Multivariate Poisson-Charlier, Meixner and Hermite-Chebycheff polynomials and Lancaster distributions

Robert Griffiths

December 15, 2014

Address
Professor R. C. Griffiths
Department of Statistics,
University of Oxford,
1 South Parks Rd, Oxford, OX1 3TG, UK
email: griff@stats.ox.ac.uk.

Abstract
Multivariate Poisson-Charlier, Meixner and Hermite-Chebycheff orthogonal polynomials are constructed from the classical 1-dimensional orthogonal polynomials and multivariate Krawtchouk polynomials. This paper studies Lancaster characterizations of bivariate distributions which have these orthogonal polynomials as eigenfunctions. The characterizations extend classical Lancaster characterizations of bivariate 1-dimensional bivariate Poisson, negative binomial and normal distributions. Transition functions of discrete and continuous time Markov chains with these polynomials as eigenfunctions are characterized.

Keywords: Canonical correlations, Lancaster distributions, multivariate Krawtchouk polynomials, multivariate Poisson-Charlier polynomials, multivariate Meixner polynomials, multivariate Hermite-Chebycheff polynomials, random elements in common.

Running Head: Multi-dimensional Lancaster distributions.
1. Introduction

Griffiths (1971) and Diaconis and Griffiths (2014) construct multivariate Krawtchouck polynomials orthogonal on the multinomial distribution and study their properties. Recent representations and derivations of orthogonality of these polynomials are in Genest et. al. (2013); Grunbaum and Rahman (2011); Iliev (2012); Mizukawa (2011). The multivariate Krawtchouk polynomials are eigenfunctions in classes of reversible composition Markov chains which have multinomial stationary distributions (Zhou and Lange, 2009). Griffiths (1975); Iliev (2012a); Genest et. al. (2014) construct and study multivariate Meixner polynomials, orthogonal on the multivariate Meixner distribution. Genest et. al. (2014a) construct multivariate Poisson-Charlier polynomials on a d-dimensional Poisson product distribution. These polynomials are also mentioned briefly as limits from the multivariate Krawtchouk and Meixner polynomials in Griffiths (1971, 1975). The three multivariate polynomial types are linked through the multivariate Krawtchouk polynomials. Authors emphasise different approaches to the multivariate orthogonal polynomials. Diaconis and Griffiths’s approach is probabilistic and directed to Markov chain applications; Iliev’s approach is via Lie groups; and Genest et. al.’s physics approach is as matrix elements of group representations on oscillator states. Xu (2013) studies discrete multivariate orthogonal polynomials which have a triangular construction of products of 1-dimensional orthogonal polynomials. These are particular cases of the polynomials studied in this paper. See Diaconis and Griffiths (2012) for the connection between the systems in multivariate Krawtchouk polynomials. In this paper multivariate Poisson-Charlier, multivariate Meixner and multivariate Hermite-Chebycheff polynomials are considered in a probabilistic approach.

A bivariate random vector $(X,Y)$ with marginal distributions $f(x)$ and $g(y)$ has a Lancaster probability distribution expansion $p(x,y)$ if

$$p(x,y) = f(x)g(y)\left\{1 + \sum_{n \geq 1} \rho_n P_n(x)Q_n(y)\right\}, \quad (1)$$

for $x$ and $y$ in the support of $X$ and $Y$, where $\{P_n(x)\}, \{Q_n(y)\}$ are orthonormal bases on $f$ and $g$, with $P_0 = Q_0 \equiv 1$. $\{\rho_n\}_{n=0}^\infty$ ($\rho_0 = 1$) is a correlation sequence and satisfies $E[P_m(X)Q_n(Y)] = \delta_{mn}\rho_n$. The problem of characterizing correlation sequences $\{\rho_n\}$ such that a bilinear series (11) is a non-negative sum and thus a bivariate probability distribution has become known as a Lancaster problem after Lancaster and colleagues who studied such bivariate distributions (Eagleson, 1964; Griffiths, 1969; Koudou, 1996).
1998; [Lancaster, 1969]). Usually \( f = g \) and \((X, Y)\) is exchangeable in these characterizations.

A discrete time reversible Markov chain can be constructed with transition functions

\[
p(y \mid x) = f(y) \left( 1 + \sum_{n \geq 1} \rho_n P_n(x)P_n(y) \right).
\]

The \( k \)-step transition functions are then

\[
p^{(k)}(y \mid x) = f(y) \left( 1 + \sum_{n \geq 1} \rho_n^k P_n(x)P_n(y) \right).
\]

A continuous time reversible Markov chain can be constructed as a Poisson imbedding of the discrete time Markov chain, whose transition functions are, for \( t \geq 0 \),

\[
\sum_{k=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^k}{k!} p^{(k)}(y \mid x) = f(y) \left( 1 + \sum_{n \geq 1} e^{-\lambda (1-\rho_n) P_n(x)P_n(y)} \right) \quad (2)
\]

[Bochner (1954)] shows that the most general continuous time reversible transition functions with stationary distribution \( f(y) \) and eigenvectors \( \{P_n(y)\} \) have eigenvalues of the form

\[
\rho_n(t) = e^{-\lambda t(1-\rho_n)}
\]

or a more general limit form obtained when \( \lambda \to \infty, \rho_n \to 1 \), with \( \lambda(1-\rho_n) \) convergent.

In the gamma, Poisson, and negative binomial classes with exchangeability and orthogonal polynomial bases, where the random variables are non-negative, a necessary and sufficient condition that (1) is a probability distribution is

\[
\rho_n = \int_0^1 z^n \varphi(dz), \quad (3)
\]

where \( \varphi \) is a probability measure on \([0,1]\). [Griffiths, 1969; Koudou, 1996, 1998; Sarmanov, 1968]. The class of bivariate distributions is a convex set with extreme correlation sequence points \( \{z^n\} \) for \( z \in [0,1] \). \( X \) and \( Y \) are independent if \( z = 0 \) and \( X = Y \) if \( z = 1 \). The bivariate Poisson characterization is equivalent to \( X = U + V, Y = W + V \), where \( U, V, W \) are
independent Poisson random variables with means $\mu(1 - Z)$, $\mu Z$, $\mu(1 - Z)$ conditional on a random variable $Z$ with distribution $\varphi$ of (3). In the normal class with exchangeability and an orthogonal polynomial basis

$$\rho_n = \int_{-1}^{1} z^n \varphi(dz),$$

(4)

where $\varphi$ is a probability measure on $[-1, 1]$ (Sarmanov and Bratoeva, 1967). The extreme sequence when $z = 0$ corresponds to independence of $X$ and $Y$ and $z = \pm 1$ corresponds to $X = \pm Y$. The characterization is equivalent to $(X,Y)$ having a standard bivariate normal distribution with correlation $Z$, conditional on $Z$, where $Z$ has distribution $\varphi$ of (4). For a general introduction to Lancaster expansions see Lancaster (1969); Koudou (1996).

Bochner characterizations of continuous time Markov chains with polynomial eigenfunctions and stationary distributions which are Poisson, negative binomial, gamma and normal are studied in Griffiths (2009). Eagleson (1969) characterized the correlation sequences $\{\rho_n\}_{n=0}^{N}$ in a bivariate binomial $(N,p \geq 1/2)$ distribution as having a representation $\rho_n = \mathbb{E}[Q_n(Z)]$, where $\{Q_n\}$ are Krawtchouk polynomials scaled so that $Q_n(0) = 1$ and $Z$ is a random variable on $0, 1, \ldots, N$. This characterization with $X$ having finite support is distinct from the moment characterizations (3) and (4). It relies on a hypergroup property of the Krawtchouk polynomials that there exists a random variable $W$ on $1, \ldots, N$ with distribution $\varphi_{xy}$ such that

$$Q_n(x)Q_n(y) = \mathbb{E}_{\varphi_{xy}}[Q_n(W)].$$

(5)

Diaconis and Griffiths (2012) study this characterization further and find another characterization in terms of bivariate binomial distributions where there are $N$ pairs of trials with random correlations. A nice connection is made with generalized Ehrenfest urn models. There is a general theory of orthogonal functions which have the hypergroup property (5) (Bakry and Huet, 2008). Let $\{p_j\}_{j=1}^{d}$ be a discrete probability distribution on $d$ points. $\{u^{(l)}\}_{l=0}^{d-1}$ ($\equiv u$) will denote a complete set of orthogonal functions with $u^{(0)} \equiv 1$, such that for $k, l = 0, 1, \ldots, d - 1$,

$$\sum_{j=1}^{d} u_j^{(k)} u_j^{(l)} p_j = \delta_{kl} \alpha_k.$$  

(6)

$u$ satisfies a hypergroup property if

$$u_j^{(r)} u_k^{(r)} = \sum_{l=1}^{d} b_{jk}(l) u_l^{(r)},$$

(7)
where \( \{b_{jk}(l)\}_{i=1}^{d} \) is a probability distribution for each \( j, k = 1, \ldots, d \). It is supposed that the orthogonal functions can be scaled so that \( u_{d}^{(r)} = 1 \), \( r = 0, \ldots, d - 1 \). (7) is equivalent to

\[
\mathcal{s}(j, k, l) = \sum_{r=1}^{d} a_{r} u_{j}^{(r)} u_{k}^{(r)} u_{l}^{(r)} \geq 0, \quad j, k, l = 1, \ldots, d
\] (8)

and \( b_{jk}(l) = p_{l} \mathcal{s}(j, k, l) \). If an orthogonal basis forms a hypergroup then a Lancaster characterization is that

\[
p_{i} p_{j} \left\{ 1 + \sum_{r=1}^{d-1} \rho_{r} a_{r}^{-1} u_{i}^{(r)} u_{j}^{(r)} \right\}, \quad i, j = 1, \ldots, d
\] (9)

is non-negative and therefore a bivariate probability distribution if and only if \( \rho_{r} = \sum_{i=1}^{d} c_{i} u_{i}^{(r)} \) for a probability distribution \( \{c_{i}\}_{i=1}^{d} \).

