APPLICATIONES MATHEMATICAE
42,2-3 (2015), pp. 231–267

G. Afendras (Buffalo, NY)
N. Papadatos (Athens)

# INTEGRATED PEARSON FAMILY AND ORTHOGONALITY OF THE RODRIGUES POLYNOMIALS: A REVIEW INCLUDING NEW RESULTS AND AN ALTERNATIVE CLASSIFICATION OF THE PEARSON SYSTEM

*Dedicated to Professor V. Papathanasiou*

*Abstract.* An alternative classification of the Pearson family of probability densities is related to the orthogonality of the corresponding Rodrigues polynomials. This leads to a subset of the ordinary Pearson system, the so-called *Integrated Pearson Family*. Basic properties of this family are discussed and reviewed, and some new results are presented. A detailed comparison between the Integrated Pearson Family and the ordinary Pearson system is presented, including an algorithm that enables one to decide whether a given Pearson density belongs, or not, to the integrated system. Recurrences between the derivatives of the corresponding orthonormal polynomials are also given.

**1. Introduction.** Karl Pearson (1895) introduced his famous family of frequency curves by means of the differential equation

$$\frac{f'(x)}{f(x)} = \frac{p_1(x)}{p_2(x)},$$

where $f$ is the probability density and $p_i$ is a polynomial in $x$ of degree at most $i$, $i = 1, 2$. Since then, a vast bibliography has been developed regarding the properties of the Pearson family distributions. The original classification given by Pearson contains twelve types (I–XII), although this

2010 *Mathematics Subject Classification*: Primary 62E15, 60E05; Secondary 62-00.
*Key words and phrases*: integrated Pearson family of distributions, derivatives of orthogonal polynomials, Rodrigues polynomials.

DOI: 10.4064/am42-2-9  

numbering system does not have a clear systematic basis (Johnson et al., 1994, p. 16). Craig (1936) proposed a new exposition and chart for Pearson curves. However, a more reasonable and convenient classification is included in a review paper by Diaconis and Zabell (1991). Extensions to discrete distributions have been introduced by Ord (1967) and an extensive review can be found in Ord (1972, Chapter 1).

In this paper we present and review a number of properties satisfied by the distributions of the Pearson family and the associated Rodrigues polynomials, which are produced by a Rodrigues-type formula. Our main focus is on a suitable subset of Pearson distributions, the *Integrated Pearson Family*, because this class subsumes all interesting properties related to the associated orthogonal polynomial systems. For example, it will be shown in Section 4 that orthogonality of Rodrigues polynomials with respect to an ordinary Pearson density $f$ results in an equivalent definition of the integrated Pearson system. This consideration entails an alternative classification of (integrated) Pearson distributions, which is essentially the one given in Diaconis and Zabell (1991).

In the context of deriving variance bounds for functions of random variables, Afendras et al. (2007, 2011) and Afendras and Papadatos (2011) have made use of the following definition, which provides the main framework of the present article.

Definition 1.1 (Integrated Pearson Family). Let $X$ be an absolutely continuous random variable with density $f$ and finite mean $\mu = \mathbb{E}\,X$. We say that $X$ (or its density) belongs to the *Integrated Pearson Family* of distributions (or *integrated Pearson system*) if there exists a quadratic polynomial $q(x) = \delta x^2 + \beta x + \gamma$ (with $\delta, \beta, \gamma \in \mathbb{R}$, $|\delta| + |\beta| + |\gamma| > 0$) such that

$$\int_{-\infty}^{x} (\mu - t) f(t)\, dt = q(x) f(x) \quad \text{for all } x \in \mathbb{R}. \tag{1.1}$$

This fact will be denoted by $X \sim \mathrm{IP}(\mu; q)$ or $f \sim \mathrm{IP}(\mu; q)$ or, more explicitly, by $X$ or $f \sim \mathrm{IP}(\mu; \delta, \beta, \gamma)$.

Despite the fact that the Integrated Pearson Family is quite restricted, compared to the usual Pearson system (see Proposition 2.1(iii) below), we believe that the reader will find here some interesting observations worth highlighting. The integrated Pearson system has many interesting properties, like recurrences on moments and on Rodrigues polynomials, covariance identities, closedness of each type under particularly useful transformations etc. Such properties are by far more complicated (if at all true) for distributions outside the Integrated Pearson Family of distributions. These features should be combined with the fact that the Rodrigues polynomials form an orthogonal system for the corresponding Pearson density if and only if the density belongs to the Integrated Pearson Family. Consequently, the Rodrigues polynomials are useful only if they are considered in the framework of the integrated Pearson system. To our knowledge, these facts have not been written explicitly elsewhere.

