Ann. Inst. Statist. Math.
Vol. 54, No. 4, 731–742 (2002)

GENERALIZED PEARSON DISTRIBUTIONS AND RELATED CHARACTERIZATION PROBLEMS

CLAUDE LEFÈVRE[1], VASILIS PAPATHANASIOU[2] AND SERGEY UTEV[3]

[1] *Institut de Statistique et de Recherche Opérationnelle, C.P.210, Université Libre de Bruxelles, Boulevard du Triomphe, B-1050 Bruxelles, Belgique,* e-mail: clefevre@ulb.ac.be
[2] *Department of Mathematics, University of Athens, Panepistemiopolis, 15784 Athens, Greece,* e-mail: bpapath@cc.uoa.gr
[3] *School of Mathematical Sciences, University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom,* e-mail: sergey.utev@nottingham.ac.uk

(Received November 24, 2000; revised September 10, 2001)

Abstract. This paper is concerned with a Rodrigues' formula for generating new classes of polynomials. The associated density functions belong to the family of generalized Pearson curves, which extends the classical Pearson family. Various properties of these polynomials (degree, orthogonality ...) are investigated and then exploited to derive several related results, especially characterizations, in probability.

Key words and phrases: Rodrigues' formula, generalized Pearson family of curves, orthogonal polynomials, covariance identity, characterizations.

1. Introduction

Any simple technique for generating new classes of polynomials, especially orthogonal, is potentially useful in different fields of mathematics, from approximation theory to probability and statistics. In this work, we are going to discuss such a polynomials generator that is constructed through a Rodrigues' formula. We will point out that the generator is closely related to different problems and characterizations in probability theory.

Specifically, we begin by considering a particular linear operator of differential form. For that, let f be some given real function, assumed to be infinitely differentiable on an interval $\mathcal{I} = (a, b)$, finite or not. Then, for any function z with similar properties, we introduce the operator D_f by stipulating that

$$D_f(z) = \frac{1}{f} D(fz), \tag{1.1}$$

where $D(h) \equiv h'$ denotes the usual differentiation. Now, let $z \equiv p$ be any polynomial of degree (denoted by $\deg(p)$) equal to n, $n \geq 0$. With the operator above, we then associate a sequence of functions $h_j \equiv h_{j|p,f}$ defined as

$$h_j = D_f{}^j(p^j) \equiv \frac{1}{f} D^j(fp^j), \qquad j \geq 0, \tag{1.2}$$

D^j denoting j-fold differentiation (thus, $h_0 = 1$, $h_1 = p' + pf'/f, \ldots$).

Our starting point is to ask the following question. Under what conditions on f the sequence of functions h_j, $j \geq 1$, constitutes a sequence of polynomials. It is worth underlining that if the h_j's are polynomials, then the identities (1.2) correspond to a *Rodrigues' formula*. A necessary condition on f is almost immediate. Indeed, from (1.2) for $j = 1$, we get that $f'/f = (h_1 - p')/p \equiv q/p$ is a ratio of two polynomials. Such a family of functions f, i.e. satisfying the ordinary differential equation

$$(1.3) \quad \frac{f'}{f} = \frac{q}{p}, \qquad q \text{ and } p \text{ being polynomials with} \deg(q) = m \text{ and } \deg(p) = n,$$

is called a family of *generalized Pearson curves*.

We recall that the classical Pearson family of curves corresponds to the special case of (1.3) with degrees $m \leq 1$ and $n \leq 2$ (see, e.g., Ord (1972), p. 9 and Johnson *et al.* (1994), p. 16). The Pearson family and the related orthogonal polynomials have received much attention in the literature. In particular, we refer the reader to the recent fascinating paper by Diaconis and Zabell (1991) which has motivated the present research. We mention that estimation by a generalized Pearson curve has been discussed by Dunning and Hanson (1977); for some other properties, see also, e.g., Ord (1972) and the Appendix.

Sections 2 and 3 of the paper are devoted to an algebraic analysis of the polynomials h_j and the associated (generalized) Pearson curves. In Section 2 we prove that the condition (1.3) on f is also sufficient to guarantee that the functions h_j, $j \geq 0$, reduce to polynomials. Furthermore, we establish that $\deg(h_j) \leq jr$ where the index r is defined as $r \equiv \max(n-1, m)$, and that strict inequality holds for some j if, and only if, $m = n - 1 \geq 1$ and under a supplementary condition on f. Conversely, in Section 3 we study some properties of the generalized Pearson curves (1.3) on the basis of the polynomials h_j's. We so obtain a characterization of the Pearson family, and we give a simple sufficient condition that allows us to determine the index r.

In the second part of the paper (Section 4), we exploit properties of the polynomials h_j to discuss several applications and characterization problems in probability theory. We will assume now that, in addition to the condition (1.3), f is a probability density function which is strictly positive on its support (a, b). Note that since Kf, for any constant K, satisfies (1.3) whenever f does, given two polynomials p and q this new assumption imposes a restriction on the support (a, b) to guarantee that f is positive on (a, b); this question will not be examined here (see, e.g., Johnson *et al.* (1994), pp. 16–25 and Dunning and Hanson (1977)).