The multivariate Krawtchouk, Poisson-Charlier, Meixner and Hermite-Chebycheff polynomials considered in this paper all have an elementary basis \( u \) in their construction. [Diaconis and Griffiths (2014)] show that the multivariate Krawtchouk polynomials satisfy a hypergroup property if and only if the elementary basis does. The multivariate Poisson-Charlier polynomials are orthogonal on the product distribution of \( d \) 1-dimensional Poisson distributions with means \( \mu \). The elementary basis is orthogonal on \( p = \mu/|\mu| \). The multivariate Meixner polynomials are orthogonal on a distribution obtained as a multinomial mixture, with the number of trials having a negative binomial distribution. The multivariate Hermite-Chebycheff polynomials are orthogonal on the product distribution of \( d \) 1-dimensional Normal distributions with means \( 0 \) and variances \( \tau \). The elementary basis \( u \) is orthogonal on \( p = \tau/|\tau| \). Lancaster characterizations are obtained for bivariate distributions which have expansions in multivariate Poisson-Charlier, Meixner and Hermite-Chebycheff polynomials. These characterizations have a similarity to the 1-dimensional ones: in the Poisson case \( X \) and \( Y \) have a structure of random elements in common with random means, but fixed marginal means; in the Normal case \( X \) and \( Y \) have a random cross-correlation matrix, which is an extension of classical canonical correlation theory. A hypergroup property of the elementary basis \( u \) plays an important role in the characterizations. Transition functions of discrete and continuous time Markov chains related to the Lancaster characterizations are constructed. The Markov chains are mixtures of birth and death chains in the Poisson and negative binomial models.
2. Orthogonal polynomials on the multinomial distribution

A brief description of the properties of orthogonal polynomials on the multinomial follows. Define a collection of orthogonal polynomials \( \{ Q_n(X; u) \} \) with \( n = (n_1, \ldots, n_{d-1}) \) and \( |n| \leq N \) on the multinomial distribution

\[
m(x; N, p) = \binom{N}{x} \prod_{j=1}^{d} p_j^{x_j}, \quad x_j \geq 0, \quad j = 1, \ldots, d, \quad |x| = N,
\]

with \( \{ p_j \}_{j=1}^{d} \) a probability distribution, as the coefficients of \( w_1^{n_1} \cdots w_{d-1}^{n_{d-1}} \) in the generating function

\[
G(x, w, u) = \prod_{j=1}^{d} \left( 1 + \sum_{l=1}^{d-1} w_l u_j^{(l)} x_j \right),
\]

where \( \{ u_j^{(l)} \}_{l=0}^{d-1} \) satisfies (6). It is straightforward to show, by using the generating function, that

\[
\mathbb{E}
[Q_m(X; u)Q_n(X; u)] = \delta_{mn} \binom{N}{n, N - |n|}.
\]

Let \( Z_1, \ldots, Z_N \) be independent identically distributed random variables such that

\[
P(Z = k) = p_k, \quad k = 1, \ldots, d.
\]

Then with

\[
X_i = |\{ Z_k : Z_k = i, k = 1, \ldots, d \}|,
\]

\[
G(X, w, u) = \prod_{k=1}^{N} \left( 1 + \sum_{l=1}^{d-1} w_l u_k^{(l)} \right).
\]

From (13)

\[
Q_n(X; u) = \sum_{\{ A_l \}} \prod_{k_1 \in A_1} u_{Z_k}^{(1)} \cdots \prod_{k_{d-1} \in A_{d-1}} u_{Z_k}^{(d-1)},
\]

where summation is over all partitions of subsets of \( \{ 1, \ldots, N \} \), \( \{ A_l \} \) such that \( |A_l| = n_l, \quad l = 1, \ldots, d - 1 \). That is, the orthogonal polynomials are symmetrized orthogonal functions in the tensor product set

\[
\bigotimes_{k=1}^{N} \{ 1, u_{Z_k}^{(i)} \}_{i=1}^{d}.
\]
The orthogonal polynomials could equally well be defined by (14) and the generating function (11) deduced. Let

\[ U_l = \sum_{k=1}^{N} u_k^{(l)} = \sum_{j=1}^{d} u_j^{(l)} X_j, \quad l = 1, \ldots, d - 1. \]

\( Q_n(X; u) \) is a polynomial of degree \( |n| \) in \( (U_1, \ldots, U_{d-1}) \) whose only term of maximal degree \( |n| \) is \( \prod_{k=1}^{d-1} U_k^{n_k} \).

The multivariate Krawtchouk polynomials diagonalize the joint distribution of marginal counts in a contingency table. Suppose \( N \) observations are placed independently into an \( r \times c \) table \( (r \leq c) \) with the probability of an observation falling in cell \((i, j)\) being \( p_{ij} \). Denote the marginal distributions as \( p_r^i = \sum_{j=1}^{c} p_{ij} \) and \( p_c^j = \sum_{i=1}^{r} p_{ij} \). Let \( p_{ij} \) have a Lancaster expansion (which is always possible, even for non-exchangeable \( p_{ij} \))

\[ p_{ij} = p_r^i p_c^j \left\{ 1 + \sum_{k=1}^{r-1} \rho_k u_i^{(k)} v_j^{(k)} \right\}, \quad (15) \]

where \( u \) and \( v \) are orthonormal function sets on \( p_r^r \) and \( p_c^c \). \( u \) is an orthonormal basis and if \( r < c \) complete \( \{v^{(k)}\}_{k=0}^{r-1} \) to a basis \( \{v^{(k)}\}_{k=0}^{c-1} \). Let \( N_{ij} \) be the number of observations falling into cell \((i, j)\) and \( X_i = \sum_{j=1}^{c} N_{ij}, \ Y_j = \sum_{i=1}^{r} N_{ij} \). \( X \) and \( Y \) are the marginal counts. Then

\[ P(X = x, Y = y) = m(x; N, p_r^r) m(y; N, p_c^c) \times \left\{ 1 + \sum_{n} \rho_1^{n_1} \cdots \rho_{r-1}^{n_{r-1}} \binom{N}{n}^{-1} Q_n(x; N, u, p_r^r)Q_{n^*}(y; N, v, p_c^c) \right\}, \quad (16) \]

where \( n^* = (n, 0_{c-r}) \). Aitken and Gonin (1935) showed (16) for a \( 2 \times 2 \) table with the usual 1-dimensional Krawtchouk polynomials and Griffiths (1971) for \( r \times c \) tables.

3. Multivariate Poisson-Charlier polynomials

Let \( X \) be a Poisson random variable with mean \( \lambda \). The Poisson-Charlier orthogonal polynomials on \( X \), \( \{C_n(X; \lambda)\}_{n=0}^{\infty} \), have a generating function

\[ \sum_{n=0}^{\infty} C_n(X; \lambda) \frac{z^n}{n!} = e^{z \left( 1 - \frac{z}{\lambda} \right)^X}. \quad (17) \]
The orthogonality relationship is

\[ \mathbb{E}[C_m(X; \lambda)C_n(X; \lambda)] = \delta_{mn} \frac{m!}{\lambda^m}. \]

Let \( X = (X_1, \ldots, X_d) \) be independent Poisson random variables with means \( \mu = (\mu_1, \ldots, \mu_d) \). The probability distribution of \( X \) is

\[ P(x; \mu) = \prod_{i=1}^{d} e^{-\mu_i} \frac{\mu_i^x_i}{x_i!}. \quad (18) \]

A natural set of multivariate Poisson-Charlier orthogonal polynomials is the product set \( \bigotimes_{j=1}^{d} \{ C_{n_j}(X_j; \mu_j) \} \). Another set is constructed by noting that the sum \( |X| \) is a Poisson random variable with mean \( |\mu| \) and the conditional distribution of \( X \) given \( |X| \) is multinomial with parameters \( (N = |X|, p = \mu/|\mu|) \).

**Theorem 1.** A set of multivariate Poisson-Charlier polynomials orthogonal on \( X \) is

\[ C_n(X; \mu, \mathbf{u}) = (n_0)!^{-1} C_{n_0}(|X| - |n_1|; |\mu|) |\mu|^{-|n_1|} Q_{n_1}(X; |X|, p, \mathbf{u}). \quad (19) \]

where \( n \in \mathbb{Z}_+^d, n_1 = (n_1, \ldots, n_{d-1}) \), and \( Q_{n_1}(X; |X|, p, \mathbf{u}) \) are multivariate Krawtchouk polynomials, orthogonal on the multinomial random variable \( X \) given \( |X| \). \( \mathbf{u} \) is an orthogonal basis satisfying (6). The orthogonality relationship is

\[ \mathbb{E}[C_m(X; \mu, \mathbf{u})C_n(X; \mu, \mathbf{u})] = \delta_{mn} |\mu|^{-|n_1|} \prod_{j=0}^{d-1} a_j^{n_j} : = \delta_{mn} h_n(|\mu|, \mathbf{u})^{-1}. \quad (20) \]

The polynomials (19) are generated by the coefficients of \( w_0^{n_0} \cdots w_{d-1}^{n_{d-1}} \) in

\[ G_{PC}(X, w, \mathbf{u}) = e^{w_0} \prod_{i=1}^{d} \left( 1 - |\mu|^{-1} w_0 + |\mu|^{-1} \sum_{j=1}^{d-1} u_j^{(j)} w_j \right)^{X_i}. \quad (21) \]

**Proof.** To show that the generating function (21) is correct it is sufficient to show that it generates the orthogonal polynomials (19) as coefficients of \( w_0^{n_0} \cdots w_{d-1}^{n_{d-1}} \). The coefficient of \( w_1^{n_1} \cdots w_{d-1}^{n_{d-1}} \) in (21) is

\[ e^{w_0} \left( 1 - |\mu|^{-1} w_0 \right)^{|X| - |n_1|} |\mu_1|^{-|n_1|} Q_{n_1}(X; |X|, p, \mathbf{u}) \quad (22) \]
so (19) holds from (22) and (17). Orthogonality of the polynomials (19) can be shown by direct argument or using the generating function (21).

\[
E[G_{PC}(X, z, u)G_{PC}(X, w, u)] = \exp\left\{ z_0 + w_0 + \sum_{i=1}^{d} \mu_i \left[ \left(1 - |\mu|^{-1} z_0 + |\mu|^{-1} \sum_{j=1}^{d-1} u_i^{(j)} z_j \right) \times \left(1 - |\mu|^{-1} w_0 + |\mu|^{-1} \sum_{j=1}^{d-1} u_i^{(j)} w_j \right) - 1\right] \right\} = \exp\left\{ |\mu|^{-1} \sum_{j=0}^{d-1} a_j z_j w_j \right\}.
\]

(23)

Now (24) follows as the coefficient of \(z_0^{m_0} \cdots z_{d-1}^{m_{d-1}} w_0^{n_0} \cdots w_{d-1}^{n_{d-1}}\) in (23).

**Corollary 1.** The transform

\[
C_n(s; \mu, u) := E\left[ \prod_{j=1}^{d} s_j^{X_j}C_n(X; \mu, u) \right] = \left( \prod_{j=0}^{d-1} n_j! \right)^{-1} e^{\sum_{i=1}^{d} \mu_i (s_i - 1)} \left(1 - S_0\right)^{n_0} \prod_{j=1}^{d-1} S_j^{n_j},
\]

(24)

where \(S_j = \sum_{i=1}^{d} p_i s_i u_i^{(j)}\).