The paper is organized as follows: In Section 2 we provide a detailed classification of the Integrated Pearson Family. It turns out that, up to an affine transformation, there are six different types of densities, included in Table 2.1. We also provide conditions guaranteeing the existence of moments, and we give recurrences as long as these moments exist. In Section 3, a detailed comparison between the Integrated Pearson Family and the ordinary Pearson system is presented. Interestingly, there exists a simple algorithm that enables one to decide whether a given ordinary Pearson density belongs to the integrated system or not. In Section 4, exploiting a result of Diaconis and Zabell (1991), we show that (under natural moment conditions) the first three Rodrigues polynomials (of degree 0, 1 and 2) are orthogonal with respect to an ordinary Pearson density if and only if this density belongs to the integrated Pearson system. Finally, in Section 5 we provide recurrences between the orthonormal polynomials and their derivatives; in fact, the derivatives themselves are orthogonal polynomials with respect to other integrating Pearson densities, having the same quadratic polynomial, up to a scalar multiple. Although we do not include any specific applications of these results here, we notice that such recurrences are particularly useful in obtaining Fourier expansions of the derivatives of a function of a Pearson variate. The main result of Section 5 is Corollary 5.4. It provides an explicit relation (in terms of $\mu$ and $q$) between the $m$th derivative of an orthonormal polynomial of degree $k \geq m$ and the corresponding orthonormal polynomial of degree $k-m$. That is, it relates the orthonormal polynomial system, associated with some $f \sim \mathrm{IP}(\mu; q)$, to the corresponding orthonormal polynomial system associated with the 'target' density $f_m \propto q^m f$.

Throughout, $X \sim \mathrm{IP}(\mu; \delta, \beta, \gamma)$ means that $X$ has finite mean $\mu$, and that $X$ admits a density $f$ (with respect to Lebesgue measure on $\mathbb{R}$) such that (1.1) is fulfilled. Define the open (bounded or unbounded) interval

$$J = J(X) := (\text{ess inf}(X), \text{ess sup}(X)).$$

If $F$ is the distribution function of $X$ then $J = (\alpha_F, \omega_F) = (\alpha, \omega)$, say, where $\alpha_F := \inf\{x : F(x) > 0\}$, $\omega_F := \sup\{x : F(x) < 1\}$. It is clear that (1.1) takes the form $0 = 0$ whenever $x = \rho$ is a zero of $q$ that lies outside the interval $(\alpha, \omega)$; thus, $f(\rho)$ may assume any value in this case. However, in order to be specific, we redefine $f(\rho) = 0$ at such points without any loss of generality. Therefore, we shall use this convention throughout without any further reference.

**2. A complete classification of the Integrated Pearson Family.** We show in this section that the Integrated Pearson Family contains six different types of distributions. These are classified in terms of the corresponding quadratic polynomial $q(x) = \delta x^2 + \beta x + \gamma$ and its discriminant $\Delta = \beta^2 - 4\delta\gamma$ as follows:

- type 1 (Normal-type, $\delta = \beta = 0$);
- type 2 (Gamma-type, $\delta = 0$, $\beta \neq 0$);
- type 3 (Beta-type, $\delta < 0$);
- type 4 (Student-type, $\delta > 0$, $\Delta < 0$);
- type 5 (Reciprocal Gamma-type, $\delta > 0$, $\Delta = 0$);
- type 6 (Snedecor-type, $\delta > 0$, $\Delta > 0$).

The first three types (with $\delta \le 0$) consist of the well-known Normal, Gamma and Beta random variables and their linear transformations. The last three types (with $\delta > 0$) consist of some less familiar distributions (see Table 2.1 below); they have finite moments of order $a$ for any $a \in [0, 1 + 1/\delta)$ while $\mathbb{E}\,|X|^{1+1/\delta} = \infty$. The proposed classification is very similar to the one given by Diaconis and Zabell (1991, Table 2 and pp. 294–296).

We start with an easily verified proposition.

Proposition 2.1. *Let* $X \sim \mathrm{IP}(\mu; q)$ *and set* $J = (\alpha, \omega) = (\mathrm{ess\ inf}(X), \mathrm{ess\ sup}(X))$. *Then:*

(i) $f(x)$ *is strictly positive for* $x$ *in* $J$ *and zero otherwise, i.e.,* $\{x : f(x) > 0\} = J$;
(ii) $f \in C^\infty(J)$, *that is,* $f$ *has derivatives of any order in* $J$;
(iii) $X$ *is a (usual) Pearson random variable supported in* $J$;
(iv) $q(x) = \delta x^2 + \beta x + \gamma > 0$ *for all* $x \in J$;
(v) *if* $\alpha > -\infty$ *then* $q(\alpha) = 0$ *and, similarly, if* $\omega < \infty$ *then* $q(\omega) = 0$;
(vi) *for any* $\theta, c \in \mathbb{R}$ *with* $\theta \neq 0$, *the random variable* $\widetilde{X} := \theta X + c \sim \mathrm{IP}(\widetilde{\mu}; \widetilde{q})$ *with* $\widetilde{\mu} = \theta\mu + c$ *and* $\widetilde{q}(x) = \theta^2 q((x - c)/\theta)$.

