APPLICATIONES MATHEMATICAE
32,1 (2005), pp. 57–67

D. POMMERET (Bruz)

APPROXIMATE POLYNOMIAL EXPANSION FOR JOINT DENSITY

Abstract. Let (X,Y) be a random vector with joint probability measure σ and with margins μ and ν. Let $(P_n)_{n\in\mathbb{N}}$ and $(Q_n)_{n\in\mathbb{N}}$ be two bases of complete orthonormal polynomials with respect to μ and ν, respectively. Under integrability conditions we have the following polynomial expansion:

$$\sigma(dx,dy)=\sum_{n,k\in\mathbb{N}}\varrho_{n,k}P_n(x)Q_k(y)\mu(dx)\nu(dy).$$

In this paper we consider the problem of changing the margin μ into $\widetilde{\mu}$ in this expansion. That is the case when μ is the true (or estimated) margin and $\widetilde{\mu}$ is its approximation. It is shown that a new joint probability with new margins is obtained. The first margin is $\widetilde{\mu}$ and the second one is expressed using connections between orthonormal polynomials. These transformations are compared with those obtained by the Sklar Theorem via copulas. A bound for the distance between the new joint distribution and its parent is proposed. Different cases are illustrated.

1. Introduction and motivations. If X and Y are two random variables with probability measures μ and ν and if $(P_n)_{n\in\mathbb{N}}$ (resp. $(Q_n)_{n\in\mathbb{N}}$) is a μ (resp. ν) complete orthonormal basis of polynomials, then the joint density of (X,Y), say σ, satisfies

$$\sigma(dx,dy)=\sum_{n,k\in\mathbb{N}}\varrho_{n,k}P_n(x)Q_k(y)\mu(dx)\nu(dy), \tag{1}$$

as soon as the series $\sum\varrho_{n,k}P_n(x)Q_k(y)$ converges in $L^2(\mu\times\nu)$, or equivalently if $\sum_{n,k\in\mathbb{N}}\varrho_{n,k}^2<\infty$. The sequence $(\varrho_{n,k})$ expresses the correlation between $P_n(X)$ and $Q_k(Y)$; that is, $\varrho_{n,k}=\mathbb{E}(P_n(X)Q_k(Y))$.

2000 *Mathematics Subject Classification*: Primary 62E10; Secondary 62E17.

Key words and phrases: bivariate density function, connection coefficient, copula, orthogonal polynomial.

Such a representation has many advantages. In particular, estimators of the coefficients $\varrho_{n,k}$ are easily obtained by empirical polynomial correlations between X and Y (see [1]). Also, independence between X and Y may be measured by the use of these coefficients (see [7]). An important case is obtained when the coefficients $\varrho_{n,k}$ vanish for $n \neq k$. In this case σ is called a *Lancaster distribution.* Such bivariate distributions have been studied in [8]. Also, [15] proved that if X and Y are unbounded then the condition $\varrho_{n,k} = 0$ for $n - k > N$ for some integer N implies that $\varrho_{n,k} = 0$ for $n \neq k$; that is, σ is a Lancaster distribution. In canonical analysis, such a situation is studied as singular value decomposition (see [2]).

In this paper we wish to investigate the structure of the bivariate probability measures satisfying (1) by changing their margins. Our purpose is to keep a polynomial expansion of the joint density function. More precisely, given any joint distribution σ satisfying a polynomial expansion (1), we study the consequences of transformations of one margin μ or ν. It is shown that replacing μ by a probability measure $\widetilde{\mu}$ in the series (1) yields a new joint distribution, say $\widetilde{\sigma}$, with a polynomial expansion with respect to its new margins $\widetilde{\mu}$ and $\widetilde{\nu}$. Next our purpose is to determine the second margin $\widetilde{\nu}$ by using connection coefficients between orthogonal polynomials. A bound for the distance between the parent joint distribution and the new one will be evaluated.