**Proof.**

\[
E\left[ \prod_{j=1}^{d} s_j^{X_j}G_{PC}(X, w, u) \right] = \exp\left\{ w_0 + \sum_{i=1}^{d} \mu_i \left[ s_i \left(1 - |\mu|^{-1} w_0 + |\mu|^{-1} \sum_{j=1}^{d-1} u_i^{(j)} w_j \right) - 1\right] \right\} = \exp\left\{ \sum_{i=1}^{d} \mu_i (s_i - 1) + w_0 (1 - S_0) + \sum_{j=1}^{d-1} S_j w_j \right\}
\]

(25)

Now (24) is the coefficient of \(w_0^{n_0} \cdots w_{d-1}^{n_{d-1}}\) in (58).
Remark 1. A representation of a Poisson random vector $X$ with independent entries and mean $\mu$ is that

$$X = e_{Z_1} + e_{Z_2} + \cdots + e_{Z_{|X|}},$$

where $e_k$ is the $k$th unit vector and $\{Z_i\}_{i=1}^{\infty}$ is a sequence of independent random variables such that $P(Z_i = k) = p_k = \mu_k/|\mu|$, $k = 1, \ldots, d$ and $|X|$ is a Poisson random variable independent of the sequence with mean $|\mu|$. A symmetrized tensor product set, similar to (14) with a random number of trials is then

$$C_n(X; \mu, u) = (n_0)!^{-1}C_n(|X| - |n_1|; |\mu|)^{-|n_1|} \times \sum_{\{A_i\}} \prod_{k_1 \in A_1} u^{(1)}_{Z_{k_1}} \cdots \prod_{k_{d-1} \in A_{d-1}} u^{(d-1)}_{Z_{k_{d-1}}}, \quad (26)$$

where summation is over all partitions of subsets of $\{1, \ldots, |X|\}, \{A_i\}$ such that $|A_i| = n_l$, $l = 1, \ldots, d - 1$. $C_n(X; \mu, u)$ is a polynomial in the sums $U_j = \sum_{i=1}^{d} u_j^{(j)} X_i$ with a single leading term of degree $|n|$ proportional to $U_0^{n_0} \cdots U_{d-1}^{n_{d-1}}$.

Remark 2. It seems natural to include a constant function $u^{(0)} = 1$ in the basis, leading to including the Poisson-Charlier polynomials in $|X|$ in the polynomial set, however it is possible to choose a basis $u$ without a constant function. Then a generating function for the multivariate polynomials is

$$\exp \left\{ - \sum_{j=0}^{d-1} \sum_{i=1}^{d} u_i^{(j)} p_i w_j \right\} \prod_{i=1}^{d} \left( 1 + |\mu|^{-1} \sum_{j=0}^{d-1} u_i^{(j)} w_j \right)^{X_i}. \quad (27)$$

This generating function is in Section 11 of [Genest et al. (2014a)], with a different notation.

Remark 3. The polynomials $\{C_n(X; \mu, u)\}$ can be obtained as a limit from a set of multivariate Krawtchouk polynomials. [Genest et al. (2014a)] show such a limit, which is also very briefly mentioned in [Griffiths (1971)]. Let $X^{(N)}$ be a multinomial $(N, p^{(N)})$ random variable of dimension $d + 1$ with $(p_i^{(N)})_{i=1}^{d} = (N + |\mu|)^{-1} \mu$ and $p_{d+1}^{(N)} = (N + |\mu|)^{-1} N$. Then from a classical limit theorem $(X_i^{(N)})_{i=1}^{d}$ converges to a Poisson random vector $X$ with mean $\mu$. Choose an orthogonal basis for the multivariate Krawtchouk polynomials of $\{v^{(j)}(N)\}_{j=0}^{d}$ with $v^{(0)}(N) = 1$, and

$$v_k^{(j)}(N) = \begin{cases} |\mu|^{-1} u_k^{(j)}, & k = 1, \ldots, d \\ 0, & k = d + 1 \end{cases}$$
where \{u^{(j)}\} is an orthogonal basis on \(\mu\). Take

\[ v_k^{(d)}(N) = \begin{cases} -|\mu|^{-1}, & k = 1, \ldots, d \\ N^{-1}, & k = d + 1. \end{cases} \]

The generating function for the multivariate Krawtchouk polynomials is

\[
\prod_{j=1}^{d+1} \left( 1 + \sum_{l=1}^{d} w_l u_j^{(l)}(N) \right)^{x_j} \\
= \prod_{j=1}^{d} \left( 1 + \sum_{l=1}^{d-1} w_l |\mu|^{-1} u_j^{(l)} - |\mu|^{-1} w_d \right)^{x_j} \left( 1 + N^{-1} w_d \right)^{N^{-1} \sum_{j=1}^{d} x_j} \\
\to e^{w_d} \prod_{j=1}^{d} \left( 1 - |\mu|^{-1} w_d + |\mu|^{-1} \sum_{l=1}^{d-1} w_l u_j^{(l)} \right)^{x_j}. \tag{28}
\]

Change the indexing by setting \(w_d\) to \(w_0\), then (28) is identical to (21). This shows that multivariate Krawtchouk polynomials with basis \(\{v^{(l)}(N)\}\) converge to the multivariate Poisson-Charlier polynomials with basis \(u\).

### 3.1. Bivariate Poisson expansions

The distribution of the marginals in a Poisson array has a diagonal expansion in the multivariate Poisson-Charlier polynomials as a natural extension from a contingency table.

**Theorem 2.** Let \((X_{ij}), i = 1, \ldots, r, j = 1, \ldots, c, r \leq c\) be an \(r \times c\) array of independent Poisson random variables with means \(\mu = (\mu_{ij})\). Let \(p_{ij} = \mu_{ij}/|\mu|\). Suppose there is a Lancaster expansion for \(i = 1, \ldots, r, j = 1, \ldots, c,\)

\[
p_{ij} = p_{i}^{r}p_{j}^{c}\left\{ 1 + \sum_{k=1}^{r-1} \rho_{ik} u^{(k)}_{i} v^{(k)}_{j} \right\}. \tag{29}
\]

where \(u\) and \(v\) are orthonormal functions on \(p^{r}\) and \(p^{c}\) respectively. Denote the marginals of the table by \(X = (X_{i})_{i=1}^{r}, Y = (Y_{j})_{j=1}^{c}\) and their means by \(\mu_{x}, \mu_{y}\). Then

\[
P(X = x, Y = y) = P(x; \mu)P(y; \mu) \\
\times \left\{ 1 + \sum_{n:|n| \geq 1} \rho_{n}^{r} \cdots \rho_{n}^{c-1} h_{n}(|\mu|) C_{n}(x; \mu_{x}, u) C_{n^{*}}(y; \mu_{y}, v) \right\}. \tag{30}
\]

where \(n \in \mathbb{Z}_{+}^{r}, n^{*} = (n, 0_{c-r}) \in \mathbb{Z}_{+}^{c}\) and \(h_{n}(|\mu|) = |\mu|^{n} \prod_{j=0}^{d-1} n_{j}!\).
Proof. The conditional distribution of \((X_{ij})\) given \(|(X_{ij})|\) is that of an \(r \times c\) contingency table with \(|(X_{ij})|\) observations. The conditional distribution of the marginals \((X, Y)\) therefore has a Lancaster expansion \((16)\), where

\[
N = |(X_{ij})| = |X| = |Y|.
\]

Then

\[
\mathbb{E} \left[ C_m(X; \mu_x, u)C_n(Y; \mu_y, v) \right]
= \mathbb{E} \left[ (m_0)!^{-1}C_{m_0}(|X| - |m_1|; |\mu|)(n_0)!^{-1}C_{n_0}(|Y| - |n_1|; |\mu|) \right.
\]
\[
\times \mathbb{E} \left[ Q_{m_1}(X; |X|, p^f, u)Q_{n_1}(Y; |Y|, p^c, v) \right] \bigg{|} (X_{ij}) \bigg{|}
\]
\[
= \mathbb{E} \left[ C_{m_0}(|X| - |m_1|; |\mu|)C_{n_0}(|X| - |n_1|; |\mu|) \left( \frac{|X|}{n_1} \right) \right]
\]
\[
\times \delta_{m_1n_1} \left( (m_0)!^{-1}(n_0)!^{-1}|\mu|^{-|m_1|+|n_1|} \rho_1^{n_1} \ldots \rho_r^{n_r-1} \right)
\]
\[
= (n_0)!^{-1} \mathbb{E} \left[ C_{m_0}(|X|; |\mu|)C_{n_0}(|X|; |\mu|) \right]
\]
\[
\times \delta_{m_1n_1} \left( (n_0!n_1! \ldots n_r! - 1)^{-1}|\mu|^{-|m_1|} \rho_1^{n_1} \ldots \rho_r^{n_r-1} \right)
\]
\[
= \delta_{mn} \left( n_0!n_1! \ldots n_r! - 1 \right)^{-1}|\mu|^{-|n_1|} \rho_1^{n_1} \ldots \rho_r^{n_r-1}
\]
\[
= \delta_{mn} \left( n_0!n_1! \ldots n_r! \right)^{-1}\rho_1^{n_1} \ldots \rho_r^{n_r-1}
\]
(31)

(30) now follows from (31).

\[ \square \]

3.2. Lancaster characterization

**Theorem 3.** Let \( u \) be an orthogonal basis on \( p = \mu/|\mu| \) with \((6)\) holding. Suppose the basis satisfies the hypergroup property \((8)\). Then

\[
P(x; \mu)P(y; \mu) \left\{ 1 + \sum_{|n| \geq 1} \rho_n h_n(|\mu|, u)C_n(x; \mu, u)C_n(y; \mu, u) \right\}
\]
(32)

is non-negative and thus a proper bivariate Poisson distribution if and only if

\[
\rho_n = \mathbb{E}_\xi \left[ \prod_{j=0}^{d-1} \rho_j(\xi)^{n_j} \right],
\]
(33)

where \( \rho_j(\xi) = \sum_{i=1}^{d-1} u_i^{(j)} \xi_i \) and \( \xi \) is a random vector in the unit simplex.
Proof. Necessity. Suppose (32) holds. Let \((X, Y)\) be a pair of Poisson random vectors with distribution (32). Then

\[
\rho_n C_n(x; \mu, u) = \mathbb{E}[C_n(Y; \mu, u) \mid X = x].
\]  (34)

C_n(x; \mu, u) has only one leading term of degree \(|n|\), proportional to

\[
m(n, x, u) = \prod_{j=0}^{d-1} \left( \sum_{i=1}^{d-1} u_i^{(j)} x_i \right)^{n_j}.
\]

Rearranging (34)

\[
\mathbb{E}[m(n, Y, u) \mid X = x] = \rho_n m(n, x, u) + R_{|n|} 1(x)  \quad (35)
\]

where \(R_{|n|} 1(x)\) is a polynomial of degree \(|n|\) in \(x\). Divide (35) by \(|x|^{|n|}\) and let \(|x| \to \infty\) such that \(x_d/|x| \to 1\) and \(x_j/|x| \to 0\), 1 \(\leq j < d\). Let \(\xi\) be a random variable with the limit distribution of \(Y/|x|\) given \(X = x\). The limit is taken through a subsequence such that the limit distribution is proper. Recalling that \(u_d^{(j)} = 1\) for \(j = 0, \ldots, d\), (33) holds from the limit. \(\rho_0(\xi) = |\xi|\) and with a particular \(n' = (n, 0, \ldots, 0)\), \(\rho_n = \mathbb{E}_\xi[|\xi|^{|n|}]\). Since \(|\rho_{n'}| \leq 1\) for all such \(n'\), \(|\xi|\) must belong to the unit simplex.

