The Annals of Statistics
1976, Vol. 4, No. 1, 216–222

TWO NECESSARY CONDITIONS ON THE REPRESENTATION OF BIVARIATE DISTRIBUTIONS BY POLYNOMIALS[1]

BY SHU-GWEI TYAN, HALUK DERIN, AND JOHN B. THOMAS

Princeton University

Let X and Y be two unbounded random variables. Then two necessary conditions are proved concerning the structure of the bivariate distribution function of X and Y when it is expanded in the orthonormal polynomials of its marginal distributions. The first condition concerns the shrinking of the polynomial representation into a diagonal form, and the second is a generalization of the Sarmanov–Bratoeva theorem.

1. Introduction. Let X and Y be two random variables and let $F_X(x)$ and $F_Y(y)$ be their distribution functions and assume they have finite moments of every order. Then a sequence of polynomials $\{P_n(x)\}_{n=0}^{\infty}$, where $P_n(x)$ is of degree n, orthonormal with respect to $F_X(x)$ can be constructed. Here ∞ is to be replaced with N if X takes on only $N+1$ values. The orthonormal polynomials associated with $F_Y(y)$ can be constructed also and are denoted by $\{Q_m(y)\}_{m=0}^{\infty}$. Here $Q_m(y)$ is of degree m, and ∞ is to be replaced with M if Y takes on only $M+1$ values. Assume that $P_0(x) \equiv Q_0(y) \equiv 1$.

Let $L^2(F)$ be the space of the square integrable real functions with respect to the distribution F. If a sequence of polynomials is complete in both $L^2(F_X)$ and $L^2(F_Y)$, and if $g(x) \in L^2(F_X)$ and $h(y) \in L^2(F_Y)$, then for any bivariate distribution function $F_{X,Y}(x, y)$ having F_X and F_Y as marginals the following is true [5]:

$$\int_{R^2} g(x)h(y)\, dF_{X,Y}(x, y) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} g_n \rho_{n,m} h_m \,. \tag{1}$$

Here R^2 is the 2-dimensional real space and

$$\rho_{n,m} = \int_{R^2} P_n(x)Q_m(y)\, dF_{X,Y}(x, y)\,,$$

$\rho_{n,0} = \rho_{0,m} = 0$ for $n, m = 1, 2, 3, \cdots$, and

$$g_n = \int_{-\infty}^{\infty} g(x)P_n(x)\, dF_X(x)\,, \qquad h_m = \int_{-\infty}^{\infty} h(y)Q_m(y)\, dF_Y(y)\,. \tag{2}$$

The double series in (1) converges absolutely. In particular,

$$F_{X,Y}(x, y) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \rho_{n,m}\, p_{x,n}\, q_{y,m} \tag{3}$$

where

$$p_{x,n} = \int_{(-\infty,x]} P_n(u)\, dF_X(u)\,, \qquad q_{y,m} = \int_{(-\infty,y]} Q_m(v)\, dF_Y(v)\,. \tag{4}$$

The structure (3) is the subject of this paper and for related general material we refer to Lancaster [10].

Received August 1974; revised June 1975.

[1] This research is supported by the National Science Foundation under Grant GK-24187 and by the U. S. Army Research Office under Contract DAHCO4-69-0012.

AMS 1970 *subject classifications*. Primary 62E10; Secondary 60E05, 42A60.

Key words and phrases. Bivariate distribution function, orthonormal polynomials.

Conversely, let ℓ^2 be the Hilbert space of square summable real sequences. Then we have the following lemma:

LEMMA 1. *Let* $\{\sigma_{n,m}\}$ *be a real double sequence where* $n, m = 0, 1, 2, \cdots$. *Assume that* $\{\sigma_{n,m}\}$ *satisfies the following two conditions*:

$$\text{(5)}\qquad \begin{aligned} &\text{(i)}\quad \sigma_{n,0} = \sigma_{0,m} = 0 \quad \textit{for} \quad n, m = 1, 2, 3, \cdots, \\ &\text{(ii)}\quad \textstyle\sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \sigma_{n,m} a_n b_m \quad \textit{converges absolutely whenever} \quad \{a_n\}_{n=0}^{\infty} \quad \textit{and} \\ &\qquad \{b_m\}_{m=0}^{\infty} \quad \textit{are in} \quad \ell^2 . \end{aligned}$$