The associated generating operator D_f defined in (1.1) is rather standard in probability. In particular, Cacoullos and Papathanasiou (1995) have derived the following covariance identity which involves D_f. Let X denote a continuous random variable with density function f. Given any differentiable function g and any function h, there exists some function $z \equiv z_{f,h}$ such that

$$(1.4) \qquad \operatorname{cov}[h(X), g(X)] = E[z(X)g'(X)],$$

provided that $z(X)g'(X)$ is integrable. This z-function is related to f and h through the differential equation

$$(1.5) \qquad D_f(z) = E(h) - h,$$

where D_f is given precisely by (1.1). Moreover, if (1.4) holds for every differentiable function g, then the functions z, f and h are related by (1.5). In the special situation where h is the identity function, z reduces to the w-function introduced by Cacoullos and Papathanasiou (1989), multiplied by $\mathrm{Var}(X)$.

Now, within our generalized framework, we begin by deriving some variant of the covariance identity (1.4), provided that an additional important condition (4.1) holds. Then, we apply this result to examine whether the system of polynomials $\{h_j, j \geq 0\}$ is (partially) orthogonal. Finally, we deduce characterizations of the Pearson, normal, gamma and beta density functions that rely on such orthogonality properties.

2. Algebraic structure of the polynomials

We propose to point out the Rodrigues' formula (1.2) as a key tool for generating polynomials h_j from a generalized Pearson curve f.

Given a polynomial t, we denote by $a(t)$ the coefficient of the power with highest degree in t. Without loss of generality, we may assume that $a(p) = 1$.

THEOREM 2.1. *Let f be of the form* (1.3). *Then, h_j defined in* (1.2) *is a polynomial such that*

$$\deg(h_j) \leq jr, \quad \text{where} \quad r = \max(n-1, m), \; j \geq 1, \tag{2.1}$$

and strict inequality holds for some j if, and only if, $m = n - 1 \geq 1$ and $a(q) = i - jn$ for some $i \in \{0, \ldots, j-1\}$.

Remark. We notice that in the special case of the Pearson family ($r \leq 1$), (2.1) reduces to $\deg(h_j) \leq j$, a result that can be found in Diaconis and Zabell ((1991), Theorem 1, p. 295). The condition yielding equality for some j then corresponds to the condition derived by Beale ((1941), p. 99–100), namely either $n = 1$, or $n = 2$ and $a(q) \notin \{-2j, 1-2j, \ldots, j-1-2j\}$.

Theorem 2.1 will be proved by using certain algebraic properties established in the four following lemmas.

LEMMA 2.1. *For any $0 \leq k \leq j$, there exists some polynomial $r_{j,k}$ of degree $k(n-1)$ such that*

$$D^k(p^j) = p^{j-k} r_{j,k}. \tag{2.2}$$

PROOF. Obviously, $D^k(p^j)$ is a polynomial of degree $jn - k$. Thus, if (2.2) holds for some polynomial $r_{j,k}$, then its degree satisfies the relation $jn - k = (j-k)n + \deg(r_{j,k})$, which yields $\deg(r_{j,k}) = k(n-1)$ as stated. To establish (2.2), we proceed by induction on k. (2.2) is true for $k = 0$. Assuming that (2.2) holds for k, we get

$$D^{k+1}(p^j) = D(p^{j-k} r_{j,k}) = p^{j-k-1}[(j-k)D(p)r_{j,k} + pD(r_{j,k})] \equiv p^{j-k-1} r_{j,k+1},$$

hence the result. □

LEMMA 2.2. *If for some $k \geq 1$, q_k is a polynomial of degree m_k, then*

$$D\left(f\frac{q_k}{p^k}\right) = f\frac{q_{k+1}}{p^{k+1}}, \tag{2.3}$$

where q_{k+1} is some polynomial of degree m_{k+1} such that $m_{k+1} = m_k + m$ if $m \geq n$, and $m_{k+1} \leq m_k + n - 1$ if $m \leq n - 1$.

PROOF. From (1.3), we get the representation

$$D\left(f\frac{q_k}{p^k}\right) = f\frac{1}{p^{k+1}}[q_k q + D(q_k)p - kq_k D(p)] \equiv f\frac{1}{p^{k+1}}q_{k+1}.$$

We then easily obtain the stated result on q_{k+1}. □

LEMMA 2.3. *Assume that $m \neq n-1$. Then, for any $i \geq 1$, there exists some polynomial q_i such that*

$$D^i(f) = f\frac{q_i}{p^i}, \tag{2.4}$$

and m_i, the degree of q_i, is such that $m_i = im$ if $m \geq n$, and $m_i \leq i(n-1)-1$ if $m \leq n-2$.