Sufficiency. Suppose that (33) holds. Let \(S_j = \sum_{i=1}^{d} p_i s_i u_i^{(j)}, j = 0, \ldots, d - 1\) and similarly for \(T_j\). Note that \(S_0 = \sum_{i=1}^{d} p_i s_i, T_0 = \sum_{i=1}^{d} p_i t_i\). The potential probability generating function (pgf) of a bivariate distribution \((X, Y)\) formed from (32) is

\[
\sum_{n:|n| \geq 0} \rho_n h_n(|\mu|, u) C_n^*(s; \mu, u) C_n^*(t; \mu, u)
= \exp \{ |\mu| (S_0 - 1 + T_0 - 1) \}
\times \mathbb{E}_\xi \left[ \exp \left\{ |\mu| |\xi| (1 - S_0) (1 - T_0) + |\mu| \sum_{k=1}^{d-1} a_k^{-1} \rho_k(\xi) S_k T_k \right\} \right]
= \exp \{ |\mu| (S_0 - 1 + T_0 - 1) \}
\times \mathbb{E}_\xi \left[ \exp \left\{ |\mu| |\xi| (1 - S_0) (1 - T_0) - S_0 T_0 + |\mu| \sum_{k=0}^{d-1} a_k^{-1} \rho_k(\xi) S_k T_k \right\} \right]
= \mathbb{E}_\xi \left[ \exp \left\{ |\mu| (1 - |\xi|) (S_0 - 1 + T_0 - 1) \right\} \right]
\times \exp \left\{ |\mu| \sum_{j,k,l=1}^{d} p_j p_k \xi s(j,k,l)(s_j t_k - 1) \right\} \right] \quad (36)
\]
The identity \( \sum_{j,k=1}^{d} p_j p_k \varepsilon(j, k, l) = 1 \), for \( l = 1, \ldots, d \) is used in the last line. (36) is a pgf because \( \varepsilon(j, k, l) \geq 0 \), completing the sufficiency proof. 

**Remark 4.** \( X \) and \( Y \) are independent if \( \rho_n = 0 \) for all \( |n| \geq 1 \). \( X = Y \) if \( \rho_n = 1 \) for \( |n| \geq 1 \), achieved when \( \xi_1 = \cdots = \xi_{d-1} = 0 \) and \( \xi_d = 1 \).

**Remark 5.** The structure of \( (X, Y) \) as marginals of a Poisson array with random means is clear from (36). 

\[
X_j =^D Z_j + \sum_{k=1}^{d} Z_{jk}, Y_k =^D Z_k' + \sum_{j=1}^{d} Z_{jk}. \tag{37}
\]

where \( Z, Z' \) and \( (Z_{jk}) \) are Poisson random variables with random means depending on \( \xi \),

\[
E[Z | \xi] = E[Z' | \xi] = (1 - |\xi|) \mu
\]

and

\[
E[Z_{jk} | \xi] = |\mu| p_j p_k \sum_{i=1}^{d} \xi_i \varepsilon(j, k, l)
= |\mu| |\xi| p_j p_k \left\{ 1 + \sum_{r=1}^{d-1} \theta_r \alpha_i^{-1} u_j^{(r)} u_k^{(r)} \right\}
= |\mu| |\xi| p_j p_k(\xi), \tag{38}
\]

where \( \theta_r = \sum_{l=1}^{d} \omega_l u_l^{(r)} \), with \( \omega_l = \xi_l / |\xi| \) and

\[
p_{jk}(\xi) = p_j p_k \left\{ 1 + \sum_{r=1}^{d-1} \theta_r \alpha_i^{-1} u_j^{(r)} u_k^{(r)} \right\} \tag{39}
\]

is a bivariate probability distribution. (38) and (39) are non-negative because of the hypergroup property satisfied by the orthogonal basis.

If \( |\xi| \equiv 1 \) then the structure is that of marginal distributions in a Poisson table as in Theorem 2.

**Remark 6.** The conditional pgf of the extreme point distribution \( Y | X = x \) is

\[
\exp \left\{ |\mu|(1 - |\xi|)(T_0 - 1) \right\} \prod_{i=1}^{d} \left( 1 - |\xi| + \sum_{j,d=1}^{d} \xi_j p_j \varepsilon(i, j, l) t_j \right)^{x_i}
= \exp \left\{ |\mu|(1 - |\xi|)(T_0 - 1) \right\} \prod_{i=1}^{d} \left( 1 - |\xi| + |\xi| \sum_{j=1}^{d} p_{ji}(\omega) t_j \right)^{x_i}, \tag{40}
\]

14
where \( p_{j|i}(\omega) = p_{ij}(\omega)/p_i \) is a conditional probability distribution. Both factors in (40) are pgfs.

**Remark 7.** An interpretation of the extreme points in the conditional distribution of \( Y \mid X \) with a pgf (40) is as the transition functions in a discrete time Markov chain. Consider a queueing system with an infinite number of servers and Poisson arrivals. The servers are of \( d \) possible types. Customers arrive at rate \( |\mu|(1 - |\xi|) \) and choose an available server of type \( i \) with probability \( p_i \). Each queue has length zero or one. At a time epoch each of the \( X_i \) customers in type \( i \) queues is either served with probability \( 1 - |\xi| \) or changes to an available server of type \( j \) with probability \( |\xi|p_{j|i}(\omega) \) and is not served in the current epoch. This describes the transition from the numbers \( X \) in the queues of \( d \) types to \( Y \). The stationary distribution of the Markov chain is Poisson with mean \( \mu \). The general transition functions are a mixture of extreme point transition functions over a distribution for \( \xi \).

**Remark 8.** The distribution of \( (X, Y) \) is asymptotically normal as \( \mu \to \infty \), independent within vectors, with means and variances \( \mu \), and random cross covariance matrix

\[
|\mu||\xi|p_{jk}(\omega).
\]

**Remark 9.** The general form of transition functions from \( x \) to \( y \) in time \( t \) in a continuous time reversible Markov process \( \{X(t)\}_{t \geq 0} \) with a stationary multivariate Poisson distribution and multivariate Poisson-Charlier polynomial eigenfunctions has an expansion

\[
P(y; \mu) \left\{ 1 + \sum_{|n| \geq 1} \rho_n(t) h_n(|\mu|, u) C_n(x; \mu, u) C_n(y; \mu, u) \right\}
\]

where

\[
\rho_n(t) = \exp \left\{ -\lambda t \left[ 1 - \rho_n \right] \right\}
\]

\[
= \exp \left\{ -\lambda t \mathbb{E}_\xi \left[ 1 - \prod_{j=0}^{d-1} \rho_j(\xi)^{n_j} \right] \right\},
\]

or a limit form, from [Bochner (1954)].

We consider a simpler particular case. Recall the notation \( \rho_k(\xi) = \sum_{i=1}^d u_i^{(k)} \xi_i \). The idea is to choose \( \lambda, \xi \) so that \( \rho_j(\xi) \to 1 \) and

\[
\lambda \mathbb{E}_\xi \left[ 1 - \prod_{j=0}^{d-1} \rho_j(\xi)^{n_j} \right] \approx \lambda \sum_{j=0}^{d-1} n_j \left( 1 - \rho_j(\mathbb{E}[\xi]) \right).
\]
To achieve this set $E[\xi_i] = \lambda^{-1} \gamma_i$, $i = 1, \ldots, d$ and $1 - E[\xi_d] = \sum_{i=1}^{d-1} \lambda^{-1} \gamma_i + \lambda^{-1} \nu$, where $\gamma, \nu \geq 0$. Then $1 - \rho_0(E[\xi]) = 1 - E[|\xi|] = \lambda^{-1} \nu$ and recalling that $u_d^{(j)} = 1$,

$$1 - \rho_j(E[\xi]) = 1 - \sum_{i=1}^{d-1} u_i^{(j)} E[\xi_i] - E[\xi_d]$$

$$= \lambda^{-1} \left( \nu + \sum_{i=1}^{d-1} \gamma_i (1 - u_i^{(j)}) \right).$$

Assume also that $\lambda \text{Var}(\rho_k(\xi)) \to 0$. Because of the hypergroup property of $u$, $|u_i^{(j)}| \leq 1$, so terms in the sum are non-negative. Let $\lambda \to \infty$. The eigenvalues have the limit form

$$\rho_n(t) = \exp \left\{ -t |n| \nu + \sum_{j=1}^{d-1} n_j \theta_j(\gamma) \right\}, \quad (44)$$

where $\theta_j(\gamma) = \sum_{i=1}^{d-1} \gamma_i (1 - u_i^{(j)})$. Now consider the infinitesimal pgf for transitions in $(t, t + \Delta t)$ as $\Delta t \to 0$, that is, $E \left[ \prod_{k=1}^d X_k(t+\Delta t) \mid X(t) = x \right]$. Only the linear terms in $\rho_n(\Delta t)$ need to be considered.

$$\rho_n(\Delta t) = \exp \left\{ -\Delta t \left( |n| \nu + \sum_{j=1}^{d-1} n_j \theta_j(\gamma) \right) \right\}$$

where $\nu = E[\nu]$, $\gamma = E[\gamma]$ and

$$\tilde{\rho}_k(E[\xi]) := \exp \left\{ -\Delta t (\nu + \theta_k(\gamma)) \right\} \approx 1 - \Delta t (\nu + \theta_k(\gamma)). \quad (45)$$

The infinitesimal pgf is now obtained by substituting the eigenvalues (45) in (40). Calculating the quantities appearing:

$$E[1 - |\xi|] = 1 - \tilde{\rho}_0(E[\xi]) \approx 1 - e^{-\nu \Delta t} \approx \nu \Delta t,$$

and

$$E[|\xi|p_{j|i}(\omega)] \approx p_j \sum_{k=0}^{d-1} \Delta t (\nu + \theta_k(\gamma)) a_k^{-1} u_i^{(k)} u_j^{(k)}$$

$$= \delta_{ij} (1 - \Delta t (\nu + |\gamma|)) + \Delta t p_j \sum_{l=1}^{d-1} \gamma_l s(i, j, l)$$

$$= \delta_{ij} (1 - \Delta t (\nu + |\gamma|)) + \Delta t |\gamma| \tilde{p}_{j|i}, \quad (46)$$
where
\[
\tilde{p}_{j|i} := p_j \sum_{l=1}^{d-1} \left( \frac{\gamma_l}{|\gamma|} \right) s(i, j, l)
\]
is a transition matrix with stationary distribution \( p \). Then (40) is, to order \( \Delta t \),
\[
\left( 1 - \Delta t |\mu| \nu \sum_{i=1}^{d} p_i t_i \right) \\
\times \prod_{i=1}^{d} \left( (1 - \Delta t (\nu + |\gamma|) t_i + \Delta t \nu + \Delta t \sum_{j=1}^{d} |\gamma| \tilde{p}_{j|i} t_j) \right)^{x_{i}}.
\]
(47)
\{X(t)\}_{t \geq 0} is now identified as a multi-type birth and death process from (47). Of course only the terms to order \( \Delta t \) matter, but the \text{pgf} is clearer in the current form. Thinking of \( X(t) \) as counts of individuals in a population at time \( t \); immigration occurs at rates \( |\mu| \nu \); deaths occur to individuals at rate \( \nu \); and changes of type from \( i \) to \( j \) occur at rates \( \gamma \tilde{p}_{j|i} \). The stationary distribution is Poisson \( (\mu) \). \|X(t)\| is a conventional birth and death process with birth rates when \( \|X(t)\| = x \) of \( |\mu| \nu \) and death rates \( |x| \nu \).