*Proof.* By (1.1), $x \mapsto q(x)f(x)$ is continuous. On the other hand, from the definition of $J = (\alpha_F, \omega_F) = (\alpha, \omega)$ it follows that $q(x)f(x)$ must vanish for all $x \le \alpha$ (if any) and for all $x \ge \omega$ (if any). Also, it must be strictly positive for $x \in J$. Indeed, if $x \in (\mu, \omega)$ then

$$q(x)f(x) = \int_x^\infty (t - \mu)f(t)\,dt \ge (x - \mu)(1 - F(x)) > 0;$$

if $x \in (\alpha, \mu)$ then

$$q(x)f(x) = \int_{-\infty}^x (\mu - t)f(t)\,dt \ge (\mu - x)F(x) > 0;$$

finally, $q(\mu)f(\mu) = \frac{1}{2}\,\mathbb{E}\,|X-\mu| > 0$. Thus, $q(x)f(x) > 0$ for all $x \in (\alpha,\omega)$. Since $q$ is continuous and has no roots in $J$ it follows that both $q(x)$ and $f(x)$ are strictly positive (and continuous) in $J$. The vanishing of $qf$ outside $J$ shows that $f(x) = 0$ for all $x \notin J$, with the possible exception at the points $x \notin J$ which are real roots of $q$. Clearly, if $\rho \in \mathbb{R} \smallsetminus (\alpha,\omega)$ is a zero of $q$ we can redefine $f(\rho) = 0$, if necessary, so that (i) and (iv) follow.

On the other hand, the function $f : (\alpha,\omega) \to (0,\infty)$ has derivatives of any order. Indeed, writing $p_1(x) = \mu - x - q'(x)$ (which is a polynomial of degree at most one) we see from (1.1) that $f : J \to (0,\infty)$ is continuous. Thus,

$$f'(x) = f(x)\frac{p_1(x)}{q(x)} \quad \text{or equivalently} \quad \frac{f'(x)}{f(x)} = \frac{\mu - x - q'(x)}{q(x)}, \quad x \in J. \tag{2.1}$$

This proves (iii).

Moreover, (2.1) shows that $f'$ is continuous in $J$ and, inductively, that $f^{(n+1)} : J \to \mathbb{R}$ is continuous, since for $x \in J$,

$$f^{(n+1)}(x) = \sum_{j=0}^{n} \binom{n}{j} f^{(j)}(x) \left(\frac{p_1(x)}{q(x)}\right)^{(n-j)}, \qquad n = 0,1,2,\ldots.$$

Now (vi) is straightforward and it remains to show (v). To this end, assume that $\omega < \infty$. Since $q(\omega) = \lim_{x\nearrow\omega} q(x)$ and $q(x) > 0$ for $x$ in a lower neighborhood of $\omega$, it follows that $q(\omega) \ge 0$. Assume now that $q(\omega) > 0$ and define $\lambda_1 := \inf_{x\in[\mu,\omega]} q(x) > 0$ and $\lambda_2 := \sup_{x\in[\mu,\omega]} |\mu - x - q'(x)| < \infty$. Then, for all $x \in [\mu,\omega)$,

$$\begin{aligned}\left|\int_\mu^x \frac{\mu - t - q'(t)}{q(t)}\,dt\right| &\le \int_\mu^x \frac{|\mu - t - q'(t)|}{q(t)}\,dt \\ &\le \int_\mu^\omega \frac{|\mu - t - q'(t)|}{q(t)}\,dt \le (\omega-\mu)\frac{\lambda_2}{\lambda_1} < \infty.\end{aligned}$$

Setting $\lambda := (\omega-\mu)\lambda_2/\lambda_1 < \infty$ and observing that

$$\ln f(x) = \ln f(\mu) + \int_\mu^x \frac{f'(t)}{f(t)}\,dt = \ln f(\mu) + \int_\mu^x \frac{\mu - t - q'(t)}{q(t)}\,dt, \qquad x \in [\mu,\omega),$$

we have

$$|\ln f(x)| \le |\ln f(\mu)| + \lambda := c < \infty, \quad \mu \le x < \omega.$$

Therefore, there exist constants $c_1, c_2$ such that $0 < c_1 \le f(x) \le c_2 < \infty$ for all $x \in [\mu,\omega)$. Thus,

$$q(\omega) = \lim_{x\nearrow\omega} q(x) = \lim_{x\nearrow\omega} \frac{1}{f(x)} \int_x^\omega (t-\mu)f(t)\,dt = 0,$$

which contradicts the assumption $q(\omega) > 0$. The case $\alpha > -\infty$ is reduced to the case $\omega < \infty$ if we consider the random variable $\widetilde{X} = -X$ with $J(\widetilde{X}) = (\widetilde{\alpha}, \widetilde{\omega}) = (-\omega, -\alpha)$. According to (vi), its density, $\widetilde{f}(x) = f(-x)$, satisfies (1.1) with $\widetilde{\mu} = -\mu$, $\widetilde{\alpha} = -\omega$, $\widetilde{\omega} = -\alpha$ and $\widetilde{q}(x) = q(-x)$. Thus, if $\alpha > -\infty$ then $\widetilde{\omega} < \infty$ and $q(\alpha) = \widetilde{q}(-\alpha) = \widetilde{q}(\widetilde{\omega}) = 0$. ∎