PROOF. For $i=1$, (2.4) corresponds to (1.3) with $m_1 = m$. Arguing by induction on i, we can then deduce (2.4) from (2.3). As for m_i, again by induction, we find from Lemma 2.2 that if $m \leq n-2$,

$$m_{i+1} \leq m_i + n - 1 \leq \cdots \leq m_1 + i(n-1) \leq n - 2 + i(n-1) = (i+1)(n-1) - 1,$$

while $m_i = im$ if $m \geq n$, which completes the proof. □

LEMMA 2.4. *Assume that $m = n-1$. Then, for any $0 \leq k \leq j$, there exists some polynomial $q_{j,k}$ such that*

$$D^k(fp^j) = fp^{j-k}q_{j,k}, \tag{2.5}$$

and its degree is such that $\deg(q_{j,k}) \leq k(n-1)$, *with strict inequality if, and only if, $n \geq 2$ and $a(q) = i - jn$ for some $i \in \{0,\ldots,k-1\}$.*

PROOF. Equation (2.5) being true for $k=0$, we proceed by induction on k. If (2.5) holds for k, we get, using (1.3),

$$D^{k+1}(fp^j) = fp^{j-k-1}[qq_{j,k} + (j-k)D(p)q_{j,k} + pD(q_{j,k})] \equiv fp^{j-k-1}q_{j,k+1}. \tag{2.6}$$

From the expression $[\ldots]$ of $q_{j,k+1}$, it is clear that $q_{j,k+1}$ is a polynomial. By induction and since $m = n-1$, we then see that its degree satisfies

$$\deg(q_{j,k+1}) \leq \deg(q_{j,k}) + n - 1 \leq \cdots \leq (k+1)(n-1).$$

Now, for $n \geq 2$, suppose that $a(q) \neq i - jn$ for all $i \in \{0,\ldots,k-1\}$. Proceeding by induction, we have that $\deg(q_{j,k}) = k(n-1)$. From $[\ldots]$, we then obtain that $a(q_{j,k+1})$, the coefficient of the highest degree in $q_{j,k+1}$, is given by

$$a(q_{j,k+1}) = [a(q) + jn - k]a(q_{j,k}). \tag{2.7}$$

Thus, we see that $a(q_{j,k+1}) \neq 0$ if, and only if, the additional condition $a(q) \neq k - jn$ holds. The result follows directly. □

PROOF OF THEOREM 2.1. We are going to discuss separately the cases where $m \neq$ or $= n-1$. We note that in all cases, $a(h_j)$ can be determined explicitly.

Case $m \neq n-1$. We will show that here, h_j is a polynomial with $\deg(h_j) = jr$. First, we observe that h_j can be rewritten as

$$h_j = \sum_{i=0}^{j} \binom{j}{i} h_{j,i}, \tag{2.8}$$

where

$$h_{j,i} \equiv [D^{j-i}(p^j)] \left[\frac{1}{f} D^i(f)\right]. \tag{2.9}$$

By Lemma 2.1, $h_{j,0}$ is a polynomial with $\deg(h_{j,0}) = j(n-1)$. For any $1 \leq i \leq j$, we get from Lemmas 2.1 and 2.3 that

$$h_{j,i} = p^i r_{j,j-i} \frac{q_i}{p^i} = r_{j,j-i} q_i.$$

This shows that $h_{j,i}, 1 \leq i \leq j$, is a polynomial, and that its degree is equal to $(j-i)(n-1) + m_i$, so that $\deg(h_{j,i}) = (j-i)(n-1) + im$ if $m \geq n$, and $\deg(h_{j,i}) \leq j(n-1) - 1$ if $m \leq n-2$. Therefore, we see from (2.8) that h_j is a polynomial. Moreover, if $m \geq n$, $h_{j,j}$ is of degree jm, while $h_{j,i}$ for $0 \leq i \leq j-1$ is of degree strictly less than jm; thus, $\deg(h_j) = jm = jr$. If $m \leq n-2$, $h_{j,i}$ for $1 \leq i \leq j$ is of degree strictly less than $j(n-1)$, which is the degree of $h_{j,0}$; thus, $\deg(h_j) = j(n-1) = jr$.

Case $m = n-1$. From (1.2) and Lemma 2.4, we have

$$h_j = \frac{1}{f} D^j(fp^j) = q_{j,j},$$

so that h_j is a polynomial with $\deg(h_j) \leq j(n-1) = jr$. Moreover, $\deg(h_j) < j(n-1)$ for some j if, and only if, $n \geq 2$ and $a(q) = i - jn$ for some $i \in \{0, \ldots, j-1\}$. □

3. Algebraic analysis of the density curves

We now propose to point out some reciprocal effects of the structure of the polynomials h_j on the generalized Pearson curve f.

The following lemma will play a central role.