This problem of change of margins is motivated by different reasons:

Practical motivation. As an illustration assume that μ is a given measure, for example a negative binomial probability, and let $(P_n)_{n\in\mathbb{N}}$ denote its associated Meixner polynomials. Assume also that ν and the coefficients $\varrho_{n,k}$ are known and that σ has the form (1). As an illustration we may consider the vector (X, Y) corresponding to the number of accident claims and the age of insured person for insurance data. Usually X and Y are supposed to be negative binomial and Gaussian distributed, respectively. It is frequent to approximate the negative binomial distribution by the Poisson one (see for instance [16]). Consider an appropriate Poisson approximation of μ, say $\widetilde{\mu}$ (for example, $\widetilde{\mu}$ and μ have common mean). We write this $\mu \approx \widetilde{\mu}$. Such an approximation yields the following approximation of σ:

$$\sigma(dx, dy) \approx \sum_{n,k\in\mathbb{N}} \varrho_{n,k} P_n(x) Q_k(y) \widetilde{\mu}(dx)\nu(dy) \equiv \widetilde{\sigma}(dx, dy).$$

Then it is of interest to study the new expression $\widetilde{\sigma}$ and to compare it with σ. Our purpose is to show that it defines a new joint probability measure with new margins. A bound for the distance of $\widetilde{\sigma}$ from its parent distribution σ will be obtained.

Theoretical motivations. An important problem, but without general solution, is the following: given two margins μ and ν with associated orthogonal polynomials P_n and Q_n, respectively, do there exist coefficients $\varrho_{n,k}$ such that (1) defines a joint probability measure? This problem is partially studied in [15] (see also [13]). It is clear that if X and Y are bounded random variables then there exist such coefficients $\varrho_{n,k}$ (for example: choose a finite family of coefficients). However, if X and Y are unbounded there is no general condition to ensure the positivity of σ defined in (1).

The theory of copulas introduced in [14] gives another viewpoint on this problem. If we express a joint distribution function as a combination of its margins, then the same combination of any margins yields a new joint distribution function.

Here we wish to investigate the following related problem: if (1) defines a joint probability measure, do the same coefficients $\varrho_{n,k}$ ensure a new joint probability measure if we change one margin? The answer is positive and this approach will be compared with the copula approach.

The paper is organized as follows. In Section 2 changes of margins preserving polynomial expansions are studied. In Section 3 changes of margins using copulas are examined. In Section 4 a bound for the distance between the new joint distribution and its parent is obtained. We give some illustrations related to connection coefficients in Section 5. In the Appendix we give a brief exposition of the connection problem for polynomials.

2. Change of margin

2.1. *General case.* Our purpose is to study the structure of bivariate distributions satisfying (1). We obtain a method for constructing new joint distributions with particular margins. The following result may be proved in the multivariate case in much the same way.

THEOREM 1. *Let σ be a joint density function satisfying* (1) *and let $\widetilde{X}$ be a random variable with distribution measure $\widetilde{\mu}$ such that the series $\sum \varrho_{n,k}^2 \mathbb{E}(P_n^2(\widetilde{X}))$ converges. Then the measure*

$$\widetilde{\sigma}(dx,dy) = \sum_{n,k\in\mathbb{N}} \varrho_{n,k} P_n(x) Q_k(y) \widetilde{\mu}(dx)\nu(dy) \tag{2}$$

is still a probability measure. Its first margin is $\widetilde{\mu}$ and the second one is given by

$$\widetilde{\nu}(dy) = \sum_{k\in\mathbb{N}} \alpha_k Q_k(y)\nu(dy), \tag{3}$$

where $\alpha_k = \sum_{n\in\mathbb{N}} \varrho_{n,k}\mathbb{E}(P_n(\widetilde{X}))$.

Proof. By construction we have $\widetilde{\sigma} \geq 0$. We now prove that its mass is equal to 1. Using the orthogonality of the sequence $\{Q_k(x)\}$ and the convergence condition we obtain

$$\int \widetilde{\sigma}(dx,dy) = \sum_{n,k\in\mathbb{N}} \varrho_{n,k} \int P_n(x)\,\widetilde{\mu}(dx) \int Q_k(y)\,\nu(dy) = \sum_{n\in\mathbb{N}} \varrho_{n,0} \int P_n(x)\,\widetilde{\mu}(dx).$$

But $\varrho_{n,0} = \mathbb{E}(P_n(X)) = 0$ for all $n > 0$. Hence

$$\int \widetilde{\sigma}(dx,dy) = \varrho_{0,0} = 1.$$

The first margin of $\widetilde{\sigma}$ is given by

$$\begin{aligned}\int \widetilde{\sigma}(dy) &= \sum_{n,k\in\mathbb{N}} \varrho_{n,k} P_n(x)\widetilde{\mu}(dx) \int Q_k(y)\,\nu(dy) = \sum_{n\in\mathbb{N}} \varrho_{n,0} P_n(x)\,\widetilde{\mu}(dx)\\ &= P_0(x)\widetilde{\mu}(dx) = \widetilde{\mu}(dx).\end{aligned}$$