Corollary 2.1. *Let $X \sim \mathrm{IP}(\mu; q)$ and assume that $\alpha = \operatorname{ess\,inf}(X)$ and $\omega = \operatorname{ess\,sup}(X)$ are the lower and upper endpoints of the distribution function of $X$. Then the support of $X$ (or of its density $f$) $S(f) = S(X) := \{x : f(x) > 0\}$ equals the open interval $J = J(X) = (\alpha, \omega)$. This interval support has the following two properties:*

(i) *$J \subseteq S^{+}(q) := \{x : q(x) > 0\}$ and*
(ii) *$J$ is a maximal open interval contained in $S^{+}(q)$, i.e., for any open interval $\widetilde{J} \subseteq S^{+}(q)$, either $\widetilde{J} \subseteq J$ or $\widetilde{J} \cap J = \emptyset$.*

Thus, the support $J$ of $X$ is a connected component of the open set $\{x : q(x) > 0\}$. Since $q$ is a polynomial of degree at most two, it is clear that the set $\{x : q(x) > 0\}$ has at most two such components. For example, if $q(x) = x^2$ then either $J = (-\infty, 0)$ or $J = (0, \infty)$; if $q(x) = x^2 - 1$ then either $J = (-\infty, -1)$ or $J = (1, \infty)$; if $q(x) = 1 - x^2$ then $J = (-1, 1)$; if $q(x) = x$ then $J = (0, \infty)$; if $q(x) = 1 + x^2$ or $q(x) \equiv 1$ then $J = \mathbb{R}$. Since $\mathbb{E}\, X = \mu \in J$, any particular choice of $\mu \in \{x : q(x) > 0\}$ characterizes the support $J$ of $X$.

We say that $q(x) = \delta x^2 + \beta x + \gamma$ is *admissible* if there exists $\mu \in \mathbb{R}$ such that $\mu \in \{x : q(x) > 0\}$. We shall show that for any admissible choice of $q$ and any $\mu \in \{x : q(x) > 0\}$ there exists an absolutely continuous random variable $X$ with density $f$ such that $\mathbb{E}\, X = \mu$ and (1.1) is fulfilled. Moreover, it will become clear that $f$ is characterized by the pair $(\mu; q)$. Therefore, the notation $X \sim \mathrm{IP}(\mu; q)$, or equivalently $f \sim \mathrm{IP}(\mu; q)$, has a well-defined meaning.

The proposed classification distinguishes between the cases $\delta = 0$, $\delta < 0$ and $\delta > 0$, as follows:

**2.1. The case $\delta = 0$.** Here we consider the two subcases $\beta = 0$ and $\beta \neq 0$.

**2.1.1.** *The subcase $\delta = 0$, $\beta = 0$.* Since $q(x) \equiv \gamma$ and $q$ is admissible, we must have $\gamma > 0$. Therefore, $J(X) = \mathbb{R}$. Fixing $\mu \in \mathbb{R}$ and solving the differential equation (2.1) we obtain

$$f(x) = \frac{1}{\sqrt{2\pi\gamma}} e^{-\frac{(x-\mu)^2}{2\gamma}}, \quad x \in \mathbb{R},$$

i.e., $X \sim N(\mu, \sigma^2)$ with $\sigma^2 = \gamma$.

**2.1.2.** *The subcase $\delta = 0$, $\beta \neq 0$.* Assume that $q(x) = \beta x + \gamma$ with $\beta \neq 0$ and fix $\mu \in \{x : q(x) > 0\}$. According to Proposition 2.1(vi) we may further assume that $\beta > 0$, $\gamma = 0$ and $\mu > 0$. To se this, it suffices to consider the random variable $\widetilde{X} = \frac{\beta}{|\beta|}(X + \frac{\gamma}{\beta})$ which has $\widetilde{q}(x) = |\beta| x$ and

$$\mathbb{E}\,\widetilde{X} = \widetilde{\mu} = \frac{\beta}{|\beta|}\left(\mu + \frac{\gamma}{\beta}\right) = \frac{q(\mu)}{|\beta|} > 0$$

(since $q(\mu) > 0$). Now, since $q(x) = \beta x$ with $\beta > 0$ we have $J(X) = (0, \infty)$. Fixing $\mu > 0$ and solving the differential equation (2.1) we obtain

$$f(x) = \frac{(1/\beta)^{\mu/\beta}}{\Gamma(\mu/\beta)} x^{\mu/\beta - 1} e^{-x/\beta}, \qquad x > 0,$$

i.e., $X \sim \Gamma(a, \lambda)$ with $a = \mu/\beta > 0$ and $\lambda = 1/\beta > 0$. Hence, a linear nonconstant $q$ corresponds to a *Gamma-type* distribution, i.e., to a linear transformation $\widetilde{X} = \theta X + c$, $\theta \neq 0$, of a Gamma random variable $X$.