LEMMA 3.1. *Let f be of the form* (1.3) *with* $n \geq 2$. *If* $\deg(h_j) < jr$ *and* $\deg(h_k) < kr$ *for some integers* $1 \leq j < k$, *then* $k - 1 \geq (k-j)(r+1)$.

PROOF. By Theorem 2.1, we know that under the above condition, we have $m = n-1 = r$ and $a(q) = i_1 - jn = i_2 - kn$ for some $i_1 \in \{0, \ldots, j-1\}$ and $i_2 \in \{0, \ldots, k-1\}$. Thus, we see that

$$k - 1 \geq i_2 = i_1 + (k-j)n \geq (k-j)n = (k-j)(r+1),$$

as stated. □

This result allows us to obtain a characterization of the Pearson family within the family of generalized Pearson curves through properties of the polynomials h_j. Let $\mathcal{L}(.)$ denote a linear span, and let us consider the subspace $\mathcal{LH} = \mathcal{L}(\{h_j, j \geq 0\})$ in a linear space of polynomials $\mathcal{LP}$. We recall that the codimension of $\mathcal{LH}$ is the dimension of its supplement (roughly, it is the difference $|\mathcal{LP}| - |\mathcal{LH}|$). We observe that $\mathcal{LH}$ is finite for $r = 0$, so that hereafter we will consider the case $r \geq 1$.

THEOREM 3.1. *Let f be of the form (1.3) with $r \geq 1$. Then, f is a Pearson curve (with $r = 1$) if, and only if,*

$$\text{(3.1)} \qquad \textit{the codimension of } \mathcal{LH} \textit{ is finite.}$$

Another equivalent characterization is that

(3.2) $\deg(h_j) \leq j$ *and* $\deg(h_k) \leq k$ *for some integers* $1 \leq j \leq k$ *such that* $3j < 2k + 1$.

PROOF. We have to establish that $r = 1$ may be expressed as in (3.1) or (3.2). For the former, we are going to show that

$$\text{(3.3)} \qquad \deg(h_j) = jr \quad \text{for all} \quad j, \text{ but a finite number.}$$

Indeed, if $n \leq 1$, then $r = m$ and by Theorem 2.1, $\deg(h_j) = jr$ for all $j \geq 1$. Consider $n \geq 2$, and let us assume that $\deg(h_i) < ir$ for some $i \geq 1$. Note that

$$j - 1 < (j - i)2 \leq (j - i)n \quad \text{for all} \quad j \geq 2i.$$

From Lemma 3.1, we then deduce that $\deg(h_j) = jr$ for all $j \geq 2i$. Therefore (3.3) is proved, and the equivalence follows now directly.

For the latter, we have by Theorem 2.1 that $r \geq 1$ yields $\deg(h_j) \leq j$ for all $j \geq 1$. Now, we will show that (3.2) implies $r = 1$. To the contrary, assume $r \geq 2$. We then have that $\deg(h_j) < jr$ and $\deg(h_k) < kr$ for some $1 \leq j < k$. Thus, from Theorem 2.1 we know that $n \geq 2$, and by Lemma 3.1 we get

$$k - 1 \geq (k - j)(r + 1) \geq (k - j)3.$$

This is in contradiction with the assumption $3j < 2k + 1$. □

Note that the condition (3.2) can be replaced equivalently by

$$\text{(3.4)} \qquad \deg(h_1) \leq 1 \quad \text{and} \quad \deg(h_k) \leq k \quad \text{for some integer } k \geq 2,$$

or also by

$$\text{(3.5)} \qquad \deg(h_2) \leq 2 \quad \text{and} \quad \deg(h_k) \leq k \quad \text{for some integer } k \geq 3,$$

(since in both cases the restriction $3j < 2k + 1$ is automatically verified). We mention that (3.4) with $k = 2$ corresponds to the characterization obtained by Diaconis and Zabell ((1991), Theorem 1, p. 296).

Lemma 3.1 being formulated for an arbitrary r, it can be applied to derive a sufficient condition for determining the value of the index r of the generalized Pearson curve.

PROPERTY 3.1. *Let f be of the form (1.3). If*

$$\text{(3.6)} \qquad \max\left(\frac{\deg(h_j)}{jw}, \frac{\deg(h_k)}{kw}\right) = 1 \textit{ for some integers } w \geq 1 \textit{ and } 1 \leq j < k \textit{ such that } (w + 2)j < (w + 1)k + 1,$$

then $r = w$.

PROOF. From Theorem 2.1 and (3.6), we can write that $jr \geq \deg(h_j) = jw$ or/and $kr \geq \deg(h_k) = kw$, so that $r \geq w$ necessarily. Let us show that $r > w$ is impossible. If $r \geq w + 1$, we get from (3.6) that $\deg(h_j) \leq jw < jr$ and $\deg(h_k) \leq kw < kr$. Thus we have $n \geq 2$, and Lemma 3.1 yields

$$k - 1 \geq (k - j)(r + 1) \geq (k - j)(w + 2),$$

which is in contradiction with the assumption $(w + 2)j < (w + 1)k + 1$. □

As a special case of the condition (3.6), we have, for instance,

$$\max\left(\frac{\deg(h_1)}{w}, \frac{\deg(h_k)}{kw}\right) = 1 \text{ for some integers } w \geq 1 \text{ and } j > 1.$$

4. Covariance identity, Orthogonality, Characterizations

Throughout this section, we assume that f is a probability density function which is strictly positive on its support (a, b). Let X be a continuous random variable with density function f.