If $F(x, y)$ *is a bivariate distribution function satisfying* (3) *with* $p_{x,n}$, $q_{y,m}$ *defined in* (4) *and* $\rho_{n,m}$ *replaced by* $\sigma_{n,m}$, *then* $F(x, y)$ *has marginals* F_X *and* F_Y. *Also* $F(x, y)$ *satisfies* (1) *for all* $g(x)$ *and* $h(y)$ *in* $L^2(F_X)$ *and in* $L^2(F_Y)$, *respectively.*

PROOF. By (5), for any $\{a_n\}_{n=0}^{\infty}$ in ℓ^2, the sequence $\{\sum_{n=0}^{\infty} \sigma_{n,m} a_n\}_{m=0}^{\infty}$ is also in ℓ^2. Since

$$\sum_{m=0}^{\infty} q_{y,m}^2 \leqq 1 , \qquad \lim_{y\to\infty} q_{y,0} = 1 ,$$

and $\sigma_{n,0} = 0$ for $n = 1, 2, 3, \cdots$, thus for any fixed x

$$\lim_{y\to\infty} F(x, y) = \sigma_{0,0} \int_{(-\infty,x]} dF_X(u) .$$

Obviously, $\sigma_{0,0} = 1$ since $F(x, y)$ is a bivariate distribution by assumption. Thus $F(x, y)$ has marginals F_X and F_Y. Furthermore (5) implies that

$$\lim_{y\uparrow y_0} F(x, y) = \sum_{m=0}^{\infty} (\sum_{n=0}^{\infty} \sigma_{n,m} p_{x,n}) \int_{(-\infty,y_0)} Q_m(v)\, dF_Y(v) .$$

Therefore, for any two bounded intervals (open, closed, or half open) A and B,

$$\text{(6)}\qquad \begin{aligned} &\int_{R^2} I_A(x) I_B(y)\, dF(x, y) \\ &\qquad = \int_{A\times B} dF(x, y) \\ &\qquad = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \sigma_{n,m} [\int I_A(x) P_n(x)\, dF_X(x)][\int I_B(y) Q_m(y)\, dF_Y(y)] . \end{aligned}$$

Here $I_A(x)$ is the characteristic function of the subset A: $I_A(x) = 1$, if $x \in A$; $I_A(x) = 0$, otherwise. Obviously (6) also holds when $I_A(x)$ and $I_B(y)$ are replaced by step functions. Since step functions are dense in both $L^2(F_X)$ and $L^2(F_Y)$, we can find $g_i(x) \to g(x)$ in $L^2(F_X)$ and $h_j(y) \to h(y)$ in $L^2(F_Y)$, where $g_i(x)$ and $h_j(y)$ are step functions. Then

$$\begin{aligned} &\int_R g_i(x) h_j(y)\, dF(x, y) \\ &\qquad = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \sigma_{n,m} [\int g_i(x) P_n(x)\, dF_X(x)][\int h_j(y) Q_m(y)\, dF_Y(y)] . \end{aligned}$$

First let $i \to \infty$ then let $j \to \infty$. It can then be shown that $F(x, y)$ satisfies (1) by using (5) and the fact that $F(x, y)$ has marginals F_X and F_Y. □

Denote by Γ the class of bivariate distribution functions which satisfy the conditions in Lemma 1; that is, $F_{X,Y}(x, y)$ has an expansion (3) with $\rho_{n,m}$ replacing $\sigma_{n,m}$ and satisfying (i) and (ii) of Lemma 1. Also denote by $\Gamma_{i,j}$, i and j are nonnegative integers, the subclass of bivariate distribution functions in Γ which have a representation (3) where $\rho_{n,m} = 0$ if $n - m > i$ or $m - n > j$; that is, the matrix $[\rho_{n,m}]$ is zero outside a uniform strip along the main diagonal.

In the following two sections, the role of unbounded random variables in the structure of a bivariate distribution function belonging to $\Gamma_{i,j}$ is studied.