**2.2. The case $\boldsymbol{\delta < 0}$.** Since $\delta < 0$ and $\{x : q(x) > 0\}$ must contain an interval, it follows that the discriminant $\beta^2 - 4\delta\gamma$ of $q$ is strictly positive. If $\rho_1 < \rho_2$ are the real roots of $q$ we write $q(x) = \delta(x - \rho_1)(x - \rho_2)$ so that the support of $X$ is the finite interval $J(X) = (\rho_1, \rho_2)$.

Now we show that for any choice of $\mu \in (\rho_1, \rho_2)$ there is a (unique) random variable $X$ with $X \sim \mathrm{IP}(\mu; q)$. To this end, it suffices to examine the particular case $q(x) = -\delta x(1 - x)$ and $0 < \mu < 1$. Indeed, the general case reduces to the particular one if we consider the random variable $\widetilde{X} = (X - \rho_1)/(\rho_2 - \rho_1)$. Fixing $\mu \in (0, 1)$, $q(x) = -\delta x(1 - x)$ and solving the differential equation (2.1) on $J(X) = (0, 1)$ we obtain

$$f(x) = \frac{1}{B(-\mu/\delta, -(1-\mu)/\delta)} x^{-\mu/\delta - 1}(1 - x)^{-(1-\mu)/\delta - 1}, \qquad 0 < x < 1,$$

i.e., $X \sim B(a, b)$ with $a = \mu/|\delta| > 0$, $b = (1 - \mu)/|\delta| > 0$. It follows that the case $\delta < 0$ corresponds to a *Beta-type* distribution, i.e., a linear transformation of a Beta random variable.

**2.3. The case $\boldsymbol{\delta > 0}$.** Here we consider the subcases where the discriminant $\Delta = \beta^2 - 4\delta\gamma$ is positive, zero or negative.

**2.3.1.** *The subcase $\delta > 0$, $\Delta < 0$.* Since $q$ has no real roots, $J(X) = \mathbb{R}$. Thus, $\mu \in \mathbb{R}$ can take an arbitrary value. Also, $q$ has the form $q(x) = \delta(x - c)^2 + \theta$ with $\delta > 0$, $\theta > 0$ and $c \in \mathbb{R}$. Without loss of generality we further assume that $c = 0$. Indeed, the general case reduces to the particular one if we consider the random variable $\widetilde{X} = X - c$. Fixing $\mu \in \mathbb{R}$, $q(x) = \delta x^2 + \theta$, (2.1) yields

$$f(x) = \frac{C}{(\delta x^2 + \theta)^{1 + 1/(2\delta)}} \exp\left(\frac{\mu}{\sqrt{\delta\theta}} \tan^{-1}(x\sqrt{\delta/\theta})\right), \qquad x \in \mathbb{R}.$$

The normalizing constant $C = C_\mu(\delta,\theta)$ can be calculated explicitly:

$$C_0(\delta,\theta) = \frac{\Gamma(1+1/(2\delta))\sqrt{\delta\theta^{1+1/\delta}}}{\Gamma(1/2+1/(2\delta))\sqrt{\pi}},$$

$$C_\mu(\delta,\theta) = \frac{2^{1/\delta}\theta^{1/2+1/(2\delta)}\sqrt{\delta}\left|\Gamma\left(1+\frac{1}{2\delta}+i\frac{\mu}{2\sqrt{\delta\theta}}\right)\right|^2}{\pi\Gamma(1+1/\delta)}.$$

The calculation is easy for $\mu = 0$ but not for $\mu \neq 0$, due to the complex argument of the Gamma function (cf. Nagahara, 1999; Nielsen, 2005).

Hence, provided $\mu = c$, the quadratic polynomial $q(x) = \delta(x-c)^2 + \theta$ (with $\delta > 0$ and $\theta > 0$) corresponds to a *Student-type* distribution centered at $c$. When $\mu \neq c$, it corresponds to some *asymmetric Student-type* distribution.

**2.3.2.** *The subcase $\delta > 0$, $\Delta = 0$.* Since $q$ has a unique real root at $\rho = -\beta/(2\delta)$, it follows that $q(x) = \delta(x-\rho)^2$, and therefore the support $J(X)$ is either $(-\infty,\rho)$ or $(\rho,\infty)$, according to whether $\mu < \rho$ or $\mu > \rho$. Without loss of generality we may assume that $q(x) = \delta x^2$ with $\delta > 0$ and $\mu > 0$. To see this, it suffices to consider the random variable $\widetilde{X} = \frac{\mu-\rho}{|\mu-\rho|}(X-\rho)$. Now, setting $J(X) = (0,\infty)$, $q(x) = \delta x^2$ $(\delta > 0)$ and $\mu > 0$ in (2.1) we find

$$f(x) = \frac{\lambda^a}{\Gamma(a)} x^{-a-1} e^{-\lambda/x}, \qquad x > 0,$$

where $\lambda = \mu/\delta > 0$ and $a = 1 + 1/\delta > 1$. Observing that $1/X \sim \Gamma(a,\lambda)$ it follows that the case $\delta > 0$, $\Delta = 0$ corresponds to a *Reciprocal Gamma-type* distribution.