The second margin is given by

$$\int \widetilde{\sigma}(dx) = \sum_{n,k\in\mathbb{N}} \varrho_{n,k} Q_k(y)\Big(\int P_n(x)\widetilde{\mu}(dx)\Big)\nu(dy),$$

which establishes the formula. ■

If the $\widetilde{\mu}$-orthogonal polynomials, denoted by $(\widetilde{P}_n)_{n\in\mathbb{N}}$, are known we may simply write

$$\widetilde{\nu}(dy) = \sum_{k\in\mathbb{N}} \alpha_k Q_k(y)\nu(dy), \tag{4}$$

where $\alpha_k = \sum_{n\in\mathbb{N}} \varrho_{n,k}\widetilde{C}_0(n)$, with $\widetilde{C}_k(n)$ denoting the connection coefficients between P_n and $\widetilde{P}_n$; that is, $P_n(x) = \sum_{k=0}^{n} \widetilde{C}_k(n)\widetilde{P}_k(x)$ (see Appendix).

The joint distribution $\widetilde{\sigma}$ can be expressed in the same way. We have

$$\widetilde{\sigma}(dx,dy) = \sum_{n,k\in\mathbb{N}} \widetilde{\varrho}_{n,k}\widetilde{P}_s(x)Q_k(y)\widetilde{\mu}(dx)\nu(dy), \tag{5}$$

where $\widetilde{\varrho}_{n,k} = \varrho_{n,k}\sum_{s\le n} \widetilde{C}_s(n)$. Also

$$\widetilde{\sigma}(dx,dy) = \sum_{n,k\in\mathbb{N}} \varrho_{n,k} \sum_{s\le n} \widetilde{C}_s(n)\widetilde{P}_s(x)Q_k(y)/N(y)\,\widetilde{\mu}(dx)\widetilde{\nu}(dy), \tag{6}$$

where $N(y)$ is such that $\widetilde{\nu}(dy) = N(y)\nu(dy)$ a.e. The associated connection coefficients appear in (5) and (6). However, from (4), we only need the first coefficients $\widetilde{C}_0$ to compute the new margin $\widetilde{\nu}$. This fact will be illustrated in Section 5.

2.2. *Transformations of Lancaster probabilities.* A particularly interesting case is obtained when the orthonormal sequences $\{P_n(x)\}$ and $\{Q_n(x)\}$ satisfy the following bi-orthogonality:

$$\mathbb{E}(P_n(X)Q_k(Y)) = \varrho_n\delta_{nk}, \tag{7}$$

where $\delta_{nk} = 1$ if $n = k$ and 0 otherwise. Then (1) becomes

$$\sigma(dx, dy) = \sum_{n\in\mathbb{N}} \varrho_n P_n(x) Q_n(y) \mu(dx) \nu(dy). \tag{8}$$

In such a situation σ is called a *Lancaster probability* and the sequence ϱ_n is referred to as the *Lancaster sequence*. Lancaster probabilities appear in Markov theory, canonical analysis (see [6], and references therein) and also have a role in disjunctive kreaging (see [9]).

Under the condition that the series (ϱ_n) converges in $l^2(\mathbb{N})$, the polynomial expansion has the simpler form (8) and Theorem 1 may be used. We just give a reformulation here:

Corollary 2. *Let σ be a Lancaster probability satisfying* (8). *Then*, *for all probability measures $\widetilde{\mu}$, the measure*

$$\begin{aligned}\widetilde{\sigma}(dx, dy) &= \sum_{n\in\mathbb{N}} \varrho_n P_n(x) Q_n(y) \widetilde{\mu}(dx) \nu(dy) \\ &= \sum_{n\in\mathbb{N}} \sum_{k\le n} \varrho_n \widetilde{C}_k(n) \widetilde{P}_k(x) Q_n(y) \widetilde{\mu}(dx) \nu(dy)\end{aligned} \tag{9}$$

is still a probability measure, *where $\widetilde{C}_k(n)$ are the connection coefficients between P_n and the $\widetilde{\mu}$ orthonormal polynomials $\widetilde{P}_n$* (*see Appendix*). *The first margin of $\widetilde{\sigma}$ is $\widetilde{\mu}$ and the second one is given by*

$$\widetilde{\nu}(dy) = \sum_{n\in\mathbb{N}} \varrho_n \widetilde{C}_0(n) Q_n(Y) \nu(dy). \tag{10}$$

It is remarkable that the diagonal summation in (8) becomes triangular in (9). Therefore the joint distribution resulting from a change of margin is not in general a Lancaster probability.

