( A_1, \ldots, A_r \), with \( r = 2^l - 1 \), is a fixed labelling of the elements of \( \mathcal{P} \). Consider any sequence \( \mathcal{C} = \{C_1, \ldots, C_k\} \) of \( k \) sets from \( \mathcal{P}\{1, \ldots, l\} \). For \( a \in \mathbb{Z}_+ \), where \( \mathbb{Z}_+ = \{0, 1, 2, \ldots\} \), write \( \sigma(C) = a \) if in the sequence \( \mathcal{C} \) the set \( A_j \) appears exactly \( a_j \) times, \( j = 1, \ldots, r \). Define \( e_a(S) \), a symmetric multi-linear function of \( s_1, \ldots, s_k \), by

\[
e_a(S) = \sum_{C: \sigma(S) = a} s_{1C} s_{2C} \cdots s_{kC}, \quad a \neq \emptyset, \tag{3.2}
\]

where for \( i = 1, \ldots, k \) and \( C \in \mathcal{P}\{1, \ldots, l\} \)

\[
s_{iC} = \begin{cases} 
\prod_{j \in C} s_{ij}, & C \neq \emptyset, \\
1, & C = \emptyset.
\end{cases}
\]

For \( a \in \mathbb{Z}_+ \) put \( a_+ = a_1 + \cdots + a_r \). Set \( \mathcal{A}_m = \{a \in \mathbb{Z}_+: a_+ = m\} \), \( m = 1, 2, \ldots \), \( \mathcal{A} = \bigcup_{m=1}^{\infty} \mathcal{A}_m \), and for \( u = (u_{A_1}, \ldots, u_{A_k}) \in [0, 1]^r \) write \( u^a \) in place of \( u_{A_1}^{a_1} u_{A_2}^{a_2} \cdots u_{A_r}^{a_r} \). Clearly, the functions \( e_a(S) \), \( a \in \mathcal{A} \), could be defined equivalently by the generating function

\[
G(u, S) = \prod_{i=1}^k \left[ 1 + \sum_{A \in \mathcal{P}_i} s_{iA} u_A \right] = 1 + \sum_{m=1}^k \sum_{a \in \mathcal{A}_m} u^a e_a(S). \tag{3.3}
\]

The simplest non-trivial case is when \( l = 2 \): then

\[
G(u, S) = \prod_{i=1}^k \left[ 1 + s_{i1} u_{\{1\}} + s_{i2} u_{\{2\}} + s_{i1} s_{i2} u_{\{1,2\}} \right].
\]

Return now to the simplification of (3.1) that results from exchangeability of \( \{X_1, \ldots, X_k\} \). Using (3.2) it follows that

\[
\log P(S) = \sum_{m=1}^k \sum_{a \in \mathcal{A}_m} \alpha_a e_a(S - J)
\]
where \( \alpha_a \) is written for the common value \( \phi_{c_a(n_1), \ldots, c_a(n_k)} \) under any permutation \( \pi \) on \( 1, \ldots, k \) (since only the multiplicities of the subscript sets need now be distinguished).

To proceed further requires an analogue of (2.4). Consider

\[
G(u, S-J) = \prod_{i=1}^{k} \left[ 1 + \sum_{A \in \mathcal{P}_i} s_{iA}^* u_A \right].
\]  

(3.4)

For \( A \in \mathcal{P}_i \), define \( v_A \) to be the coefficient of \( s_{iA} \) in the expansion of

\[
1 + \sum_{A \in \mathcal{P}_i} s_{iA}^* u_A
\]

as a linear function of \( s_{iA} \), \( A \in \mathcal{P}_i \). Setting successively \( S = J \) and \( S = 0 \), a zero matrix, shows that the constant term in this expansion is \( 1 \ v_+ \) where

\[
v_+ = \sum_{A \in \mathcal{P}_i} v_A = \sum_{A \in \mathcal{P}_i} (-1)^{|A|+1} u_A
\]