**2.3.3.** *The subcase $\delta > 0$, $\Delta > 0$.* Assuming that $\rho_1 < \rho_2$ are the roots of $q$ we can write $q(x) = \delta(x-\rho_1)(x-\rho_2)$ and the support $J(X)$ has to be either $(-\infty,\rho_1)$ or $(\rho_2,\infty)$, according to whether $\mu < \rho_1$ or $\mu > \rho_2$. By considering the random variable $\widetilde{X} = -(X-\rho_1)$ when $\mu < \rho_1$ and $\widetilde{X} = X - \rho_2$ when $\mu > \rho_2$ it is easily seen that both cases reduce to $\widetilde{\mu} > 0$, $J(\widetilde{X}) = (0,\infty)$ and $\widetilde{q}(x) = \delta x(x+\theta)$ with $\delta, \theta = \rho_2 - \rho_1 > 0$. Thus, there is no loss of generality in assuming $\mu > 0$, $J(X) = (0,\infty)$ and $q(x) = \delta x(x+\theta)$ with $\delta,\theta > 0$. Then (2.1) yields

$$f(x) = \frac{1}{B(a,b)} \theta^a x^{b-1} (x+\theta)^{-a-b}, \qquad x > 0,$$

with $a = 1 + 1/\delta > 1$ and $b = \mu/(\delta\theta) > 0$. Equivalently, $\theta/(X+\theta) \sim B(a,b)$. It follows that the case $\delta > 0$, $\Delta > 0$ corresponds to a *Snedecor-type* distribution.

All the above possibilities are summarized in Table 2.1 opposite; compare with Diaconis and Zabell (1991, Table 2, p. 296).

**Table 2.1.** Densities of the Integrated Pearson Family $\mathrm{IP}(\mu;\delta,\beta,\gamma) \equiv \mathrm{IP}(\mu;q)$ (*)

| Type and usual notation | Density $f(x)$ | Support | $q(x)$ | Parameters | Mean $\mu$ | Classification rule |
|---|---|---|---|---|---|---|
| **1.** Normal-type $X \sim N(\mu,\sigma^2)$ | $\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ | $\mathbb{R}$ | $\sigma^2$ | $\gamma = \sigma^2 > 0$ | $\mu \in \mathbb{R}$ | $\delta = \beta = 0$ |
| **2.** Gamma-type $X \sim \Gamma(a,\lambda)$ | $\frac{\lambda^a}{\Gamma(a)} x^{a-1} e^{-\lambda x}$ | $(0,\infty)$ | $\frac{x}{\lambda}$ | $a,\ \lambda > 0$ | $\frac{a}{\lambda} > 0$ | $\delta = 0,\ \beta \neq 0$ |
| **3.** Beta-type $X \sim B(a,b)$ | $\frac{x^{a-1}(1-x)^{b-1}}{B(a,b)}$ | $(0,1)$ | $\frac{x(1-x)}{a+b}$ | $a,\ b > 0$ | $\frac{a}{a+b} > 0$ | $\delta = \frac{-1}{a+b} < 0$ |
| **4.** Student-type (1) | $\frac{C \exp\left(\frac{\mu \tan^{-1}(x\sqrt{\delta/\gamma})}{\sqrt{\delta\gamma}}\right)}{(\delta x^2+\gamma)^{1+\frac{1}{2\delta}}}$ (2) | $\mathbb{R}$ | $\delta x^2 + \gamma$ | $\delta,\ \gamma > 0$ | $\mu \in \mathbb{R}$ | $\delta > 0,\ \beta^2 < 4\delta\gamma$ |
| **5.** Reciprocal Gamma-type | $\frac{\lambda^a}{\Gamma(a)} x^{-a-1} e^{-\frac{\lambda}{x}}$ | $(0,\infty)$ | $\frac{x^2}{a-1}$ | $a > 1,\ \lambda > 0$ | $\frac{\lambda}{a-1} > 0$ | $\delta = \frac{1}{a-1} > 0,\ \beta^2 = 4\delta\gamma$ <br> $\frac{1}{X} \sim \Gamma(a,\lambda)$ |
| **6.** Snedecor-type (3) | $\frac{\theta^a}{B(a,b)} x^{b-1}(x+\theta)^{-a-b}$ | $(0,\infty)$ | $\frac{x(x+\theta)}{a-1}$ | $a > 1,\ b,\theta > 0$ | $\frac{b\theta}{a-1} > 0$ | $\delta = \frac{1}{a-1} > 0,\ \beta^2 > 4\delta\gamma$ <br> $\frac{\theta}{X+\theta} \sim B(a,b)$ |

(*) A random variable $X$ belongs to the Integrated Pearson Family if and only if there exist constants $c_1 \neq 0$ and $c_2 \in \mathbb{R}$ such that the density of $\widetilde{X} = c_1 X + c_2$ is contained in the table.
(1) For $n > 1$ and if $\mu = 0$ and $\delta = 1/(n-1) = \gamma/n$ then $X \sim t_n$.
(2) $C = C_\mu(\delta,\gamma) > 0$, with $C_0(\delta,\gamma) = \Gamma(1+1/(2\delta))\sqrt{\delta\gamma^{1+1/\delta}}/\Gamma(1/2+1/(2\delta))\sqrt{\pi}$, $C_\mu(\delta,\gamma) = 2^{1/\delta}\gamma^{1/2+1/(2\delta)}\sqrt{\delta}|\Gamma(1+1/(2\delta)+i\mu/(2\sqrt{\delta\gamma}))|^2/\Gamma(1+1/\delta)\pi$.
(3) For $n > 0$, $m > 2$ and if $a = m/2$, $b = n/2$, $\theta = m/n$ then $X \sim F_{n,m}$.