An important fact is that the Lancaster sequence ϱ_n can be expressed in terms of the new margins as follows:

Proposition 3. *Let $(\widetilde{X}, \widetilde{Y})$ have a new joint distribution $\widetilde{\sigma}$ given in* (9). *Then*

$$\varrho_n = \mathbb{E}(P_n(\widetilde{X}) Q_n(\widetilde{Y})) \mathbb{E}(P_n(\widetilde{X}))^{-1}.$$

Proof. This follows from the orthogonality of the polynomials Q_n. ■

Thus the coefficients ϱ_n express correlations between $P_n(\widetilde{X})$ and $Q_n(\widetilde{Y})$ weighted by $\mathbb{E}(P_n(\widetilde{X}))^{-1}$. Note also that the form of the summation in (9) easily permits one to truncate the expansion and thus to get estimates.

3. Copula approach. Next it is of interest to compare the change of margins in Theorem 1 with the Sklar Theorem (see [14]) within the frame of copulas. This famous result permits one to change marginal distribution functions to obtain new joint distribution functions. Note that Sklar's Theorem allows any changes of margins although Theorem 1 only ensures one change of margins (see the counterexample given in Section 5). Obviously, some cases permit many changes of margins, like the independent case when σ is the product of μ and ν.

However, in the case of multivariate margins Sklar's Theorem breaks down (see [3]). In such a case, when the two margins are multidimensional, Theorem 1 is valid and then provides a way of generating new joint distributions with fixed multivariate margins.

From the Sklar Theorem (see [10] for a fuller treatment of copulas) it is well known that if a joint distribution function is expressed in terms of marginal ones, namely

$$F(x,y) = C(F_X(x), F_Y(y)), \tag{11}$$

where $C : [0,1] \times [0,1] \to [0,1]$ is a copula (that is, $C(u,0) = C(0,v) = 0$, $C(u,1) = C(1,v) = 1$ and C is 2-increasing), then changing F_X and F_Y in such an expression yields a new joint distribution function with some nice properties (see [14]). We can apply this result to the expansion given in (1). Write F, F_X and F_Y for the distribution functions associated to σ, μ and ν, respectively. Then there exists a copula, say C, such that (11) occurs. Differentiating twice this equality yields (under the condition of existence)

$$f(x,y) = f_X(x) f_Y(y) C''(F_X(x), F_Y(y)),$$

where f, f_X and f_Y denote the densities associated to σ, μ and ν, respectively. It is required here that the relevant derivatives exist. To simplify notation we write C'' instead of $(\partial/\partial x)(\partial/\partial y)C$. Comparing this expression with (1) we obtain the following relation:

$$C''(u,v) = \sum_{n,k\in\mathbb{N}} \varrho_{n,k} P_n(F_X^{-1}(u)) Q_k(F_Y^{-1}(v)), \tag{12}$$

assuming that F_X and F_Y are invertible.

THEOREM 4. *Let σ be a joint density function satisfying* (1) *and denote by F_X the distribution functions associated to the margins μ. Let $\widetilde{X}$ be a random variable with distribution measure $\widetilde{\mu}$ and with distribution function $\widetilde{F}_X$. Then the measure*

$$\widetilde{\sigma}(dx,dy) = \sum_{n,k\in\mathbb{N}} \varrho_{n,k} P_n(F_X^{-1}(\widetilde{F}_X(x))) Q_k(y) \widetilde{\mu}(dx) \nu(dy)$$

is still a probability measure with margins $\widetilde{\mu}$ and ν.