(3.5)

with \(|A|\) denoting the cardinality of the set \( A \). Thus \( v_A \) is defined by the identity

\[
1 + \sum_{A \in \mathcal{P}_i} s_{iA}^* u_A = 1 \ v_+ + \sum_{A \in \mathcal{P}_i} s_{iA} v_A.
\]

(3.6)

Hence

\[
v_A = \sum_{B \supseteq A} (-1)^{|B \setminus A|} u_B
\]

(3.7)

where the summation is over all \( B \in \mathcal{P}_i \) satisfying the stated inclusion. Expression (3.7) inverts to

\[
u_A = \sum_{B \supseteq A} v_B.
\]

(3.8)

This can be seen directly by consideration of the identity

\[
1 + \sum_{A \in \mathcal{P}_i} s_{iA} u_A = 1 - v_+ + \sum_{A \in \mathcal{P}_i, j \notin A} \prod_{i} (s_{ij} + 1) u_A,
\]

or by recognition of one of the simplest special cases of Möbius inversion (cf. Aigner [1, Section 4.19]). From (3.4) and (3.6) it follows that

\[
G(u, S-J) = \prod_{i=1}^{k} \left[ 1 - v_+ + \sum_{A \in \mathcal{P}_i} s_{iA} v_A \right].
\]

(3.9)

Finally, by expanding the products in (3.4) and (3.9) we obtain

\[
1 + \sum_{m=1}^{k} \sum_{a \in \mathcal{A}_m} u^* e_a(S-J) = G(u, S-J)
\]

\[
= (1 - v_+)^k + \sum_{m=1}^{k} \sum_{a \in \mathcal{A}_m} (1 - v_+)^k u^* e_a(S)
\]

(3.10)

which is the required analogue of (2.4).
Now define a linear function $\mathcal{L}$ acting on the vector space of all real coefficient polynomials in the $r$ variables $u_{A_1}, \ldots, u_{A_k}$ by

$$
\mathcal{L}[1] = 1, \quad \mathcal{L}[u^a] = \alpha_a, \quad a \in \mathcal{A}.
$$

(In analogy with Section 2, $\mathcal{L}$ depends on $\alpha_a, a \in \mathcal{A}$, though here we prefer to avoid this complexity of notation.) This linear functional acts also on polynomials in $v_{A_1}, \ldots, v_{A_k}$, since by (3.7), these are linear functions of $u_{A_1}, \ldots, u_{A_k}$. Then, using the outer expressions of (3.10), we have for $k = 1, 2, \ldots$

$$
\sum_{m=1}^{k} \sum_{a \in \mathcal{A}_m} \alpha_a e_a(S - J) = \mathcal{L}[(1 - v_+)^k - 1] + \sum_{m=1}^{k} \sum_{a \in \mathcal{A}_m} \mathcal{L}[(1 - v_+)^{k+1} e_a(S)] e_a(S),
$$

(3.11)

Observe (cf. [8, Theorem 2.1]) that (3.11) is the logarithm of a pgfn iff

$$
\mathcal{L}[(1 - v_+)^{k-a} e_a] \geq 0, \quad a \in \mathcal{A}, \quad a_+ \leq k, \quad k = 1, 2, \ldots
$$

By Theorem A3 of the appendix, these conditions are necessary and sufficient for $\beta_a = \mathcal{L}[v^a], a \in \mathcal{A}$, to be representable in the form (A3). This observation, together with consistency which follows as in the proof of Theorem 1, yields the following result.

**Theorem 3.** In order that $P(s_1, \ldots, s_k), \ k = 1, 2, \ldots$ be the sequence of finite-dimensional pgfns of some infinite sequence $\mathcal{X} = \{X_1, X_2, \ldots\}$ of non-degenerate random $l$-vectors, it is necessary and sufficient that

$$
\log P(s_1, \ldots, s_k) = \sum_{A \in \mathcal{P}_1} \lambda(A) \sum_{i=1}^{k} (s_i - 1) + 
\int_{\mathcal{F}_r \setminus \{o\}} \left\{ \prod_{i=1}^{k} \left[ 1 - v_+ + \sum_{A \in \mathcal{P}_1} s_i A v_A \right] - 1 \right\} \Lambda(dv),
$$