Remark 2.1. Since
$$(\mu-x)\frac{\exp\big(\frac{\mu}{\sqrt{\delta\gamma}}\tan^{-1}(x\sqrt{\delta/\gamma})\big)}{(\delta x^2+\gamma)^{1+1/(2\delta)}}=\frac{d}{dx}\frac{\exp\big(\frac{\mu}{\sqrt{\delta\gamma}}\tan^{-1}(x\sqrt{\delta/\gamma})\big)}{(\delta x^2+\gamma)^{1/(2\delta)}},$$
it follows that $\mathbb{E}\,X=\mu$ for the Student-type densities (type 4), while for all other cases it is evident that the mean is as displayed in Table 2.1. Next, it is easily verified that the densities of Table 2.1 satisfy the assumptions (B) of Proposition 3.3, below, with $\mu=\mathbb{E}\,X$, $p_2(x)=q(x)$ and $p_1(x)=\mu-x-q'(x)$, where $\mu$ and $q$ are as in the table. Hence, according to Proposition 3.3, all these densities are, indeed, integrated Pearson.

Corollary 2.2. *Assume that* $X\sim\mathrm{IP}(\mu;\delta,\beta,\gamma)$.

(a) *If* $\delta\le 0$ *then* $\mathbb{E}\,|X|^{\alpha}<\infty$ *for any* $\alpha\in[0,\infty)$.
(b) *If* $\delta>0$ *then* $\mathbb{E}\,|X|^{\alpha}<\infty$ *for any* $\alpha\in[0,1+1/\delta)$, *while* $\mathbb{E}\,|X|^{1+1/\delta}$ $=\infty$.

*Proof.* If $X\sim\mathrm{IP}(\mu;\delta,\beta,\gamma)$ then we can find constants $c_1\neq 0$ and $c_2\in\mathbb{R}$ such that the density of $\widetilde{X}=c_1X+c_2$ is contained in Table 2.1. Then, according to Proposition 2.1(vi), $\widetilde{X}\sim\mathrm{IP}(\widetilde{\mu};\widetilde{\delta},\widetilde{\beta},\widetilde{\gamma})$ with $\widetilde{\delta}=\delta$. The assertion follows from the fact that $\mathbb{E}\,|X|^{\alpha}<\infty$ if and only if $\mathbb{E}\,|c_1X+c_2|^{\alpha}<\infty$. ∎

Next, we shall obtain a recurrence for the moments and the central moments of a random variable $X\sim\mathrm{IP}(\mu;q)$. To this end we first prove a simple lemma.

Lemma 2.1. *If* $X\sim\mathrm{IP}(\mu;\delta,\beta,\gamma)$ *has support* $J(X)=(\alpha,\omega)$ *and* $\mathbb{E}\,|X|^n$ $<\infty$ *for some* $n\ge 1$ *(that is,* $\delta<1/(n-1)$*) then*
$$\lim_{x\nearrow\omega}x^kq(x)f(x)=\lim_{x\searrow\alpha}x^kq(x)f(x)=0,\qquad k=0,1,\dots,n-1,$$
*and, in general, for any* $c\in\mathbb{R}$,
$$\lim_{x\nearrow\omega}(x-c)^kq(x)f(x)=\lim_{x\searrow\alpha}(x-c)^kq(x)f(x)=0,\qquad k=0,1,\dots,n-1.$$

*Proof.* See arXiv:1205.2903v2, pp. 9–10. ∎

Lemma 2.2. *If* $X\sim\mathrm{IP}(\mu;\delta,\beta,\gamma)$ *and* $\mathbb{E}\,|X|^n<\infty$ *for some* $n\ge 2$ *(that is,* $\delta<1/(n-1)$*) then for any* $c\in\mathbb{R}$, *the central moments about* $c$ *satisfy the recurrence*
$$\mathbb{E}(X-c)^{k+1}=\frac{(\mu-c+kq'(c))\,\mathbb{E}(X-c)^k+kq(c)\,\mathbb{E}(X-c)^{k-1}}{1-k\delta},$$
$$k=1,\dots,n-1,$$
*with initial conditions* $\mathbb{E}(X-c)^0=1$, $\mathbb{E}(X-c)^1=\mu-c$, *where* $q(c)=\delta c^2+\beta c+\gamma$, $q'(c)=2\delta c+\beta$. *In particular,*