4. Multivariate Meixner polynomials

Let \( X \) be a negative binomial random variable with probability distribution
\[
\frac{\Gamma(\alpha + x)}{\Gamma(\alpha) x!} \frac{\theta^x}{(1 + \theta)^{x+\alpha}}, \quad x = 0, 1, \ldots, \alpha, \theta > 0.
\]
(48)
The Meixner orthogonal polynomials on \( X \), \( \{M_n(X; \alpha, \gamma)\}_{n=0}^{\infty} \) have a generating function
\[
\sum_{n=0}^{\infty} \frac{\Gamma(\alpha + n)}{\Gamma(\alpha) n!} M_n(x; \alpha, \kappa) z^n = (1 - z/\kappa)^x (1 - z)^{-(x+\alpha)},
\]
(49)
with \( \kappa = \theta/(1 + \theta) \). The orthogonality relationship for the 1-dimensional Meixner polynomials is
\[
\mathbb{E}[M_m(X; \alpha, \kappa) M_n(X; \alpha, \kappa)] = \delta_{mn} \frac{\Gamma(\alpha) n!}{\Gamma(\alpha + n) \kappa^n}.
\]
(50)
The multivariate Meixner distribution is
\[
P(x; \alpha, \theta, p) = \frac{\Gamma(\alpha + |x|)}{\Gamma(\alpha) |x|!} \frac{\theta^{|x|}}{(1 + \theta)^{\alpha + |x|}} \left( \frac{|x|}{x} \right)^d p_i^{x_i}.
\]
(51)
for \( x \in \mathbb{Z}_+^d \). The conditional distribution of \( X \) given \(|X|\) is \( m(x; |x|, p) \).

The pgf of (51) is

\[
1 - \theta \sum_{i=1}^{d} p_i (s_i - 1)^{-\alpha}. \tag{52}
\]

Let

\[
P(x; \mu, p) = \prod_{i=1}^{d} e^{-\mu p_i} (\mu p_i)^{x_i}/x_i!.
\]

The multivariate Meixner distribution is a Poisson-Gamma mixture

\[
P(x; \alpha, \theta, p) = E_{\mu} [ P(x; \mu, p)] , \tag{53}
\]

where \( \mu \sim \text{Gamma} (\alpha, \theta) \) with density

\[
\frac{\mu^{\alpha-1} e^{-\mu/\theta}}{\theta^\alpha \Gamma(\alpha)}, \mu > 0.
\]

**Theorem 4.** Let \( X \) have a multivariate Meixner distribution \( P(x; \alpha, \theta, p) \). A set of multivariate Meixner polynomials, orthogonal on \( X \) is

\[
M_n(X; \alpha, \theta, p, u) = \frac{\Gamma(\alpha + n_0)}{\Gamma(\alpha) n_0!} M_{n_0}(|X| - |n_1|; \alpha, \theta/(1 + \theta)) Q_{n_1}(X; |X|, p, u), \tag{54}
\]

where \( n \in \mathbb{Z}_+^d \), \( n_1 = (n_1, \ldots, n_{d-1}) \), and \( Q_{n_1}(X; |X|, p, u) \) are multivariate Krawtchouk polynomials, orthogonal on the multinomial random variable \( X \) given \(|X|\). \( u \) is an orthogonal basis satisfying (6). The orthogonality relationship is

\[
E [ M_m(X; \alpha, \theta, p, u) M_n(X; \alpha, \theta, p, u) ] = \delta_{mn} \frac{\Gamma(|n| + \alpha)}{\Gamma(n_0!) \cdots n_{d-1}!} \left( \frac{1 + \theta}{\theta} \right)^{n_0} \theta^{(|n|-n_0)} \prod_{j=1}^{d-1} a_j^{n_j}. \tag{55}
\]

The multivariate Meixner polynomials are coefficients of \( w_0^{n_0} \ldots w_{d-1}^{n_{d-1}} \) in the generating function [Griffiths, 1975]

\[
G_M(X, w, \alpha, \theta, u) = (1-w_0)^{-|X|+\alpha} \prod_{i=1}^{d} \left( 1-w_0 \frac{1 + \theta}{\theta} + \sum_{j=1}^{d-1} u_i^{(j)} w_j \right)^{X_i}. \tag{56}
\]
These polynomials have a representation as a mixture of the multivariate Poisson-Charlier polynomials:

\[
P(X; \alpha, \theta, p) \left( \frac{\theta}{1 + \theta} \right)^{n_0} M_n(X; \alpha, \theta, p, u) = E_{\mu} \left[ P(X; \mu, p)^{|n|} C_n(X; \mu, u) \right],
\]

where \( \mu \equiv |\mu| \).

**Corollary 2.** The transform of the multivariate Meixner polynomials is

\[
M^*_n(s; \alpha, \theta, p, u) := E \left[ \prod_{j=1}^d s_i X_i M_n(X; \alpha, \theta, p, u) \right]
\]

\[
= \frac{\Gamma(\alpha + |n|)}{\Gamma(\alpha)n_0! \cdots n_{d-1}!} \left[ 1 - \theta(S_0 - 1) \right]^{-(\alpha + |n|)}
\]

\[
\times (1 + \theta)^{n_0} (1 - S_0)^{n_0 \theta - n_0} \prod_{j=1}^{d-1} S_{nj}.
\]

The transform of the 1-dimensional Meixner polynomials is

\[
M^*_n(s; \alpha, \kappa) := E \left[ s^X M_n(X; \alpha, \kappa) \right] = (1 + \theta)^n (1 - s)^n (1 - \theta(s - 1))^{-(\alpha + n)}.
\]

The proofs in Theorem 4 and Corollary 2 follow those for the multivariate Poisson-Charlier polynomials so are omitted for brevity.

**Remark 10.** The multivariate Poisson-Charlier polynomials can be obtained as a limit from the multivariate Meixner polynomials. In the generating function (57) set \( w_0 = w'_0 \theta / |\mu| \), and \( w'_j = w_j / |\mu|, j = 1, \ldots, d-1 \). Now let \( \alpha \to \infty, \theta \to 0 \) with \( \alpha \theta \to |\mu| \). The generating function for the multivariate Meixner polynomials (57) converges to the generating function for the multivariate Poisson polynomials (21) with dummy variables \( w' \). Therefore

\[
M_n(X; \alpha, \theta, p, u) \theta^{n_0} |\mu|^{-|n|} \to C_n(X; \mu, u).
\]

A characterization of bivariate multivariate negative binomial distributions whose eigenfunctions are the multivariate Meixner polynomials is more difficult than the Poisson and Normal cases.

**Theorem 5.** Let \( u \) be an orthogonal basis on \( p = \mu / |\mu| \) with (9) holding. Suppose the basis satisfies the hypergroup property (8).
Then
\[ P(x; \alpha, \theta, p)P(y; \alpha, \theta, p) \times \left\{ 1 + \sum_{|n| \geq 1} \rho_n h_n(\alpha, \theta, u)M_n(x; \alpha, \theta, p, u)M_n(y; \alpha, \theta, p, u) \right\}, \tag{60} \]

with \( \{\rho_n\} \) not depending on \( \alpha \), is non-negative and thus a proper bivariate negative binomial distribution for all \( \alpha > 0 \) if and only if
\[ \rho_n = \mathbb{E}_\xi \left[ \prod_{j=0}^{d-1} \rho_j(\xi)^{n_j} \right], \tag{61} \]

where \( \rho_j(\xi) = \sum_{i=1}^{d-1} u_i^{(j)} \xi_i \) and \( \xi \) is a random vector in the unit simplex satisfying a set of conditions. Recall that
\[ p_{ij}(\omega) = p_i p_j \sum_{l=1}^{d} \omega_{l}s(i, j, l). \]

where \( \omega = \xi/|\xi| \). For non-random \( \xi \) the conditions are
\[ p_{ij}(\omega) \geq \frac{\theta(1 - |\xi|)}{1 + \theta(1 - |\xi|)} p_i p_j, \ i, j = 1, \ldots, d. \tag{62} \]

Sufficient conditions when \( \xi \) is random are also (62) however we do not have a proof of the necessity of (62) then.