2. A necessary condition on $\Gamma_{i,j}$. Let X and Y be two random variables with distributions F_X and F_Y, and let their joint distribution be $F_{X,Y}(x, y)$. Assume that F_X and F_Y have finite moments of every order and that $P_n(x)$ and $Q_m(y)$ are orthonormal polynomials as described at the beginning of Section 1.

LEMMA 2. *If $F_{X,Y}(x, y)$ is in Γ then the conditional expectations*

$$E[P_n(X) \mid Y] = \sum_{m=0}^{\infty} \rho_{n,m} Q_m(Y) \tag{7}$$

$$E[Q_m(Y) \mid X] = \sum_{n=0}^{\infty} \rho_{n,m} P_n(X) \tag{8}$$

are true. The series converge in quadratic mean.

PROOF. From Lemma 1, for any Borel subset B,

$$\int_{R^2} P_n(x) I_B(y)\, dF_{X,Y}(x, y) = \sum_{m=0}^{\infty} \rho_{n,m} \int I_B(y) Q_m(y)\, dF_Y(y) .$$

By (5) $\{\rho_{n,m}\}_{m=0}^{\infty}$ is in ℓ^2 for any fixed n, thus (7) is true; (8) is then obvious. □

The following theorem due to Derin and Thomas [6] was proved by Brown [4] for the case $\Gamma_{0,0}$.

THEOREM 1. *Suppose $F_{X,Y}(x, y)$ is in Γ. Then $F_{X,Y}(x, y)$ is in $\Gamma_{i,j}$ if and only if*

(i) *$E(X^k \mid Y) =$ a polynomial in Y of degree less than or equal to $k + j$,*

and

(ii) *$E(Y^k \mid X) =$ a polynomial in X of degree less than or equal to $k + i$*

for $k = 0, 1, 2, \cdots$. Here equality is in quadratic mean and hence almost surely.

PROOF. The necessity is a direct result of Lemma 2. The sufficiency follows directly from the fact that

$$\rho_{n,m} = E[P_n(X) Q_m(Y)] = E\{E[P_n(X) \mid Y] Q_m(Y)\} ,$$

which is zero if $m - n > j$. By analogy, $\rho_{n,m} = 0$ if $n - m > i$. Thus $F_{X,Y}(x, y) \in \Gamma_{i,j}$. □

The polynomial regression property expressed in (i) and (ii) of Theorem 1 is interesting. However, very little is known about the class $\Gamma_{i,j}$ except for the special case $\Gamma_{0,0}$ which has a diagonal expansion (see Bochner [3], Sarmanov and Bratoeva [11], Askey [2], Gasper [8] and [9], and Eagleson [7]). For applications of Γ to nonlinear analysis, refer to Cambanis and Liu [5].

In the following it is shown that if $F_{X,Y}(x, y)$ is in $\Gamma_{i,j}$ but does not belong to $\Gamma_{0,0}$ then at least one of the random variables X and Y is bounded.

LEMMA 3. *If Y is unbounded, and if there exists a $k \geqq 0$ such that $\rho_{n,m} = 0$ for all $m - n > k$, then $\rho_{n,m} = 0$ if $m > n$.*

PROOF. Assume $\rho_{j,j+k} \neq 0$ for some j; otherwise k is replaced with $k - 1$ and so on until a nonzero term is obtained. Then

$$E[P_j(X) \mid Y] = \sum_{m=0}^{j+k} \rho_{j,m} Q_m(Y)$$

and

$$E[P_j^{2k}(X) \mid Y] \geqq \{E[P_j(X) \mid Y]\}^{2k} = [\sum_{m=0}^{j+k} \rho_{j,m} Q_m(Y)]^{2k} . \tag{9}$$

The L.H.S. of (9) is a polynomial of degree at most $2kj + k$ while the R.H.S. of (9) is a polynomial of degree at least $2k(j + k)$. Since Y is unbounded, the inequality (9) holds only if

$$2kj + k \geqq 2k(j + k) .$$

This implies that $k = 0$. □

THEOREM 2. *If both X and Y are unbounded and if $F_{X,Y}(x, y)$ belongs to $\Gamma_{i,j}$, then $F_{X,Y}(x, y)$ belongs to $\Gamma_{0,0}$.*