(3.12)

where $\lambda(A) \in [0, \infty), \ A \in \mathcal{P}_1$, and $\Lambda$ is a not necessarily finite measure on $\mathcal{F}_r \setminus \{o\}$ satisfying

$$
\int_{\mathcal{F}_r \setminus \{o\}} v_+ \Lambda(dv) < \infty
$$

and

$$
\sum_{A \in \mathcal{P}_1} \lambda(A) + \int_{\mathcal{F}_r \setminus \{o\}} v_+ \Lambda(dv) > 0.
$$

The constants $\lambda(A), A \in \mathcal{P}_1$ and the measure $\Lambda$ are unique. Further, when (3.12) is satisfied $\mathcal{X}$ is an exchangeable sequence of Poisson random vectors that are jointly infinitely divisible.
Remark 6. The case where \( X_1, X_2, \ldots \) are all degenerate at zero could be covered by allowing \( \lambda(A) = 0, A \in P \), and taking \( \Lambda \) to be the zero measure.

Remark 7. In the sense of giving a representation for \( \alpha_a, a \in \mathcal{A} \), in terms of the variables \( u_{A_1}, \ldots, u_{A_n} \) rather than \( v_{A_1}, \ldots, v_{A_n} \), there is no direct analogue when \( r > 1 \) of Theorem 2. However, suppose that \( \alpha_a \) can be represented as at (A3). Then, since under the linear map from \( \mathbb{R}^r \) to \( \mathbb{R}^r \) defined by \( \lambda(A) = 0, A \in P \), and \( \Lambda \) on \( \mathcal{Y}_r \backslash \{o \} \) is transformed to some measure \( \Lambda^* \) on \( [0, 1]^r \backslash \{0 \} \), there is a representation

\[
\alpha_a = \sum_{j=1}^{r} \lambda_j \delta_{a_j} + \int_{[0, 1]^r \backslash \{0 \}} u^* \Lambda^*(du), \quad a \in \mathcal{A},
\]

where otherwise the notation is as in (A3). Of course in the case \( r - 1 \) the two representations coincide, and this case has been dealt with independently in Section 2.

Suppose addition of sequences is defined element-wise. Consider any sequence \( \mathcal{X} = \{ X_1, X_2, \ldots \} \) having finite-dimensional log pgfns given by (3.12). Then \( \mathcal{X} \) can be decomposed as \( \mathcal{X} = \mathcal{U} + \mathcal{V} \) where \( \mathcal{U} \) and \( \mathcal{V} \) are independent sequences of random vectors such that the elements \( U_1, U_2, \ldots \) of \( \mathcal{U} \) are mutually independent and identically distributed Poisson random vectors each with log pgfn

\[
\sum_{A \in P} \lambda(A) (s_A - 1)
\]

and \( \mathcal{V} \) is an exchangeable sequence of Poisson random vectors having finite-dimensional log pgfns given by (3.12) with \( \lambda(A) = 0, A \in P \). A sequence of the form \( \mathcal{U} \) will be said to consist of independent components. Further, a sequence of the form \( \mathcal{V} \) will be said to have no independent components. Observe that (3.13) is simply a compact expression for a multi-linear function of \( s_1, \ldots, s_j \) which is the log pgfn of a general infinitely divisible Poisson random \( l \)-vector, as derived by Dwass and Teicher [5]. As these authors pointed out, there is an obvious interpretation in terms of random elements in common.