**Proof.** The proof of the necessity of (61) follows that in the Poisson case in Theorem 3. The potential pgf of \( Y \) given \( X = x \) in (60) is
\[
\mathbb{E}\left[ \prod_{j=1}^{d} t_j^{Y_j} \mid X = x \right]
\]

\[
= \mathbb{E}_\xi \sum_{|n| \geq 0} \rho_{|n|} h_n(\alpha, \theta, u) M_n(x; \alpha, \theta, p, u) M_n^*(t; \alpha, \theta, p, u)
\]

\[
= (1 - \theta(T_0 - 1))^{-\alpha} \mathbb{E}_\xi \left[ \left( 1 - \frac{\theta|\xi|}{1 - \theta(T_0 - 1)} \right)^{-\alpha} \right]
\]

\[
\times \prod_{i=1}^{d} \left( 1 + \frac{-\theta \xi_j (1 - T_0) + \sum_{j=1}^{d-1} \rho_j \xi_j a_j^{-1} T_j}{1 - \theta(T_0 - 1)} \right)^{x_i}
\]

\[
= \mathbb{E}_\xi \left[ (1 - \theta(1 - |\xi|)(T_0 - 1))^{-\alpha} \right]
\]

\[
\times \prod_{i=1}^{d} \left( 1 - \theta - |\xi|((T_0 - 1) + |\xi| \sum_{j=1}^{d} (p_{ji}(\omega) - p_j)t_j \right)^{x_i}.
\] (63)

If \( \xi \) is non-random a necessary and sufficient condition that (63) be a pgf for all \( x \in \mathbb{Z}_d^+ \) and \( \alpha > 0 \) is that for each \( i = 1, \ldots, d \)

\[
\frac{1 - \theta - |\xi|(1 + \theta))(T_0 - 1) + |\xi| \sum_{j=1}^{d} (p_{ji}(\omega) - p_j)t_j}{1 - \theta - |\xi|(1 + \theta))(T_0 - 1)}
\]

\[
= 1 + \frac{|\xi|}{1 + \theta(1 - |\xi|)} \sum_{j=1}^{d} p_{ji}(\omega)t_j - 1 \]

\]

are pgfs, where

\[
\beta = \frac{|\xi|}{1 + \theta(1 - |\xi|)}, \quad \kappa = \frac{\theta(1 - |\xi|)}{1 + \theta(1 - |\xi|)}.
\]

The constant term in (64) is equal to

\[
\frac{(1 - |\xi|)(1 + \theta)}{1 + \theta(1 - |\xi|)} = \frac{1 + \theta}{\theta} \kappa \geq 0.
\]

The coefficient of \( \prod_{j=1}^{d} t_j^{n_j} \) for \( |n| > 0 \) in (64) is \( \beta \) times

\[
\left( \kappa(1 - |\xi|) \right)^{|n| - 1} \prod_{j=1}^{d} p_{ji}(\omega)^{|n| - (n_j - 1)} \prod_{l=1}^{d} p_{li}^{n_l - \delta_{ij}} - \left( \kappa - 1 \right)^{|n|} \prod_{l=1}^{d} p_{li}^{n_l}
\]

\[
= \kappa^{|n|} \left( \frac{|n|}{n} \right) \prod_{l=1}^{d} p_{li}^{n_l} \left\{ \kappa^{-1} \sum_{j=1}^{d} n_j p_{ji}(\omega)/p_j - 1 \right\}.
\] (65)
The term in the brackets in (65) is non-negative for all \( n \in \mathbb{Z}^d_\text{+} \) if and only if (62) holds. We cannot argue the necessity in the same way if \( \xi \) is random.

**Remark 11.** A sufficient condition for (62) to hold is that for \( i, j = 1, \ldots, d \)

\[
p_{ij}(\omega) \geq \frac{\theta}{1 + \theta} p_i p_j.
\]

(66)

**Remark 12.** The joint pgf of \((X, Y)\) with pgf (60) is

\[
E\xi \left\{ (1 - \theta(S_0 - 1))(1 - \theta(T_0 - 1)) - |\xi|\theta(1 + \theta)(1 - S_0)(1 - T_0) - \theta|\xi|\sum_{j=1}^{d} (p_{ij}(\omega) - p_i p_j)s_it_j \right\}^{-\alpha}
\]

(67)

Three particular cases are:

1. \(|\xi| = 0\); then \(X\) and \(Y\) are independent.

2. \(|\xi| = 1\); then (67) is equal to \(E_\omega(1 + \theta - \theta \sum_{i,j=1}^{d} p_{ij}(\omega)s_it_j)^{-\alpha}\) which corresponds to a bivariate multinomial distribution where the number of trials \(N\) has a 1-dimensional negative binomial distribution. The conditions (66) are then always satisfied assuming that \(u\) has the hypergroup property.

3. \(p_{ij}(\omega) = p_i p_j, i, j = 1, \ldots, d\); then \(X\) and \(Y\) are conditionally independent given \((|X|, |Y|)\).

**Corollary 3.** \((X, Y)\) has a distribution with a Lancaster expansion (60) if and only if, in the extreme points, the conditional pgf of \(Y|X = x\) has the form

\[
\frac{(1 - \kappa)^\alpha}{(1 - \kappa \sum_{j=1}^{d} p_j t_j)^\alpha} \prod_{i=1}^{d} \left( 1 - \beta + \frac{\beta(1 - \kappa)\sum_{j=1}^{d} q_{ji}t_j}{1 - \kappa \sum_{j=1}^{d} p_j t_j} \right)^{x_i},
\]

(68)

where \(0 \leq \beta \leq 1, \kappa = (1 - \beta)\theta/(1 + \theta)\) and \(Q = (q_{ji})\) is a transition matrix, reversible with respect to \(p\), with eigenvalues \(u\).

**Proof.** In Theorem 5, the conditional pgf (63) and simplification (64), make the 1-1 mapping

\[
\beta = \frac{|\xi|}{1 + \theta(1 - |\xi|)}, \quad \kappa = \frac{\theta(1 - |\xi|)}{1 + \theta(1 - |\xi|)} = \frac{(1 - \beta)\theta}{1 + \theta}, \quad q_{ji} = \frac{p_{ji}(\omega) - \kappa p_j}{1 - \kappa}.
\]

(69)
with $p$ fixed. $Q$ is a transition matrix since it has rows sums unity and is non-negative because of (66). In the inverse map to (69) any $0 \leq \beta \leq 1$, $0 < \kappa < 1 - \beta$ and $Q$ determine

$$\theta = \frac{\kappa}{1 - \beta - \kappa}, \quad |\xi| = \frac{\beta}{1 - \kappa}, \quad p_{ji}(\omega) = (1 - \kappa)q_{ji} + \kappa p_j$$

(70) satisfying (62). The calculation showing (63) is equivalent to (68) follows. Divide by $1 + \theta (1 - |\xi|)$ within the brackets raised to the powers $x_i$ and express $p_{ji}(\omega)$ as in (70) to show that the pgf is

$$\frac{1 - \theta (1 - |\xi|)(T_0 - 1)}{1 + \theta (1 - |\xi|)} (1 - \frac{\theta (1 - |\xi|)}{1 + \theta (1 - |\xi|)} T_0)^{-|\mathbf{x}|}$$

$$\times \prod_{i=1}^{d} \left(1 - \frac{|\xi|}{1 + \theta (1 - |\xi|)} - \frac{\theta (1 - |\xi|)}{1 + \theta (1 - |\xi|)} T_0 \right)$$

$$+ \frac{|\xi|(1 - \kappa)}{1 + \theta (1 - |\xi|)} \sum_{j=1}^{d} q_{ji} t_j + \frac{\kappa |\xi|}{1 + \theta (1 - |\xi|)} T_0 \right)^{x_i}$$

$$= (1 - \kappa)^\alpha (1 - \kappa T_0)^{-\alpha} (1 - \kappa T_0)^{-|\mathbf{x}|}$$

$$\times \left((1 - \beta)(1 - \kappa T_0) + \beta (1 - \kappa) \sum_{j=1}^{d} q_{ji} t_j \right)^{x_i}$$

(71)

(68) now follows from (71) by dividing inside the bracketed terms by $1 - \kappa T_0$.

**Remark 13.** An interpretation of the extreme points in the conditional distribution of $Y | X = x$ with a pgf (68) is as the transition functions in a discrete time Markov chain. Consider a population of individuals of $d$ types with configuration $x$. In a transition each individual dies with probability $1 - \beta$, or gives birth with probability $\beta$ according to the multivariate negative binomial pgf $(1 - \kappa)(1 - \kappa \sum_{i=1}^{d} p_i t_i)^{-1}$. Parents then change their types according to the transition matrix $Q$. Immigration occurs according to the pgf $(1 - \kappa)^\alpha (1 - \kappa \sum_{i=1}^{d} p_i t_i)^{-\alpha}$. The resulting population configuration is then $Y$. The stationary distribution of the Markov chain is multivariate Meixner with parameters $\alpha$, $\theta$, $p$ specified by (70). If $(\beta, Q)$ are random it may be possible that $Q$ has negative entries but the expected value of (68) is still a pgf.
Remark 14. A continuous time reversible Markov process \( \{X(t)\}_{t \geq 0} \) with a stationary multivariate Meixner distribution and multivariate Meixner polynomial eigenfunctions can be constructed similarly to the multivariate Poisson in Remark 9. Choose \( \xi \) so that (43) - (46) hold. The conditions that (62) hold is then, to order \( \Delta t \),

\[
|\xi|p_{ji} \geq \frac{\theta}{1 + \theta}|\xi|p_j; \text{ or }
\delta_{ij}(1 - (\nu + |\gamma|)\Delta t) + |\gamma|\Delta t\bar{p}_{ji} \geq \nu \theta \Delta tp_j.
\]

That is

\[
|\gamma|\bar{p}_{ji} \geq \nu \theta p_j, \ j \neq i \tag{72}
\]

and the conditions are satisfied if \( i = j \) as \( \Delta t \to 0 \). Consider the factor (64) arising in the conditional pgf written as

\[
1 - \frac{|\xi|}{(1 + \theta(1 - |\xi|))(1 - \kappa T_0)} + \frac{|\xi|}{(1 + \theta(1 - |\xi|))(1 - \kappa T_0)} \sum_{j=1}^{d} p_{ji}(\omega)t_j. \tag{73}
\]

The first term in (73) is, to order \( \Delta t \),

\[
1 - \frac{1 - \nu \Delta t}{(1 + \theta \nu \Delta t)(1 - \theta \nu \Delta t T_0)} \approx \nu \Delta t + \theta \nu \Delta t(1 - T_0)
\]

and the second term is

\[
(1 - \theta \nu \Delta t + \theta \nu \Delta t T_0) \sum_{j=1}^{d} \left( \delta_{ij}(1 - (\nu + |\gamma|)\Delta t) + |\gamma|\Delta t\bar{p}_{ji} \right)t_j.
\]

It is sufficient to replace \( |\xi| \) and \( |\xi|p_{ji}(\omega) \) by their means in calculations because only linear terms occur to order \( \Delta t \). The infinitesimal probability generating of \( X(t + \Delta t) | X(t) = x \) function follows as

\[
(1 - \theta \nu \alpha + \theta \nu \alpha \sum_{j=1}^{d} p_j t_j)
\]

\[
\times \prod_{i=1}^{d} \left( \nu(1 + \theta)\Delta t + \theta \nu t_i \Delta t \sum_{j=1}^{d} p_j t_j + \Delta t \sum_{j=1}^{d} \bar{r}_{ji} t_j + (1 - \nu(1 + \theta)\Delta t - |\gamma|\Delta t) t_i \right)^{x_i} \tag{74}
\]
where $\tilde{r}_{ji} = \gamma \tilde{p}_{ji} - \theta \nu p_j \geq 0$. As in the Poisson model note all that is required are terms of order $\Delta t$, but more are retained for a better understanding. $\{X(t)\}_{t \geq 0}$ is now identified as a multitype birth and death process in a population of individuals of $d$ types; immigrants arrive at rates $\theta \nu \alpha p_j$; births occur to individuals at rates $\nu \theta p_j$, not depending on parental type; deaths occur to individuals at rate $\nu (1 + \theta)$; and changes of type from $i$ to $j$ occur at rate $\tilde{r}_{ji}$. The stationary distribution is multivariate Meixner. $|X(t)|$ is a conventional birth and death process with birth rates when $|X(t)| = |x|$ of $\theta \nu (\alpha + |x|)$ and death rates $\nu (1 + \theta) |x|$ with a negative binomial stationary distribution.