PROOF. That $F_{X,Y}(x, y) \in \Gamma_{i,j}$ implies that $\rho_{n,m} = 0$ if $m - n > j$ or $n - m > i$. Since Y is unbounded, it follows from Lemma 3 that $\rho_{n,m} = 0$ for all $m > n$. Since X is also unbounded, thus $\rho_{n,m} = 0$ for all $n > m$. Therfore $F_{X,Y}(x, y) \in \Gamma_{0,0}$. □

3. Coefficients of $\Gamma_{0,0}$. Sarmanov and Bratoeva [11] showed that a bivariate density function $f(x, y)$ can be represented as a diagonal expansion in normalized Hermite polynomials

$$f(x, y) = \frac{1}{2\pi} [1 + \sum_{k=1}^{\infty} c_k H_k(x) H_k(y)] \exp\{-(x^2 + y^2)/2\}$$

where the series converges in the mean, if and only if the sequence $\{c_k\}$ is the moment sequence of some probability distribution concentrated in the open interval $(-1, 1)$ and $\sum_{k=1}^{\infty} c_k^2 < \infty$.

In this section we show that the necessity part of the Sarmanov–Bratoeva theorem is true for general orthonormal polynomials provided the common marginal distribution has unbounded support. The method of proof is suggested in [11].

THEOREM 3. *Let the bivariate distribution function $F_{X,Y}(x, y)$ of X and Y belong to $\Gamma_{0,0}$ and let $F_X(u) \equiv F_Y(u) \equiv F(u)$. Write*

$$F_{X,Y}(x, y) = \sum_{n=0}^{\infty} c_n p_{x,n} p_{y,n} , \tag{10}$$

where

$$p_{x,n} = \int_{(-\infty,x]} P_n(u)\, dF(u) .$$

The $P_n(u)$, $n = 0, 1, 2, \cdots$, are the polynomials orthonormal with respect to $F(u)$. If the random variable X is unbounded then

$$c_n = \int_{[-1,1]} u^n \, dG(u)$$

where $G(u)$ is a probability distribution function with $c_0 = 1$.

PROOF. Clearly $c_0 = \rho_{0,0} = 1$ as shown in Lemma 1. By Lemma 2

$$E[P_n(X) \mid Y] = c_n P_n(Y)\,, \qquad n = 0, 1, 2, \cdots .$$

Obviously, for arbitrary real numbers $z_0, z_1, z_2, \cdots$ and real x

$$\sum_{i=0}^{n} \sum_{j=0}^{n} z_i z_j x^{i+j} \geqq 0\,.$$

Let $x^n = \sum_{k=0}^{n} a_{n,k} P_k(x)$ for $n = 0, 1, 2, \cdots$ and define

$$e_i(x) = \sum_{k=0}^{i} a_{i,k} c_k P_k(x)\,, \qquad i = 0, 1, 2, \cdots .$$

Then $E(X^n \mid Y) = e_n(Y)$ and

$$\sum_{i=0}^{n} \sum_{j=0}^{n} z_i z_j e_{i+j}(Y) = E[\sum_{i=0}^{n} \sum_{j=0}^{n} z_i z_j X^{i+j} \mid Y] \geqq 0 \tag{11}$$

almost surely. Since the left hand side of (11) is a polynomial and hence is continuous in y, thus

$$\sum_{i=0}^{n} \sum_{j=0}^{n} z_i z_j e_{i+j}(y) \geqq 0$$

for all y in the support, denoted by S, of F. Therefore by [1], the sequence $\{e_n(y)\}_{n=0}^{\infty}$ is a moment sequence for all $y \in S$, and $\{e_n(y)y^{-n}\}_{n=0}^{\infty}$ is also a moment sequence for any y, $y \neq 0$ and $y \in S$. Thus

$$c_n = \lim_{|y| \to \infty} e_n(y) y^{-n}$$

is also a moment sequence. By (5), c_n is bounded; hence

$$c_n = \int_{[-1,1]} u^n \, dG(u)\,. \qquad \square$$

COROLLARY 1. *If, in addition,* $X \geqq a$ *almost surely or if* $X \leqq b$ *almost surely, where* a *and* b *are both finite real numbers, then*