The following construction yields sequences with no independent components. Let \( \mathcal{N} \) be a Poisson process defined on \( (0, 1] \) and having mean measure \( \mu \) satisfying \( \int_{(0,1]} t \mu(dt) < \infty \). Take a doubly indexed sequence,

\[
\{ \mathcal{N}_{ij} : i = 1, 2, \ldots; j = 1, \ldots, l \},
\]

of identical copies of this process. Conditional on the Poisson process, points are deleted or retained, independently for each \( i \), according to the following scheme. For given \( i \), a point at \( t \) in the original process is deleted with probability \( \frac{1}{l} \) from all of the copies \( \mathcal{N}_{ij}, j = 1, \ldots, l \), and for \( A \in P_l \), retained in those \( \mathcal{N}_{ij} \) for which \( j \in A \) and deleted from those \( \mathcal{N}_{ij} \) for which \( j \not\in A \), this being done with probability \( Y_A(t) \) where \( \{ Y_A(t) : A \in P_l \} \) is a random element of \( \mathcal{Y}_l \backslash \{0 \} \) such that

\[
\sum_{A \in P_l} Y_A(t) = t.
\]
The entire deletion procedure is carried out independently for each point in the original process. Let $X_{ij}$ be the number of points retained in $N_{ij}$, $i = 1, 2, \ldots, k$, $j = 1, \ldots, l$. Then the joint pgfn of these $kl$ random variables has logarithm

$$
\int_{(0,1]} \int_{\mathcal{A} \setminus \{o\}} \left\{ \prod_{i=1}^{k} \left[ 1 - t + \sum_{A \in \mathcal{P}_j} s_{iA} Y_A(t) \right] - 1 \right\} \Phi_i(dy) \mu(dt) \quad (3.14)
$$

where $\Phi_i$ is the distribution of $\{ Y_A(t): A \in \mathcal{P}_i \}$. To see that (3.14) agrees with the second term of (3.12) take $\Lambda$ to be the measure defined by

$$
\Lambda(B) = \int_{(0,1]} \int_B \Phi_i(dy) \mu(dt)
$$

where $B$ is any Borel set of $\mathcal{A} \setminus \{o\}$.

**Remark 8.** Let $\mathcal{X} = \{X_1, X_2, \ldots\}$ be an infinitely divisible exchangeable sequence of Poisson random $l$-vectors such that the elements within each vector are mutually independent. Then the log pgfn of $X_1$ must have the form

$$
\sum_{j=1}^{l} \lambda_j (s_{1,j} - 1)
$$

where as from (3.12) it can be represented as

$$
\sum_{A \in \mathcal{P}_i} \lambda(A)(s_{1,A} - 1) + \sum_{A \in \mathcal{P}_i} \int_{\mathcal{A} \setminus \{o\}} n_A(dv)(s_{1,A} - 1).
$$

The identity of these two expressions implies that $\lambda(A) = 0$ and

$$
\int_{\mathcal{A} \setminus \{o\}} v_A A(dv) = 0
$$

whenever $|A| > 1$, $A \in \mathcal{P}_i$. Thus $\Lambda$ is concentrated on the set

$$
\{ v \in \mathcal{A}: v_A = 0, |A| > 1 \}
$$

and can therefore be regarded as a measure on the $l$-simplex

$$
\mathcal{A}_l = \{ v \in [0,1]^k: v_+ \leq 1 \}
$$

where now we write $v_j = v_{(j)}$, $j = 1, \ldots, l$. The joint log pgfn of $X_1, \ldots, X_k$ then has the form

$$
\sum_{j=1}^{l} \lambda_j \sum_{i=1}^{k} (s_{ij} - 1) + \int_{\mathcal{A} \setminus \{o\}} \left\{ \prod_{i=1}^{k} \left[ 1 - t_j + \sum_{j=1}^{l} s_{ij} t_j \right] - 1 \right\} \Lambda(dv). \quad (3.15)
$$