Proof. Applying Sklar's Theorem we can change F_X into $\widetilde{F}_X$. Hence

$$\widetilde{F}(x,y) = C(\widetilde{F}_X(x), F_Y(y)).$$

By taking second derivatives we finally obtain a new joint density as

$$\widetilde{f}(x,y) = \widetilde{f}_X(x) f_Y(y) C''(\widetilde{F}_X(x), F_Y(y)),$$

and using (11) we get

$$\widetilde{f}(x,y) = \widetilde{f}_X(x) f_Y(y) \sum_{n,k\in\mathbb{N}} \varrho_{n,k} P_n(F_X^{-1}(\widetilde{F}_X(x))) Q_k(y),$$

that is,

$$\widetilde{\sigma}(dx,dy) = \sum_{n,k\in\mathbb{N}} \varrho_{n,k} P_n(F_X^{-1}(\widetilde{F}_X(x))) Q_k(y) \widetilde{\mu}(dx)\nu(dy). \ \blacksquare$$

It is important to note that we can also change F_Y into $\widetilde{F}_Y$ to obtain a new joint distribution. However, it appears that the orthogonal functions $P_n(F_X^{-1}(\widetilde{F}_X))$ are polynomials if and only if there exists a polynomial h such that $\widetilde{F}(x) = F(h(x))$. Thus, to keep a polynomial expansion of the new joint distribution the change of the marginal variable X must have the form $\widetilde{X} = h^{-1}(X)$, where h is a polynomial. Very few interesting examples have this property. For instance we may consider $\widetilde{X} = -X$ or $\widetilde{X} = \sqrt{X}$ or any affinities. But if the degree of h is ≥ 2 it appears very difficult to obtain a simple expression for h^{-1}.

4. Distance of a new joint distribution from its parent. It is natural now to try to measure the distance between models (1) and (2). For instance, we may consider the case where $\widetilde{\mu}(dx) = \mu(dx)(1+p(x))$, as in an Edgeworth expansion. The following proposition shows how the distance depends of this quantity $p(x)$.

PROPOSITION 5. *Let σ be a joint distribution with polynomial expansion* (1). *Assume that $\widetilde{\mu}$ is absolutely continuous with respect to μ and write $h = d\widetilde{\mu}/d\mu$; that is, $\widetilde{\mu}(dx) = h(x)\mu(dx)$. Denote by $\widetilde{\sigma}$ the new joint distribution with margin changed from μ into $\widetilde{\mu}$. Then*

$$\int |\sigma - \widetilde{\sigma}|(dx,dy) \leq \Big(\int h(x)\,\widetilde{\mu}(dx) - 1\Big)^{1/2} \sum_{n,k\in\mathbb{N}} |\varrho_{n,k}|.$$

Proof. We have

$$\begin{aligned}\int |\sigma - \widetilde{\sigma}|(dx,dy) &= \sum_{n,k\in\mathbb{N}} |\varrho_{n,k}| \int |Q_k(y)|\,\nu(dy) \int |P_n(x)|\,|\mu - \widetilde{\mu}|(dx)\\ &\leq \sum_{n,k\in\mathbb{N}} |\varrho_{n,k}| \int |P_n(x)|\,|1-h(x)|\,\mu(dx)\\ &\leq \sum_{n,k\in\mathbb{N}} |\varrho_{n,k}| \Big(\int (1-h(x))^2\,\mu(dx)\Big)^{1/2},\end{aligned}$$

which establishes the formula. $\blacksquare$

Example 1. Let μ and $\widetilde{\mu}$ be Gaussian with unit variance and with two different means 0 and m. Consider the Lancaster distribution σ with margins $\mu = \nu$ with correlation $-1 < t < 1$. The Lancaster sequence is given by $\varrho_n = t^n$ (see [6]) and the associated orthonormal polynomials are Hermite polynomials. Then we have $h(x) = \exp(mx - m^2/2)$ and

$$\Big(\int h(x)\,\widetilde{\mu}(dx) - 1\Big)^{1/2} = (\exp(m^2) - 1)^{1/2}.$$

We obtain

$$\int |\sigma - \widetilde{\sigma}|(dx, dy) \le (\exp(m^2) - 1)^{1/2}(1 - |t|)^{-1}.$$

Example 2. Let μ be a normal distribution and let

$$\widetilde{\mu}(dx) = \mu(dx)\Big\{1 + \sum_{j=1}^{J} \alpha_j H_j(x)\Big\}$$

be an Edgeworth expansion, where H_j are Hermite polynomials and α_j are suitable constants. Then

$$\int |\sigma - \widetilde{\sigma}|(dx, dy) \le \sum_{j=1}^{J} |\alpha_j| \sum_{n,k\in\mathbb{N}} |\varrho_{n,k}|.$$

5. Illustrations

5.1. *Calculating some connection coefficients.* We use the recurrence relations given in [11] to find the values of some coefficients. The important simplification is that we only need the coefficients $\widetilde{C}_0$. Some examples are reported in Table 1, where $M_n^{(a,b)}, C_n^a, K_n^{(a,b)}$ are the Meixner, Charlier and Krawtchouk polynomials, respectively (see [5] for notation and properties of orthogonal polynomials) and where $(x)_n = x(x+1)\cdots(x+n-1)$ denotes the Pochhammer symbol.