$$c_n = \int_{[0,1]} u^n \, dG(u)\,.$$

PROOF. We consider only the case when $X \leqq a$ almost surely; the proof of the other case is similar. Define

$$\tilde{e}_i(x) = \sum_{k=0}^{i} \sum_{j=0}^{k} \binom{i}{k} (-a)^{i-k} a_{k,j} c_j P_j(x)\,,$$

for $i = 0, 1, 2, \cdots$. Then $\tilde{e}_n(Y) = E[(X - a)^n \mid Y]$ and

$$\sum_{i=0}^{n} \sum_{j=0}^{n} z_i z_j \tilde{e}_{i+j}(Y) \geqq 0\,, \qquad \sum_{i=0}^{n} \sum_{j=0}^{n} z_i z_j \tilde{e}_{i+j+1}(Y) \geqq 0$$

almost surely. By the continuity of $\tilde{e}_n(y)$, the sequence $\{\tilde{e}_n(y)\}_{n=0}^{\infty}$ satisfies the Stieltjes moment problem [1] for all $y \in S$. For $y > 0$ with $y \in S$, the sequence $\{\tilde{e}_n(y)y^{-n}\}_{n=0}^{\infty}$ is also a Stieltjes moment sequence. Since the support S has no upper bound, thus

$$c_n = \lim_{y \to \infty} \tilde{e}_n(y) y^{-n} = \int_{[0,1]} u^n \, dG(u)\,. \qquad \square$$

However, the converse is not true (it is true for some F's; for example, normal, Poisson, generalized gamma, and negative-binomial). A counterexample is the density function

$$f(x) = \frac{1}{\pi} |\Gamma(\tfrac{1}{2} + ix)|^2 = \frac{1}{\cosh(\pi x)}\,, \quad -\infty < x < \infty\,.$$

Let $c_n = \rho^n$, $0 < \rho < 1$, and let $F'(x) = f(x)$. The orthonormal polynomials $P_n(x)$ associated with $F(x)$ are the Pollaczek polynomials of infinite interval [12]. If $F_{X,Y}(x, y)$, constructed as in (10), is a bivariate distribution, then

$$\phi(u, -u) = \int_{R^2} e^{iux-iuy}\, dF_{X,Y}(x, y)$$

should be a characteristic function. But

$$\phi(u, -u) = \frac{1 + \cosh(\theta)}{\sinh(\theta)} \int_{-\infty}^{\infty} \frac{\sin(\theta x)}{\sin(\pi x)}\, e^{iux}\, dx$$

where $\cosh(\theta) = (1 + \rho)/(1 - \rho)$ and is not a characteristic function.

COROLLARY 2. *Let F be fixed. Let $\{c_n\}_{n=0}^{\infty}$ and $\{d_n\}_{n=0}^{\infty}$ be two sequences such that for each the $F_{X,Y}(x, y)$ of* (10) *is a bivariate distribution* (X *and* Y *may be bounded or not*). *Then $\{c_n d_n\}_{n=0}^{\infty}$ also generates a bivariate distribution of the structure* (10).

PROOF. Let

$$F_{X,Z}(x, z) = \sum_{n=0}^{\infty} c_n p_{x,n} p_{z,n}\,,$$
$$F_{Y,Z}(y, z) = \sum_{n=0}^{\infty} d_n p_{y,n} p_{z,n}\,,$$

and let A, B, and C be three arbitrary Borel subsets. Then

$$\int_{R^2} I_A(x) I_C(z)\, dF_{X,Z}(x, z) = \int_C g_A(z)\, dF(z) \geqq 0$$

where

$$g_A(z) = \sum_{n=0}^{\infty} c_n \left[\int_A P_n(x)\, dF(x)\right] P_n(z)\,,$$

and where the convergence is in $L^2(F)$. Clearly $g_A(z) \geqq 0$ almost surely. Likewise, define $h_B(z)$ by

$$h_B(z) = \sum_{n=0}^{\infty} d_n \left[\int_B P_n(y)\, dF(y)\right] P_n(z)\,,$$

which is also nonnegative almost surely. Thus

$$\sum_{n=0}^{\infty} c_n d_n \left[\int_A P_n(x)\, dF(x)\right]\left[\int_B P_n(y)\, dF(y)\right] = \int_{-\infty}^{\infty} g_A(z) h_B(z)\, dF(z) \geqq 0\,.$$