**Remark 9.** Now let $\mathcal{X} = \{X_1, X_2, \ldots\}$ be an infinitely divisible exchangeable sequence of Poisson random $l$-vectors such that the sequences

$$
\mathcal{X}_j = \{X_{ij}: i = 1, 2, \ldots\}, \quad j = 1, \ldots, l
$$

are mutually independent. Then from (2.10) the joint log pgfn of $X_1, \ldots, X_k$ has the form

$$
\sum_{j=1}^{l} \lambda_j \sum_{i=1}^{k} (s_{ij} - 1) + \sum_{j=1}^{l} \int_{(0,1)} \left[ \prod_{i=1}^{k} \left( 1 - u + us_{ij} \right) - 1 \right] A_j(du) \tag{3.16}
$$

with $\lambda_j, A_j, j = 1, \ldots, l$, satisfying conditions as in Theorem 2. In particular each measure $A_j$ is concentrated on $(0, 1)$. Observe that (3.12) has this form when $A(A) = 0$ if $|A| > 1$, $A \in \mathcal{D}$, and $A$ is concentrated on

$$
\bigcup_{j=1}^{l} \{ v \in \mathcal{P}_j : 0 < v_{(j)} \leq 1; v_A = 0, A \in \mathcal{P}_{\setminus \{j\}} \}.
$$

Clearly (3.15) has the form (3.16) when $A$ is concentrated on

$$
\bigcup_{j=1}^{l} \{ v \in \mathcal{P}_j : 0 < v_j \leq 1; v_i = 0, i \neq j \}.
$$

Appendix: Moment problems

**A1. The classical moment problem on [0, 1]**

Previous writings on this problem have tended to adopt either an analytic, perhaps functional analytic, approach (cf. Akhieser [2, p. 73ff], Karlin and Studden [9, Chapter IV], and Shohat and Tamarkin [12, pp. 8–9]), or a probabilistic approach as in Feller [6, VII.3]. Feller’s approach comes close to using linear functionals, though they are never explicit. The linear functional notation is well-suited to the needs of the present paper: see for example the proofs of Theorems 1 and 3.

Let $\mathcal{Q}$ denote the vector space of all polynomials in $u \in [0, 1]$ and having real coefficients. For a given sequence $\{ (\alpha_0, \alpha_1, \ldots) \}$ of real numbers, a linear functional, $\mathcal{L}_\alpha$, can be defined on $\mathcal{Q}$ starting from

$$
\mathcal{L}_\alpha[u^m] = \alpha_m, \quad u \in [0, 1], \tag{A1}
$$

for $m = 0, 1, \ldots$, or on the obvious subspace of $\mathcal{Q}$ starting from (A1) for $m = 1, 2, \ldots$, and extending by linearity. The solution of the classical moment problem can be stated as follows.

**Theorem A1.** A sequence $\{ \alpha_0, \alpha_1, \ldots \}$ of real numbers with $\alpha_0 = 1$ is the moment sequence of some probability measure, $\rho$, on $[0, 1]$ iff the linear functional $\mathcal{L}_\alpha$ defined by (A1) for $m = 0, 1, \ldots$ satisfies

$$
\mathcal{L}_\alpha[u^m(1 - u)^{k-m}] \geq 0, \quad m = 0, 1, \ldots, k, \quad k = 0, 1, \ldots.
$$

The measure $\rho$ is unique.

**Corollary 1.** If $\alpha_0 \neq 1$ then, with the modification that $\rho$ is now a finite measure (having total mass $\alpha_0$) on $[0, 1]$, the above result still holds.
The following result is used in the proof of Theorem 1 of this paper.

**Corollary 2.** A sequence \( \{\alpha_1, \alpha_2, \ldots\} \) has a representation

\[
\alpha_n = \int_0^1 u^{n-1} \rho(du), \quad n = 1, 2, \ldots,
\]

where \( \rho \) is a finite measure on \([0, 1]\), iff the linear functional \( L_\alpha \) defined by (A1) for \( m = 1, 2, \ldots \) satisfies

\[
L_\alpha[u^m(1-u)^{k-m}] \geq 0, \quad m = 1, 2, \ldots, k, \ k = 0, 1, \ldots.
\]

The measure \( \rho \) is unique.

Theorem A1 can be proved in various ways. The non-probabilistic proofs of [2, p. 73ff] and [12, pp. 8-9] do give the necessary and sufficient condition in functional form. A probabilistic proof can be given as for Theorem 2 in the first edition of Feller [6, VII.3] but making explicit use of linear functionals. (Notice that the structure and simplicity of the first edition proof has not been preserved in the second edition.) We prefer the latter approach since a proof of Corollary 2 is easily constructed in a similar way, and the approach extends to relevant multivariate situations (Theorems A2 and A3 in the next section).