4.1 *Covariance identity*

We are going to derive some variant of the covariance identity (1.4). For that, we will impose the condition

$$\lim_{x \to a} x^i p(x) f(x) = 0 = \lim_{x \to b} x^i p(x) f(x) \text{ for all integers } i \geq 0. \tag{4.1}$$

Note that under (1.3), (4.1) implies that $r \geq 1$. Indeed, by Theorem 2.1, $\deg(h_1) \leq r$, and by (4.1), $\int_a^b f(x) h_1(x) dx = 0$ so that $\deg(h_1) \geq 1$.

THEOREM 4.1. *Let f be of the form* (1.3) *and satisfy* (4.1). *Then, for any infinitely differentiable function g such that its derivatives $D^k(g)$, $k \geq 0$, do not grow faster than polynomials,*

$$(-1)^j E[h_j(X) g(X)] = E[p^j(X) D^j g(X)], \qquad j \geq 1. \tag{4.2}$$

PROOF. Integrating by parts we have

$$\begin{aligned} \int_a^b h_j(x) g(x) f(x) dx &= \int_a^b g D^j(fp^j) dx = \int_a^b g dD^{j-1}(fp^j) \\ &= \left[\lim_{x \to b} g D^{j-1}(fp^j) - \lim_{x \to a} g D^{j-1}(fp^j)\right] \\ &\quad - \int_a^b D(g) D^{j-1}(fp^j) dx. \end{aligned} \tag{4.3}$$

Arguing as for (2.6), we find that for all $k < j$, $(1/f) D^k(fp^j) = pu$ where u is a polynomial. Thus, the condition (4.1) and the assumption made on g imply that the term $[\ldots]$ in (4.3) reduces to 0, giving

$$\int_a^b g D^j(fp^j) dx = -\int_a^b D(g) D^{j-1}(fp^j) dx.$$

By iteration, we then deduce that

$$\int_a^b gh_j f dx = (-1)^j \int_a^b D^j(g) f p^j dx,$$

which is (4.2). □

Remark. Taking $g(x) = x^j$ in (4.2), multiplying both sides of this identity by $z^j/j!^2, z \in I\!R$, and summing up, we get the formal series expansion

$$E \exp[zp(X)] = \sum_{j=0}^{\infty} \frac{z^j}{j!^2} E[h_j(X)(-X)^j], \tag{4.4}$$

a result which seems to be new. For instance, let us choose $p(x) \equiv 1$. Then, $\{h_j = (1/f)D^j(f), j \geq 0\}$, and

$$\exp(z) = \sum_{j=0}^{\infty} \frac{z^j}{j!^2} E[h_j(X)(-X)^j],$$

which includes a remarkable identity for a standard normal distribution when $\{h_j(x)(-\sqrt{2})^j, j \geq 0\}$ are Hermite polynomials. Similarly, taking $p(x) \equiv x$ (resp. $x(1-x)$) yields a series expansion (4.4) which includes a remarkable identity for a gamma distribution when $\{h_j(x)/j!, j \geq 0\}$ are the associated Laguerre polynomials (resp. for a beta distribution when $\{h_j(x)(-1)^j/j!, j \geq 0\}$ are the associated Jacobi polynomials).

4.2 *Orthogonality*

As a direct corollary, we now show that the polynomials h_j are, after some indice lag, orthogonal with respect to such a density function f.

COROLLARY 4.1. *Let f be of the form* (1.3) *and satisfy* (4.1). *Then, h_j and h_k are orthogonal whenever $j > kr$, that is*

$$E[h_j(X)h_k(X)] = 0 \quad \textit{for all} \quad j > kr. \tag{4.5}$$

PROOF. Writing (4.2) for $g(x) = x^i, 1 \leq i \leq j-1$, we see that $h_j \perp \mathcal{L}\{1, x, \ldots, x^{j-1}\}$. By Theorem 2.1, $\deg(h_k) \leq kr$, so that $h_j \perp h_k$ if $j - 1 \geq kr$. □

Thus, for $r \geq 2$, we can easily extract an orthogonal subsystem of polynomials from the system of polynomials $\{h_j, j \geq 1\}$.