5. Multivariate Hermite-Chebycheff polynomials

Let $X$ be a $N(0, \tau)$ random variable. The Hermite-Chebycheff polynomials, $\{H_n(X; \tau)\}_{n=0}^{\infty}$, orthogonal on $X$ which is $N(0, \tau)$ have a generating function

$$\sum_{n=0}^{\infty} H_n(X; \tau) \frac{z^n}{n!} = \exp \left\{ xz - \frac{1}{2} \tau z^2 \right\}. \quad (75)$$

The orthogonality relationship is

$$\mathbb{E} [H_m(X; \tau)H_n(X; \tau)] = \delta_{mn} n! \tau^n.$$

Let $X = (X_1, \ldots, X_d)$ be independent normal random variables with variances $\tau = (\tau_1, \ldots, \tau_d)$ and density

$$f(x; \tau) = \prod_{i=1}^{d} \frac{1}{\sqrt{2\pi \tau_i}} e^{-\frac{x_i^2}{2\tau_i}}, \quad x \in \mathbb{R}^d. \quad (76)$$

A set of multivariate Hermite-Chebycheff polynomials on $X$ is the product set $\bigotimes_{i=1}^{d} \{H_n(X_i; \tau_i)\}_{n=0}^{\infty}$. Denote $p = \tau/|\tau|$ and let $u$ be an orthogonal basis on $p$, satisfying (6). Another set is obtained by considering the mapping

$$X \rightarrow \vec{X} = \left( \sum_{i=1}^{d} u_i^{(j)} X_i \right)_{j=0}^{d-1} \quad (77)$$

which preserves normality and independence because

$$\text{Cov}(\vec{X}_j, \vec{X}_k) = \sum_{i=1}^{d} u_i^{(j)} u_i^{(k)} \tau_i = |\tau| \delta_{jk} a_k, \quad j, k = 0, 1, \ldots, d - 1.$$
Theorem 6. The set of multivariate Hermite-Chebycheff polynomials associated with the mapping (77) is

\[ H_n(X; \tau, u) = d^{-1} \prod_{j=0}^{d-1} H_{\hat{X}_j}(\tau, |\tau| a_j), \]  

where \( n_1 = (n_1, \ldots, n_{d-1}) \).

The orthogonality relationship is

\[ E[H_m(X; \tau, u)H_n(X; \tau, u)] = \delta_{mn} |\tau|^{|n|} \prod_{j=0}^{d-1} a_{nj}^{n_j} := \delta_{mn} h_n(|\tau|, u)^{-1}. \]  

A generating function for the polynomials (78) is

\[ G_{HC}(X, w, u) = \sum_{n \geq 0} H_n(X; \tau, u) \prod_{j=0}^{d-1} w_j^{n_j} \prod_{j=0}^{d-1} a_{nj}^{n_j}! = \exp \left\{ \sum_{j=0}^{d-1} \hat{X}_j w_j - \frac{1}{2} |\tau| \sum_{j=0}^{d-1} a_j w_j^2 \right\}. \]  

Remark 15. The multivariate Poisson-Charlier polynomials tend to the multivariate Hermite-Chebycheff polynomials as the elements of \( \mu \) tend to infinity. This is to be expected as \( |\mu|^{1/2}(X - \mu) \) converges to a normal random vector \( Z \) with independent entries and variances \( p \).

\[ n_0! \cdots n_{d-1}! |\mu|^{n/2} (-1)^{n} C_n(|\mu|^{1/2} Z + \mu; \mu, u) \to H_n(Z; p, u). \]  

The limit (81) is seen from a generating function argument.
normal distribution as $\alpha \to \infty$ with mean 0 and covariance matrix $\theta^2 p p^T + \theta \text{ diag}(p)$. The variables in the limit are not independent, however

$$\hat{Z}_j = \sum_{i=1}^d Z_i u_i^{(j)}, \quad j = 0, \ldots, d-1$$

are independent with $\text{Var}(\hat{Z}_0) = \theta(1 + \theta)$ and $\text{Var}(\hat{Z}_j) = \theta a_j, \quad j \geq 1$.

$$n_1! \cdots n_d! \theta^{n_0} M_n(\alpha^{1/2} Z + \alpha \theta p; \alpha, p, u) \to H'_n(Z; p, u), \quad (82)$$

where $H'_n$ is similarly defined to $H_n$ in (78), but with $a_0 = \theta(1 + \theta)$ instead of $a_0 = 1$. The generating function convergence showing (82), with a similar calculation to the Poisson case, is

$$\log G_M(\alpha^{1/2} + \alpha \theta p, \alpha^{-1/2} w, \alpha, \theta, u)$$

$$\to -\frac{1}{\hat{Z}_0} w_0 + \sum_{j=1}^{d-1} \hat{Z}_j w_j - \frac{1}{2} \left(1 + \theta\right) \theta^{-1} w_0^2 - \frac{1}{2} \theta \sum_{j=1}^{d-1} w_j^2.$$  

**Remark 16.**

$$G_{HC}(x, e_0, u) = \exp \left\{ |X| w_0 - \frac{1}{2} |\tau| w_0^2 \right\}. \quad (83)$$

generates Hermite-Chebycheff polynomials in the sum $|X|$.

**Corollary 4.** The transform

$$H_n^*(s; \tau, u) := E \left[ \prod_{i=1}^d \exp \left\{ \phi_i X_i \right\} H_n(X; \tau, u) \right]$$

$$= \exp \left\{ \sum_{i=1}^d \tau_i \phi_i^2 \right\} \prod_{j=0}^{d-1} \left( \sum_{i=1}^d \tau_i u_i^{(j)} \phi_i \right)^{n_j}. \quad (84)$$

**Proof.**

$$E \left[ \exp \left\{ \sum_{j=0}^{d-1} \phi_j X_j \right\} G_{HC}(x, w, u) \right]$$

$$= \exp \left\{ \sum_{i=1}^d \tau_i \phi_i^2 \right\} \exp \left\{ \sum_{i=1}^d \sum_{j=0}^{d-1} \tau_i u_i^{(j)} \phi_i w_j \right\}$$

therefore (84) holds.
5.1. Normal Table: Canonical correlations

**Theorem 7.** Let \((X_{ij}), i = 1, \ldots, r, j = 1, \ldots, c, r \leq c\), be an \(r \times c\) array of independent normal random variables with variances \(\tau = (\tau_{ij})\). Let \(p_{ij} = \tau_{ij}/|\tau|\). Suppose there is a Lancaster expansion \(^{[29]}\). Denote the marginals of the table by \(X = (X_i)_{i=1}^r, Y = (Y_j)_{j=1}^c\). Then

\[
P(X = x, Y = y) = f(x; \tau)f(y; \tau)
\times \left\{ 1 + \sum_{n:|n| \geq 1} \theta_1^{n_1} \cdots \theta_r^{n_r-1} h_n(|\tau|, u) H_n(x; \mu_x, u) H_n(y; \mu_y, v) \right\},
\]

where \(n = (n_0, n_1, \ldots, n_{r-1}), n^* = (n, 0_{c-r}).\)

**Proof.**

\[
\text{Cov}\left(\sum_{i=1}^r X_i u_i^{(l)}, \sum_{j=1}^c X_j v_j^{(m)}\right) = \sum_{i=1}^r \sum_{j=1}^c \tau_{ij} u_i^{(l)} v_j^{(m)} = \delta_{lm} |\tau| \rho_l.
\]

Therefore

\[
\mathbb{E}[G_{HC}(X, \omega, \theta)G_{HC}(Y, \zeta, \theta)]
= \exp \left\{ \frac{1}{2} |\tau| \sum_{l=0}^{r-1} w_i^2 + \frac{1}{2} |\tau| \sum_{m=0}^{c-1} z_m^2 + |\tau| \sum_{l,m=0}^{r-1} \delta_{lm} \rho_l w_i z_m \right\}
\times \exp \left\{ -\frac{1}{2} |\tau| \sum_{l=0}^{r-1} w_i^2 - \frac{1}{2} |\tau| \sum_{m=0}^{c-1} z_m^2 \right\}
= \exp \left\{ |\tau| \sum_{l=0}^{r-1} \rho_l w_i z_l \right\}.
\]

**Remark 17.** In the classical theory of canonical correlations in normal random vectors \(X \oplus Y\) is a direct sum of \(r + c, r \leq c\) normal random vectors with a partitioned covariance matrix

\[
V = \begin{bmatrix}
V_{11} & V_{12} \\
V_{21} & V_{22}
\end{bmatrix}.
\]

The objective is to find linear transformations \(X \rightarrow \hat{X}, Y \rightarrow \hat{Y}\) such that the random variables are independent within sets and \(\{(\hat{X}_k, \hat{Y}_k)\}_{k=0}^{r-1}\)
are independent pairs with correlations \( \{ \rho_k \}_{k=0}^{r-1} \). The indexing is chosen to agree with that in Theorem 7. Without loss of generality \( X \) and \( Y \) are independent within vectors. Theorem 7 is a special case of canonical correlations. This occurs because in the table \((X_{ij})\), \( \hat{X}_k = \sum_{i=1}^r u_i^{(k)} X_i \), \( \hat{Y}_k = \sum_{i=1}^c v_i^{(k)} Y_i \) and \( \hat{X}_0 = \hat{Y}_0 \).

5.2. Lancaster expansion

**Theorem 8.** Let \( \{ u^{(j)} \}_{j=0}^{d-1} \) be an orthogonal basis on \( p = \tau/|\tau| \) with \( (6) \) holding. Then

\[
 f(x, \tau)f(y, \tau) \left\{ 1 + \sum_{|n| \geq 1} \rho_n h_n(|\tau|, u) H_n(x; \tau, u) H_n(y; \tau, u) \right\} \tag{88}
\]

is non-negative and thus a proper bivariate normal distribution if and only if

\[
 \rho_n = \mathbb{E}\left[ \prod_{j=0}^{d-1} \xi_{n,j}^{n_j} \right], \tag{89}
\]

where \( \xi \) is a random vector with elements in \([-1, 1]\). An equivalent statement to the characterization \( (89) \) is that, conditional on \( \xi \), \( (X, Y) \) is a bivariate normal pair of random vectors, independent within vectors, with means \((0, 0)\), variances \((\tau, \tau)\) and cross covariances \( V(\xi) = (v_{ij}(\xi)) \), where

\[
 v_{ij}(\xi) = \text{Cov}(X_i, Y_j \mid \xi) = |\tau| p_i p_j \sum_{l=0}^{d-1} \xi_{l} a_{i}^{-(l)} u_{i}^{(l)} u_{j}^{(l)}. \tag{90}
\]

**Proof. Necessity.** Suppose \( (88) \) holds. Let \( (X, Y) \) be a pair of normal random vectors with distribution \( (89) \). Then

\[
 \rho_n H_n(x; \tau, u) = \mathbb{E}\left[ H_n(Y; \tau, u) \mid X = x \right]
\]

and there is only one leading term of degree \(|n|\), proportional to

\[
 m(n, \bar{x}, u) = \prod_{j=0}^{d-1} \bar{x}_{n,j}^{n_j}
\]

on the left side. Rearranging

\[
 \mathbb{E}\left[ m(n, \hat{Y}, u) \mid X = x \right] = \rho_n m(n, \hat{x}, u) + R_{|n|-1}(\hat{x}) \tag{91}
\]
where $R_{|n|-1}(\hat{x})$ is a polynomial of degree $|n| - 1$ in $\hat{x}$. Divide \([91]\) by $m(n, \hat{x}, u)$ and let $\hat{x}_j \to \infty$, $j = 0 \ldots, d - 1$. Let $\xi$ be a random variable with the limit distribution of $(\hat{Y}_j / \hat{x}_j)$ given $\hat{X} = \hat{x}$. Then \((89)\) holds.