Table 1. Connection coefficients between polynomials

Polynomials	$C_0(n)$
$C_n^{(a)} \to C_n^{(b)}$	$(b-a)^n$
$M_n^{(1,a)} \to M_n^{(1,b)}$	$n\{(b-a)/(1-a)(1-b)\}^n$
$K_n^{(a)}(\cdot, N) \to k_n^{(b)}(\cdot, N)$	$(N+1)_n(b-a)^n$
$M_N^{(c,1)} \to C_n^{(b)}$	$(-b)^n(n+1)!/(c)_n$

5.2. *Examples with connection coefficients.* Consider a joint distribution with margins $\mu = \mathcal{P}(m_0)$ (Poisson with mean m_0) and $\nu = \mathcal{N}B(1,p)$ (corresponding to the geometric distribution) and with expansion

$$\sigma(dx, dy) = \sum_{n,k\in\mathbb{N}} \varrho_{n,k} C_n^{(m_0)} M_n^{(p,1)} \mathcal{P}(m_0)(dx)\mathcal{N}B(1,p)(dy). \tag{13}$$

Poisson to Poisson. Changing μ into a new Poisson margin $\mathcal{P}(m_1)$ yields, from Table 1, a new joint distribution with margins $\widetilde{\mu} = \mathcal{P}(m_1)$ and

$$\widetilde{\nu}(dy) = \sum_{n,k\in\mathbb{N}} \varrho_{n,k}(m_0 - m_1)^n M_n^{(p,1)} \mathcal{NB}(p,1)(dy).$$

Negative binomial to Poisson. In (13), changing $\mathcal{NB}(1,p)$ into $\mathcal{P}(m_1)$ yields a new joint distribution with new margins $\widetilde{\nu} = \mathcal{P}(m_1)$ and

$$\widetilde{\mu}(dx) = \sum_{n,k\in\mathbb{N}} \varrho_{n,k}(-m_0)^n (n+1)!/(p)_n C_n^{(m_0)}\, \mathcal{P}(m_0)(dx).$$

5.3. *Counterexample.* We give here a counterexample to point out the limitations of Theorem 1. It proves that changing both μ and ν in (1) does not yield a new joint distribution in general.

Consider a Lancaster distribution with margins $\mu = \nu$ distributed as a Gaussian $\mathcal{N}(0,1)$ with correlation $-1 < t < 1$. Let $P_n = Q_n$ be the associated orthonormal Hermite polynomials. Changing both μ and ν into arbitrary new measures $\widetilde{\mu} = \widetilde{\nu} = \varphi$ in (8) yields

$$\widetilde{\sigma}(dx,dy) = \Big(1 + \sum_{n>0} \varrho_n P_n(x) P_n(Y)\Big)\varphi(dx)\varphi(dy),$$

where $\varrho_n = t^n$. Then

$$\iint \widetilde{\sigma}(dx,dy) = 1 + \sum_{n>0} \varrho_n \Big\{\int P_n(x)\,\varphi(dx)\Big\}^2 > 1,$$

as soon as $t > 0$, and $\widetilde{\sigma}$ is a positive measure but not with mass equal to 1.

Appendix: Connection polynomials. Consider two sequences of polynomials $\{P_n(x)\}$ and $\{Q_n(x)\}$; that is, each element $P_n(x)$ (resp. $Q_n(x)$) is an nth degree polynomial in x. The *connection problem* between them consists in finding the coefficients $C_k(n)$ in the following expression:

$$P_n(x) = \sum_{k=0}^{n} C_k(n) Q_k(x). \tag{14}$$

When both sequences are orthonormal, i.e.