PROPERTY 4.1. *Given a subsequence of indices $j(r) = 1 + r + \cdots + r^{j-1}$, $j \geq 1$, the polynomials $\{k_j \equiv h_{j(r)}, j \geq 1\}$ are orthogonal.*

Now, let us consider the case $r = 1$ (Pearson family). Taking $g = h_j$ in (4.2)—if it is valid—yields

$$(-1)^j E[h_j^2(X)] = j! a(h_j) E[p^j(X)]. \tag{4.6}$$

This means that either $a(h_j) \neq 0$ (i.e. $\deg(h_j) = j$), or $h_j \equiv 0$.

Note that in general, $h_j \equiv 0$ may happen. For instance, define $f \equiv c/p^j$, c being a normalizing factor. When $j = 1$ and $\deg(p) = 2$, f belongs to the Pearson family. Obviously, $D^j(fp^j) \equiv 0$, i.e. $h_j \equiv 0$. It is shown below, however, that this is not possible under the condition (4.1). The result, not new, is recalled for reasons of clarity.

PROPERTY 4.2. *Let f be a Pearson family satisfying* (4.1). *Then, the Beale condition holds for all $j \geq 1$, so that $\{h_j, j \geq 0\}$ is a complete orthogonal system.*

PROOF. Let us assume to the contrary that the Beale condition is not true for all $j \geq 1$. By Lemma 2.4 and (2.7), we then know that for some $j \geq 1$, $a(h_j) = a(q_{j,j}) = 0$. Thus, (4.6) gives $h_j \equiv 0$, i.e. $D^j(fp^j) \equiv 0$, so that

$$(4.7) \qquad f(x) = p_k(x)/p^j(x) \text{ for some polynomial } p_k \text{ of degree } k \leq j-1.$$

As a consequence, (4.1) implies that a and b are finite. Furthermore, let us put $t = j - k$. By Theorem 2.1, $h_{j+t} = (1/f)D^{j+t}(fp^{j+t})$ is a polynomial. We also observe that $fp^{j+t} = p^t p_k$ is a polynomial and, since $\deg(p) \leq 2$, of degree

$$\deg(p^t p_k) = k + 2t = k + 2(j-k) = 2j - k = j + t.$$

Thus, h_{j+t} can be expressed as $h_{j+t} = cp^j/p_k$, so that it is a polynomial if, and only if, the roots of p_k correspond precisely to the roots of p. Since p has at most two roots and condition (4.1) holds, we deduce from (4.7) that f is of the form

$$(4.8) \qquad f = c(x-a)^\alpha (b-x)^\beta \text{ for some integers } \alpha \text{ and } \beta.$$

Now, f being integrable, we have α and $\beta > -1$. Moreover, from (4.8) we see that $a(q) = \alpha + \beta$. Therefore, $a(q) > -2 \geq -j-1$, which means that the Beale condition is satisfied, hence the contradiction. The implication that $\{h_j, j \geq 0\}$ is a complete orthogonal system, then follows from Corollary 4.1. □

Remark. Without condition (4.1), a Pearson family can give rise to a finite subsystem of orthogonal polynomials. For instance, given any real $\varepsilon > 1$, define $f(x) = k(1+x^2)^{-\varepsilon}, x \in \mathbb{R}$ (a Student's-type distribution). Thus, f is a Pearson curve with $f'(x)/f(x) = -(2\varepsilon x)/(1+x^2)$. Obviously, (4.1) is not satisfied. Using integration by parts and Lemma 2.4, it can be shown that any polynomials h_j, h_k are orthogonal whenever $j \neq k < \varepsilon$.

4.3 *Characterizations*

We start with the following characterization of the Pearson family, which seems to be little known.

PROPERTY 4.3. *Let f be of the form* (1.3) *and satisfy* (4.1). *Then, $\{h_j, j \geq 0\}$ is a (complete) orthogonal system if, and only if, f belongs to the Pearson family.*

PROOF. By Property 4.2, it suffices to establish the necessity part. Let us suppose to the contrary that $r > 1$. By Lemma 3.1, we can then find some integer $k \geq 1$ such that $\deg h_{2k} = 2kr$. From (4.2) with $j = 2kr$ and $g = h_{2k}(x)$, we now find that

$$E[h_{2kr}(X)h_{2k}(X)]$$
$$= (-1)^{2kr}E[p^{2kr}(X)D^{2kr}h_{2k}(X)] = (-1)^{2kr}(2kr)!a(h_{2k})E[p^{2kr}(X)] \neq 0,$$

so that $h_{2kr} \not\perp h_{2k}$. □

Remark. It is well-known (see, e.g., Ord (1972)) that within the Pearson family, the only density functions which generate orthogonal polynomials through a Rodrigues' formula are:

(C.1) on the real line $\mathbb{R} = (-\infty, \infty)$: the normal densities $f(x) = k \exp[-(x-\mu)^2/2\sigma^2]$, which generate Hermite polynomials,

(C.2) on a half real line $[a, \infty)$: the gamma densities $f(x) = k(x-a)^\alpha \exp(-\theta x)$ with $\alpha > -1, \theta > 0$, which generate Laguerre polynomials,

(C.3) on a finite interval $[a, b]$: the beta densities $f(x) = k(x-a)^\alpha (b-x)^\beta$ with $\alpha, \beta > -1$, which generate Jacobi polynomials (and in particular Legendre (when $\alpha = \beta = 0$) and Gegenbauer (when $\alpha = \beta$) polynomials).