**A2. Multivariate extensions**

Multivariate moment results are required for Section 3 of this paper. These are a natural generalisation of the univariate results in the same way as a multinomial distribution is a natural generalisation of a binomial distribution. Another type of generalisation, to products of binomial distributions, is considered in Shohat and Tamarkin [12, p. 9].

Let \( Q \) denote now the vector space of all real coefficient polynomials in the \( r \) variables \( v_1, \ldots, v_r \) where \( v = (v_1, \ldots, v_r) \in F_r = \{ v \in [0, 1]^r : v_+ \leq 1 \} \) and \( v_+ = v_1 + \cdots + v_r \). Further, let \( A = \bigcup_{m=0}^\infty A_m \) where \( A_m = \{ a \in \mathbb{Z}^r_+ : a_+ = m \} \) with \( \mathbb{Z}^r_+ = \{0, 1, \ldots\} \). Set \( A_0 = A \cup \{0\} \), and \( v^a = v_1^{a_1}v_2^{a_2} \cdots v_r^{a_r} \).

For a given collection \( \{\alpha_a : a \in A^0\} \) define a linear functional, \( L_\alpha \), on \( Q \) starting from

\[
L_\alpha[v^a] = \alpha_a, \quad v \in F_r,
\]

for \( a \in A^0 \), or on the obvious subspace of \( Q \) starting from (A2) for \( a \in A \), and extending by linearity. The following result extends Theorem A1 to this setting.

**Theorem A2.** A collection \( \{\alpha_a : a \in A^0\} \) of real numbers with \( \alpha_0 = 1 \) is the set of moments of some probability measure, \( \rho \), on \( F_r \), iff the linear functional defined by (A2) for \( a \in A^0 \) satisfies

\[
L_\alpha[v^a(1-v_+)^{k-a_+}] \geq 0, \quad a \in A, \ a_+ \leq k, \ k = 0, 1, \ldots.
\]

The measure \( \rho \) is unique.
Corollary. If $\alpha_0 \neq 1$ then, with the modification that $\rho$ is now a finite measure (with total mass $\alpha_0$) on $\mathcal{F}_r$, the above result still holds.

A proof of Theorem A2 can be constructed along the lines of Theorem 2 in the first edition of Feller [6, VII.3]. A different approach can be based on Bernstein polynomials on $\mathcal{F}$. (This is suggested, in the case $r = 2$, in [6, Chapter VII, Problem 3].)

In Section 3 we need the following analogue of Corollary 2 of Theorem A1.

**Theorem A3.** A collection $\{\beta_a : a \in \mathcal{A}\}$ of real numbers has a representation

$$\beta_a = \sum_{j=1}^{r} \lambda_j \beta_{a,j} + \int_{\mathcal{F}_r \setminus \{o\}} v^a \Lambda(dv), \quad a \in \mathcal{A}, \quad (A3)$$

where, for $j = 1, \ldots, r$, $0 \leq \lambda_j < \infty$ and $1_j$ denotes the $r$-vector with a unit in the $j$th position but zeros elsewhere, and where $\Lambda$ is a not necessarily finite measure on $\mathcal{F}_r \setminus \{o\}$ and satisfying

$$\int_{\mathcal{F}_r \setminus \{o\}} v_j \Lambda(dv) < \infty, \quad (A4)$$

iff the linear functional $\mathcal{L}$ defined on the vector space of all $r$-variable polynomials in $v^a$, with $v \in \mathcal{F}_r$, by

$$\mathcal{L}[v^a] = \beta_a, \quad a \in \mathcal{A}, \quad (A5)$$

satisfies

$$\mathcal{L}[v^a (1 - v_+)^k] \geq 0, \quad a \in \mathcal{A}, \quad a_+ \leq k, \quad k = 1, 2, \ldots. \quad (A6)$$

The constants $\lambda_1, \ldots, \lambda_r$ and the measure $\Lambda$ are unique.