**Sufficiency** Suppose that \((89)\) holds. Let $\rho_r = \sum_{i=1}^d u^{(r)}_i \xi_i$. The transform of a potential exchangeable bivariate distribution $(X, Y)$ formed from \((88)\) is

$$
E \left[ \prod_{i,j=1}^d \exp \left\{ \phi_i X_i + \psi_i Y_j \right\} \right]
= \sum_{n: |n| \geq 0} \rho_n h_n(|\tau|, u) H_n^*(\tau; \tau, u) H_n^*(t; \tau, u)
= \exp \left\{ \sum_{i=1}^d \tau_i (\phi_i^2 + \psi_i^2) \right\}
\times E \left[ \exp \left\{ \sum_{r=0}^{d-1} |\tau| a_r^{-1} \rho_r \left( \sum_{i=1}^d p_i \phi_i u^{(r)}_i \right) \left( \sum_{j=1}^d p_j \psi_j u^{(r)}_j \right) \right\} \right]
= \exp \left\{ \sum_{i=1}^d \tau_i (\phi_i^2 + \psi_i^2) \right\}
\times E \left[ \exp \left\{ |\tau| \sum_{i,j=1}^d p_i p_j \sum_{r=0}^{d-1} \xi_r a_r^{-1} u^{(r)}_i u^{(r)}_j \phi_i \psi_j \right\} \right]
= \exp \left\{ \sum_{i=1}^d \tau_i (\phi_i^2 + \psi_i^2) \right\}
\times E \left[ \exp \left\{ \tau \sum_{i,j=1}^d p_i p_j \sum_{r=0}^{d-1} \xi_r a_r^{-1} u^{(r)}_i u^{(r)}_j \phi_i \psi_j \right\} \right]
= \delta_{rs} |\tau| \xi_r a_r.
$$

which is a proper moment generating function with the covariance structure \((90)\). To see this note that the conditional distribution of $(\hat{X}, \hat{Y})$ consists of independent bivariate normal pairs $\{(\hat{X}_i, \hat{Y}_i)\}_{i=1}^d$. The covariance structure is

$$
\text{Cov}(\hat{X}_r, \hat{Y}_s \mid \xi) = \sum_{i,j=1}^d u^{(r)}_i u^{(s)}_j \text{Cov}(X_i, Y_j \mid \xi)
= \sum_{i,j=1}^d u^{(r)}_i u^{(s)}_j |\tau| p_i p_j \sum_{l=0}^{d-1} \xi_l a_l^{-1} u^{(l)}_i u^{(l)}_j
= \delta_{rs} |\tau| \xi_r a_r.
$$

That is

$$
\text{Var}(\hat{X}_r \mid \xi) = \text{Var}(\hat{Y}_r \mid \xi) = |\tau| a_r, \quad \text{Corr}(\hat{X}_r, \hat{Y}_r \mid \xi) = \xi_r.
$$

\[ \square \]
Remark 18. $X$ and $Y$ are independent if $\rho_n = 0$ for all $|n| \geq 1$. $X = Y$ if $\rho_n = 1$ for $|n| \geq 1$, when $\xi_0 = \cdots = \xi_{d-1} = 1$. These two cases are true from general theory, and also follow from evaluating (92).

Remark 19. Theorem 8 includes the classical case of canonical correlation of a pair of exchangeable normal vectors $(X, Y)$ when $\xi$ is constant. There is an insistence that the cross correlation matrix have an expansion (90). This is like a Lancaster expansion, but non-negativity is not required.

Theorem 8 in general is an extension of a canonical correlation expansion for the distribution of $(X, Y)$ with the particular transformation to $(\hat{X}, \hat{Y})$. The characterization is similar to a canonical correlation characterization, with a random covariance structure such that the marginal vectors are normal. The joint moment generating function of $(X, Y)$ is

$$\mathbb{E}_\xi \left[ \exp \left\{ \frac{1}{2} s^T s + \frac{1}{2} t^T t + s^T V(\xi)t \right\} \right]$$

(93)

where $V(\xi)$ is defined in (90). The conditional moment generating function of $Y \mid X$ is

$$\mathbb{E}_\xi \left[ \exp \left\{ \frac{1}{2} t^T (I - V(\xi)^2) t + t^T V(\xi) X \right\} \right].$$

(94)

Remark 20. The conditional distribution of $Y \mid X = x$ has an interpretation as a transition function in a discrete time Markov chain, where transitions are made from $x$ to $Y$ which is normal with mean $V(\xi)x$ and covariance matrix $I - V(\xi)^2$, with the cross covariance matrices $V(\xi)$ identically distributed at each epoch.

References

Aitken, A. C. and Gonin, H. T. (1935) On fourfold sampling with and without replacement. Proc. Roy. Soc. Edinb. 55 114–125.

Bakry, D. and Huet, N. (2006) The hypergroup property and representation of Markov Kernels. Séminaire de Probabilités XLI, Lecture notes in Mathematics, Vol 1934, 295–347, Springer.

Bochner, S. (1954). Positive zonal functions on spheres, Proc. Nat. Acad. Sci. USA 40 1141–1147.

Diaconis, P. and Griffiths R. C. (2012) Exchangeable pairs of Bernoulli random variables, Krawtchouk polynomials, and Ehrenfest urns. Aust. NZ J. Stat. 54 81–101.
Diaconis, P. and Griffiths R. C. (2014) An introduction to multivariate Krawtchouk polynomials and their applications. *Journal of Statistical Planning and Inference* **154** 39–53.

Eagleson, G. K. (1964) Polynomial expansions of bivariate distributions. *Ann. Math. Statist.* **35** 1208–1215.

Eagleson, G. K. (1969) A characterization theorem for positive definite sequences on the Krawtchouk polynomials. *Austral. J. Statist.* **11** 29–38.

Genest, V. X., Vinet, L. and Zhedanov, A. (2013). The multivariate Krawtchouk polynomials as matrix elements of the rotation group representations on oscillator states. *J. Phys A: Math. Theor.* **46** 505203.

Genest, V. X., Miki, H., Vinet, L., and Zhedanov, A. (2014). The multivariate Meixner polynomials as matrix elements of SO (d, 1) representations on oscillator states. *J. Phys A: Math. Theor.* **47** 045207.

Genest, V. X., Miki, H., Vinet, L., and Zhedanov, A. (2014a). The multivariate Charlier polynomials as matrix elements of the Euclidean group representation on oscillator states. *J. Phys A: Math. Theor.* **47** 215204.

Griffiths R. C. (1969) The canonical correlation coefficients of bivariate gamma distributions. *Ann. Math. Statist.* **40** 1401–1408.

Griffiths, R. C. (1971) Orthogonal polynomials on the multinomial distribution. *Austral. J. Statist.* **13** 27–35. Corrigenda (1972) *Austral. J. Statist.* **14** 270.

Griffiths, R. (1975). Orthogonal polynomials on the negative multinomial distribution. *J. Multivariate Anal.* **5** 271–277.

Griffiths, B. [R.C.] (2009). Stochastic Processes with Orthogonal Polynomial Eigenfunctions, *J. Comput. Appl. Math.* **23** 739–744.

Grunbaum, F. and Rahman, M. (2011) A system of multivariable Krawtchouk polynomials and a probabilistic application. *SIGMA* **7** 119–136.

Hoare, M. and Rahman, M. (1983) Cumulative Bernoulli trials and Krawtchouk processes. *Stochastic Process. Appl.* **16** 113–139.

Hoare, M., Rahman, M. (2008). A probabilistic origin for a new class of bivariate polynomials. *SIGMA* **4** 18 pages.
Iliev, P. and Xu, Y. (2007). Discrete orthogonal polynomials and difference equations of several variables. *Advances in Mathematics* 212 1–36.

Iliev, P. (2012) A Lie-theoretic interpretation of multivariate hypergeometric polynomials. *Compositio Mathematica* 148 991–1002.

Iliev, P. (2012a). Meixner polynomials in several variables satisfying bispectral difference equations. *Advances in Applied Mathematics* 49 15–23.

Ishmail, M. E. H. (2005) *Classical and Quantum Orthogonal Polynomials in one variable*, Volume 98 of *Encyclopedia of Mathematics and its Applications*. Cambridge: Cambridge University Press.

Koudou A. E. (1996) Probabilités de Lancaster, *Exposition. Math.* 1 247–275.

Koudou A. E. (1998) Lancaster bivariate probability distributions with Poisson, Negative Binominal and Gamma margins, *Sociedad de Estadística e Investigación Operativa Test*. 7 95–110.

Lancaster, H. (1965) The Helmert matrices. *Amer Math. Monthly* 72, 4–12.

Lancaster H. (1969) *The chi-squared distribution*, John Wiley & Sons.

Mizukawa, H. (2010) Finite Gelfand pair approaches for Ehrenfest diffusion model. arXiv:1009.1205.

Mizukawa, H. (2011) Orthogonality relations for multivariate Krawtchouk polynomials. arXiv:1009.1203.

Mizukawa, H. and Tanaka, H. (2004). $(n + 1, m + 1)-$ hypergeometric functions associated to character algebras. *Proceedings of the American Mathematical Society* 132, 2613–2618.

Sarmanov, I. O. (1968) A generalized symmetric gamma-correlation, *Dokl. Akad. Nauk SSSR* 179 (1968) 1279–1281.

Sarmanov I. O. and Bratoeva, Z. N. (1967) Probabilistic properties of bilinear expansions of Hermite polynomials. *Theor. Probability Appl.* 12 (1967) 470–481.

Xu, Y. (2013) Hahn, Jacobi, and Krawtchouk polynomials of several variables. *Journal of Approximation Theory*. In press.
Zhou, H. and Lange, K. (2009) Composition Markov chains of multinomial type. Adv. Appl. Probab. 41 270–291.