Note that all of these satisfy the condition (4.1).

This property leads us to a characterization of the normal, gamma and beta densities.

COROLLARY 4.2. *Let f be of the form* (1.3) *with* $\max(n-1, m) \geq 1$ *on the real line* $\mathbb{R}$. *Then,* $\{h_j, j \geq 0\}$ *is a (complete) orthogonal system if, and only if, f is a normal density.*

PROOF. The sufficiency part follows immediately from (C.1) above. By Theorem 3.1, we know that if $\{h_j, j \geq 0\}$ is a complete orthogonal system, then f is a Pearson curve, which implies that f is a normal density by the previous remark. Thus, it remains to show that if $\{h_j, j \geq 0\}$ is an orthogonal system, then it is complete. First, observe that by Lemma 3.1, h_j is of degree jr for all $j \geq j_0$, so that $\mathcal{LH}$ is infinite. Now, f being of the form (1.3), f' has at most a finite number of roots. As for $j \geq 1$, $\int_{-\infty}^{\infty} |x|^j f(x)dx < \infty$ (by the assumption of orthogonality), we then have

$$\lim_{x \to -\infty} x^j f(x) = 0 = \lim_{x \to \infty} x^j f(x), \qquad j \geq 0. \tag{4.9}$$

This is equivalent to the condition (4.1) at the point $b = \infty$, hence the result by Property 4.2. □

COROLLARY 4.3. *Let f be of the form* (1.3) *with* $\max(n-1, m) \geq 1$ *on a half real line. Then,* $\{h_j, j \geq 0\}$ *is a (complete) orthogonal system if, and only if, f is a gamma density.*

PROOF. An argument similar to the one in the proof of Corollary 4.2 above is applicable, so that it remains to show that f satisfies (4.1). Let us take, for instance, $a < \infty$ and $b = \infty$. As above, we see that the condition (4.9), and its equivalent (4.1), are satisfied at the point $b = \infty$. Since h_1 is orthogonal to $h_0 = 1$, we have

$$0 = \int_a^\infty h_1(x) f(x) dx = \int_a^\infty (fp)'(x) dx = -p(a) f(a),$$

and thus,

$$\lim_{x \to a} x^j p(x) f(x) = a^j p(a) f(a) = 0, \qquad j \geq 0,$$

as announced. □

COROLLARY 4.4. *Let f be of the form* (1.3) *on a finite interval and satisfy* (4.1). *Then,* $\{h_j, j \geq 0\}$ *is a (complete) orthogonal system if, and only if,* f *is a beta density.*

Remark. The proof is straightforward once (4.1) is assumed. It seems difficult, however, not to insert that condition in the case of a finite interval (a, b). We notice that (a, b) is usually chosen as the maximum interval where p does not have roots, thus implying that $p(a) = p(b) = 0$.

4.4 *On condition* (4.1)

A condition analogous to (4.1) is given in several papers related to the present work, especially through the concept of z-function in the sense of (1.5) (see, e.g., Papathanasiou (1995) and Cacoullos and Papathanasiou (1995) and references therein).

Let us assume that p corresponds to the z-function associated with some function h, that is

$$p(x)f(x) = \int_a^x [Eh(X) - h(t)]f(t)dt, \qquad t \in [a, b].$$

Then, pf satisfies

$$\lim_{x \to a} p(x)f(x) = 0 = \lim_{x \to b} p(x)f(x).$$

Clearly, this restriction to a finite interval $[a, b]$ is equivalent to (4.1).

Now, let us consider the special case where $h(x) = x$. As shown by Korwar (1991), the Pearson family of curves can be characterized on the basis of the associated $z := w$-function. Specifically, it can be seen that if w is a polynomial of degree at most 2, then f is a Pearson curve, and that the converse holds true under the additional condition (4.1).

Note that in general, the w-function associated with a Pearson curve family is not necessarily a polynomial. For instance, define $f(x) := 1/(x \ln 2), x \in [1, 2]$. Thus, f is a Pearson curve with $f'(x)/f(x) = -x/x^2$. We see that (4.1) is not satisfied while the Beale condition holds for all $j \geq 1$ (since $a(q) = -1 \neq i - 2j, 0 \leq i \leq j - 1$). Moreover, we obtain that $w(x) = x \ln x / \ln 2 - x(x - 1)$ and is not a polynomial.

Acknowledgements

We thank an anonymous referee for valuable remarks and suggestions. The research of V. Papathanasiou and S. Utev was partially supported by a grant of the Fonds National de la Recherche Scientifique Belge. It was achieved during their visit at the Institut de Statistique et de Recherche Opérationnelle of the Université Libre de Bruxelles.