Obviously, the function $F_{X,Y}(x, y) = \sum_{n=0}^{\infty} c_n d_n p_{x,n} p_{y,n}$ is a bivariate distribution function. □

Define $F(x, y; u) = \sum_{n=0}^{\infty} u^n p_{x,n} p_{y,n}$, where $-1 \leqq u \leqq 1$. Consider the set $T = \{u : F(x, y; u)$ is a bivariate distribution function$\}$. Then it is trivial to show that T is closed. By Corollary 2, T is closed under multiplication. We give a partial converse to Theorem 3.

COROLLARY 3. *Let $G(u)$ be a probability distribution function with its support contained in T. Then $F(x, y) = \sum_{n=0}^{\infty} c_n p_{x,n} p_{y,n}$ is a bivariate distribution where $c_n = \int_{-1}^{1} u^n\, dG(u)$.*

PROOF. Since $F(x, y; u)$ is continuous in u and since

$$\sum_{n=0}^{\infty} c_n p_{x,n} p_{y,n} = \int_T F(x, y; u)\, dG(u)\,,$$

it follows that $F(x, y)$ is a bivariate distribution in $\Gamma_{0,0}$. □

In this section, only the symmetric case, i.e. $F_X(u) \equiv F_Y(u)$, is discussed; for the asymmetric case, we refer to Tyan and Thomas [13], and Tyan [14].

REFERENCES

[1] AKHIEZER, N. I. (1965). *The Classical Moment Problem and Some Related Questions in Analysis*. Hafner, New York.

[2] ASKEY, R. (1970). Orthogonal polynomials and positivity. *Studies in Applied Mathematics* 6, *Wave Propagation and Special Functions*. SIAM, Philadelphia.

[3] BOCHNER, S. (1954). Sturm-Liouville and heat equations whose eigenfunctions are ultraspherical polynomials or associated Bessel functions. *Proc. Conf. on Differential Equations*. Univ. of Maryland, 23–48.

[4] BROWN, J. L. (1958). A criterion for the diagonal expansion of a second-order probability distribution in orthogonal polynomials. *IRE Trans. Information Theory* **4** 72.

[5] CAMBANIS, S. and LIU, B. (1971). On the expansion of a bivariate distribution and its relationship to the output of a nonlinearity. *IEEE Trans. Information Theory* **17** 17–25.

[6] DERIN, H. and THOMAS, J. B. (1972). On the representation of bivariate distributions. *Proc. Tenth Annual Allerton Conference on Circuit and System Theory*. 293–302 Univ. of Illinois, Urbana-Champaign.

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

[8] GASPER, G. (1972). Banach algebras for Jacobi series and positivity of a kernel. *Ann. of Math.* **95** 261–280.

[9] GASPER, G. (1973). Nonnegativity of a discrete Poisson kernel for the Hahn polynomials. *J. Math. Anal. Appl.* **42** 438–451.

[10] LANCASTER, H. O. (1969). *The Chi-squared Distribution*. Wiley, New York.

[11] SARMANOV, O. V. and BRATOEVA, Z. N. (1967). Probabilistic properties of bilinear expansions of Hermite polynomials. *Theor. Probability Appl.* **12** 470–481.

[12] SZEGÖ, G. (1967). *Orthogonal Polynomials*, Amer. Math. Soc. Colloq. Pub. **23** Amer. Math. Soc., Providence.

[13] TYAN, S. and THOMAS, J. B. (1975). Characterization of a class of bivariate distribution functions. *J. Multivariate Anal.* **5** 227–235.

[14] TYAN, S. (1975). The structure of bivariate distribution functions and their relation to Markov processes. Ph. D. dissertation, Dept. of Electrical Engineering, Princeton Univ.

DEPARTMENT OF ELECTRICAL ENGINEERING
PRINCETON UNIVERSITY
PRINCETON, NEW JERSEY 08540