$$\int_E P_n(x)P_k(x)\,\mu(dx) = \begin{cases} 1 & \text{if } n = k,\\ 0 & \text{otherwise,}\end{cases}$$

$$\int_E Q_n(x)Q_k(x)\,\nu(dx) = \begin{cases} 1 & \text{if } n = k,\\ 0 & \text{otherwise,}\end{cases}$$

they satisfy three-term recurrence relations

$$xP_n(x) = a_n P_{n+1}(x) + b_n P_n(x) + c_n P_{n-1}(x), \tag{15}$$

$$xQ_n(x) = \alpha_n Q_{n+1}(x) + \beta_n Q_n(x) + \gamma_n Q_{n-1}(x). \tag{16}$$

Then, multiplying (14) by x, using (15)–(16), and identifying with respect to Q_n yields

$$(17)\qquad (\beta_k - b_n)C_k(n) + \alpha_{k-1}C_{k-1}(n) + \gamma_{k+1}C_{k+1}(n) - a_nC_k(n+1) - c_nC_k(n-1) = 0.$$

Recurrence relations for connection coefficients are studied in various papers (see [11], [4], [12]). Note that in Theorem 1 we only need particular connection coefficients which simplify the use of the relation (17).

Acknowledgements. The author thanks the referee for his comments and suggestions which have helped to improve the original manuscript.

References

[1] R. Blacher, *Higher order correlation coefficients*, Statistics 25 (1993), 1–15.

[2] A. Buja, *Remarks on functional canonical variates, alternating least squares methods and ACE*, Ann. Statist. 18 (1990), 1032–1069.

[3] C. Genest, J. J. Quesada Molina et J. A. Rodríguez Lallena, *De l'impossibilité de construire des lois à marges multidimensionnelles données à partir de copules*, C. R. Acad. Sci. Paris Sér. I Math. 320 (1995), 723–726.

[4] E. Godoy, A. Ronveaux, A. Zarzo and I. Area, *Minimal recurrence relations for connection coefficients between classical orthogonal polynomials: Continuous case*, J. Comput. Appl. Math. 84 (1997), 257–275.

[5] R. Koekoek and R. F. Swarttouw, *The Askey-scheme of hypergeometric orthogonal polynomials and its q-analogue*, report no. 94-05, Faculty of Technical Mathematics and Informatics, Delft University of Technology, 1994.

[6] A. E. Koudou, *Lancaster bivariate probability distibutions with Poisson, negative binomial and gamma margins*, Test 7 (1998), 95–110.

[7] H. O. Lancaster, *The Chi-squared Distribution*, Wiley, New York, 1969.

[8] —, *Joint probability distributions in the Meixner classes*, J. Roy. Statist. Soc. Ser. B 37 (1975), 434–443.

[9] G. Matheron, *Isofactorial models and change of support*, in: Geostatistics for Natural Resources Characterization, G. Verly *et al.* (eds.), NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. 122, Reidel, Dordrecht, 1984, Part 1, 449–467.

[10] R. B. Nelsen, *An Introduction to Copulas*, Lecture Notes in Statist. 139, Springer, New York, 1999.

[11] A. Ronveaux, S. Belmehdi, E. Godoy and A. Zarzo, *Recurrence relation approach for connection coefficients. Application to classical discrete orthogonal polynomials*, in: Symmetries and Integrability of Difference Equations, CRM Proc. Lecture Notes 9, Amer. Math. Soc., Providence, RI, 1996, 319–335.

[12] J. Sánchez-Ruiz and J. S. Dehesa, *Some connection and linearization problems for polynomials in and beyond the Askey scheme*, J. Comput. Appl. Math. 133 (2001), 579–591.

[13] O. V. Sarmanov and Z. N. Bratoeva, *Probabilistic properties of bilinear expansions of Hermite polynomials*, Theoret. Probab. Appl. 12 (1967), 470–481.

[14] A. Sklar, *Fonctions de répartition à n dimensions et leurs marges*, Publ. Inst. Statist. Univ. Paris 8 (1959), 229–231.

[15] S. G. Tyan, H. Derin and J. B. Thomas, *Two necessary conditions on the representation of bivariate distributions by polynomials*, Ann. Statist. 4 (1976), 216–222.

[16] W. Vervaat, *Upper bounds for the distance in total variation between the binomial or the negative binomial and the Poisson distribution*, Statist. Neerlandica 23 (1970), 79–86.

CREST-ENSAI
rue Blaise Pascal
BP 37203
35172 Bruz Cedex, France
E-mail: pommeret@ensai.fr

Received on 27.4.2004;
revised version on 27.8.2004 (1741)