Appendix

We point out here a special class of generalized Pearson curves.

Clearly, a generalized Pearson curve is *logconcave* if $q'p \leq qp'$. Moreover, it can be shown that a *symmetric* curve can be represented as $f'/f = q_o/p_e$, where $q_o(x) = \sum_i a_i x^{2i+1}$ and $p_e(x) = \sum_i a_i x^{2i}$.

Within the Pearson family, the only symmetric logconcave curves are the normal, uniform and symmetric beta densities. Symmetric logconcave densities are rather frequent in the generalized Pearson family. In particular, this is the case for f of the form $f(x) = \exp(-\sum_{i=0}^k a_i^2 x^{2i}) \equiv \exp[-l(x)]$; such a density is named quasi normal.

To close, it is worth deriving an upper bound for the variance of a quasi normal distribution. Following Borovkov and Utev (1983), let us consider, for any r.v. X, the functional $R_X = \sup_g\{\mathrm{Var}[g(X)]/E[g'(X)]^2\}$, the sup being taken over all absolutely continuous functions g with $0 < E[g'(X)]^2 < \infty$. It is known that $\mathrm{Var}(X) \leq R_X$, and equality holds if, and only if, X has a normal distribution. Moreover, the characterization is robust in the sense that an upper bound for $d_{TV}(X, N_{[EX,\mathrm{Var}(X)]})$, the total variation distance between X and a normal random variable $N_{[EX,\mathrm{Var}(X)]}$, is provided by $d_{TV}(X, N_{[EX,\mathrm{Var}(X)]}) \leq 3\sqrt{R_X/\mathrm{Var}(X) - 1}$ (Utev (1989) and Cacoullos *et al.* (1994)). Now, assume that X has a quasi normal density. To estimate R_X, we can apply the variational inequality for logconcave densities of Brascamp and Lieb (1976), giving

$$\mathrm{Var}[g(X)] \leq E\left\{\frac{[g'(X)]^2}{\{-\ln[f(X)]\}''}\right\} = E\left\{\frac{[g'(X)]^2}{\left[\sum_{i=1}^k \alpha_i^2 2i(2i-1)X^{2(i-1)}\right]}\right\} \leq \frac{E[g'(X)]^2}{2\alpha_1^2},$$

so that $R_X \leq 1/2\alpha_1^2$. We then deduce that

$$\mathrm{Var}(X) \leq 1/2\alpha_1^2, \qquad \text{and} \qquad d_{TV}(X, N_{[0,\mathrm{Var}(X)]}) \leq 3\sqrt{1/2\alpha_1^2\,\mathrm{Var}(X) - 1}.$$

Note that the bound for $\mathrm{Var}(X)$ and the related characterization can also be easily derived from (4.2) with $j = 1$, which yields a Stein-type identity $E[l'(X)g(X)] = E[g'(X)]$.

References

Beale, F. S. (1941). On a certain class of orthogonal polynomials, *Ann. Math. Statist.*, **12**, 97–103.

Borovkov, A. A. and Utev, S. A. (1983). On an inequality and a related characterization of the normal distribution, *Theory Probab. Appl.*, **28**, 209–218.

Brascamp, H. J. and Lieb, E. H. (1976). On extensions of the Brunn-Minkowski and Prekopa-Leindner theorems including inequalities for the log concave functions and with applications to the diffusion equation, *J. Funct. Anal.*, **22**, 366–389.

Cacoullos, T. and Papathanasiou, V. (1989). Characterizations of distributions by variance bounds, *Statist. Probab. Lett.*, **7**, 351–356.

Cacoullos, T. and Papathanasiou, V. (1995). A generalization of covariance identity and related characterizations, *Math. Methods Statist.*, **4**, 106–113.

Cacoullos, T., Papathanasiou, V. and Utev, S. A. (1994). Variational inequalities with examples and an application to the central limit theorem, *Ann. Probab.*, **22**, 1607–1618.

Diaconis, P. and Zabell, S. (1991). Closed form summation for classical distributions: Variations on a theme of De Moivre, *Statist. Sci.*, **6**, 284–302.

Dunning, K. A. and Hanson, J. N. (1977). Generalized Pearson distributions and nonlinear programming, *J. Statist. Comput. Simulation*, **6**, 115–128.

Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994). *Continuous Univariate Distributions*, Vol. 1, 2nd ed., Wiley, New York.

Korwar, R. M. (1991). On characterizations of distributions by mean absolute deviation and variance bounds, *Ann. Inst. Statist. Math.*, **43**, 287–295.

Ord, J. K. (1972). *Families of Frequency Distributions*, Charles Griffin, London.

Papathanasiou, V. (1995). A characterization of the Pearson system of distributions and the associated orthogonal polynomials, *Ann. Inst. Statist. Math.*, **47**, 171–176.

Utev, S. A. (1989). Probability problems connected with a certain integrodifferential inequality, *Siberian Math. J.*, **30**, 490–493.