**Proof.** Suppose that (A6) is satisfied. Define linear functionals $\mathcal{L}_j$, $j = 1, 2, \ldots, r$, on $\mathcal{F}_r$ by

$$\mathcal{L}_j[v^a] = \mathcal{L}[v^a v_j], \quad a \in \mathcal{A}^0, \quad j = 1, \ldots, r.$$  

Then, for each $j$,

$$\mathcal{L}_j[v^a (1 - v_+)^k] \geq 0, \quad a \in \mathcal{A}, \quad a_+ \leq k, \quad k = 0, 1, \ldots.$$  

By the corollary to Theorem A2, for each $j$ there exists a finite measure $\rho_j$ on $\mathcal{F}_r$ such that

$$\mathcal{L}_j[v^a] = \int_{\mathcal{F}_r} v^a \rho_j(dv), \quad a \in \mathcal{A}^0.$$  

Since

$$\mathcal{L}[v^a v_j] = \int_{\mathcal{F}_r} v^a v_j \rho_j(dv) = \int_{\mathcal{F}_r} v^a \rho_j(dv)$$
for \( a \in \mathcal{A}^0 \) and \( i, j \in \{ 1, \ldots, r \} \), and since a finite measure on a bounded set is uniquely determined by its moments, it follows that

\[
v_i \rho_j (dv) = v_j \rho_i (dv), \quad i, j \in \{1, \ldots, r\}.
\]

(A7)

From (A7) we deduce that

\[
\rho_i (\{ v \in \mathcal{H}_r : v_i = 0, v_+ > 0 \}) = 0, \quad i \in \{1, \ldots, r\},
\]

(A8)

and, for each pair \( i, j \in \{1, \ldots, r\} \),

\[
v_i^{-1} \rho_j (dv) = v_j^{-1} \rho_i (dv)
\]

on the set \( \{ v \in \mathcal{H}_r : v_0 > 0, v_+ > 0 \} \). Further, for \( a \in \mathcal{A} \) such that \( a_i \geq 1 \),

\[
\mathcal{P}[v^a] = \int_{\mathcal{H}_r} v^{a-1} \rho_j (dv) = \delta_a \rho_j (\{o\}) + \int_{\mathcal{H}_r \setminus \{o\}} v^a v_i^{-1} \rho_i (dv)
\]

where, because of (A8), the last integral need to be taken only over \( \mathcal{H}_r \setminus \{v : v_i > 0\} \).

In view of (A8) and (A9) a measure \( \Lambda \) on \( \mathcal{H}_r \setminus \{o\} \) can be defined unambiguously by specifying the restrictions \( \Lambda^E, 0 \subset E \subset \{1, \ldots, r\} \) of \( \Lambda \) to the respective sets

\[
\mathcal{H}_r^E = \{ v \in \mathcal{H}_r : v_j > 0, j \in E; v_k = 0, k \notin E; j, k \in \{1, \ldots, r\} \}
\]

as

\[
\Lambda^E (dv) = v_j^{-1} \rho_j (dv), \quad v \in \mathcal{H}_r^E,
\]

for some \( j \in E \). Further, define constants \( \lambda_1, \ldots, \lambda_r \) by

\[
\lambda_j = \rho_j (\{o\}), \quad j = 1, \ldots, r.
\]

Then for \( a \in \mathcal{A} \) we have the representation

\[
\mathcal{P}[v^a] = \sum_{j=1}^r \lambda_j \delta_a \rho_j (dv) + \int_{\mathcal{H}_r \setminus \{o\}} v^a \Lambda (dv).
\]

Notice that, although \( \Lambda \) is not a finite measure, we do have

\[
\int_{\mathcal{H}_r \setminus \{o\}} v_i \Lambda (dv) < \infty, \quad i = 1, \ldots, r,
\]

or equivalently

\[
\int_{\mathcal{H}_r \setminus \{o\}} v_+ \Lambda (dv) < \infty. \quad \square
\]

References