(i) *the usual moments* ($c = 0$) *satisfy the recurrence*

$$\mathbb{E}\, X^{k+1} = \frac{(\mu + k\beta)\, \mathbb{E}\, X^k + k\gamma\, \mathbb{E}\, X^{k-1}}{1 - k\delta}, \qquad k = 1, \ldots, n-1,$$

*with initial conditions* $\mathbb{E}\, X^0 = 1$ *and* $\mathbb{E}\, X^1 = \mu$;

(ii) *the central moments* ($c = \mu$) *satisfy the recurrence*

$$\mathbb{E}(X-\mu)^{k+1} = \frac{kq'(\mu)\, \mathbb{E}(X-\mu)^k + kq(\mu)\, \mathbb{E}(X-\mu)^{k-1}}{1 - k\delta}, \qquad k = 1, \ldots, n-1,$$

*with initial conditions* $\mathbb{E}(X-\mu)^0 = 1$ *and* $\mathbb{E}(X-\mu)^1 = 0$.

*Proof.* See arXiv:1205.2903v2, pp. 10–11. ∎

**3. Comparison with the ordinary Pearson system.** The ordinary Pearson family consists of absolutely continuous random variables $X$ supported in some (open) interval $(\alpha, \omega)$ such that their density $f$, which is assumed to be strictly positive and differentiable in $(\alpha, \omega)$, satisfies the *Pearson differential equation*

$$\frac{f'(x)}{f(x)} = \frac{p_1(x)}{p_2(x)}, \qquad \alpha < x < \omega, \tag{3.1}$$

where $p_1$ is a polynomial of degree at most one and $p_2$ is a polynomial of degree at most two. Since we can multiply the numerator and the denominator of (3.1) by the same nonzero constant, it is usually assumed, for convenience, that $p_1$ is a monic linear polynomial of degree one, e.g., $p_1(x) = x + a_0$. Although this restriction specifies both $p_1$ and $p_2$ whenever $p_1$ is nonconstant, it is not satisfactory for our purposes because it eliminates all rectangular (uniform over some interval) distributions and several $B(a, b)$ densities (those with $a + b = 2$)—see Table 2.1. Therefore, when we say that a function $f$ satisfies the Pearson differential equation (3.1) it will be assumed that $p_1$ is *any polynomial* of degree at most one (the cases $p_1 \equiv 0$ and $p_1 \equiv c \neq 0$ are allowed) and $p_2 \not\equiv 0$ is any polynomial of degree at most two. Note that common zeros of $p_1$ and $p_2$ are allowed inside the interval $(\alpha, \omega)$. Also, $p_1$ and $p_2$ may have common zeros outside the interval $(\alpha, \omega)$, which happens in the case of an exponential density.

Clearly, the ordinary Pearson family contains some random variables whose expectation does not exist, e.g., Cauchy. Sometimes it is asserted that, under finiteness of the first moment, (1.1) and (3.1) are equivalent—see, e.g., Korwar (1991, pp. 292–293). However, this is true only in particular cases, i.e., when we have made a 'correct' choice of $p_2$ (that is, $p_2(x) = \theta q(x)$ for some $\theta \neq 0$) and provided that a solution $f$ of (3.1) is considered in a maximal subinterval of the support of $p_2$, $\{x : p_2(x) \neq 0\}$. The following algorithmic procedure will always decide correctly if a given Pearson density belongs to the Integrated Pearson Family. The algorithm makes a correct choice of $p_2$, if it exists, as follows:

THE INTEGRATED PEARSON ALGORITHM

STEP 0. Assume that a Pearson density $f$ with finite (unknown) mean and (known) support $S(f) = \{x : f(x) > 0\} = (\alpha, \omega)$ satisfies $f'/f = \widetilde{p}_1/\widetilde{p}_2$ for given (real) polynomials $\widetilde{p}_1$, $\widetilde{p}_2$ (with $\widetilde{p}_2 \not\equiv 0$), of degree at most one and two, respectively.

STEP 1. Cancel the common factors of $\widetilde{p}_1$ and $\widetilde{p}_2$, if any. Then the resulting polynomials, say $\widetilde{p}_1^{(1)}$ and $\widetilde{p}_2^{(1)}$, are irreducible—they do not have any common zeros in $\mathbb{C}$. In case $\widetilde{p}_1 \equiv 0$, it suffices to define $\widetilde{p}_1^{(1)} \equiv 0$, $\widetilde{p}_2^{(1)} \equiv 1$.

STEP 2. If $\alpha > -\infty$ and $\widetilde{p}_2^{(1)}(\alpha) \neq 0$ then multiply both $\widetilde{p}_1^{(1)}$ and $\widetilde{p}_2^{(1)}$ by $x - \alpha$ and name the resulting polynomials $\widetilde{p}_1^{(2)}$ and $\widetilde{p}_2^{(2)}$; otherwise (i.e. if either $\alpha = -\infty$, or $\alpha > -\infty$ and $\widetilde{p}_2^{(1)}(\alpha) = 0$) set $\widetilde{p}_1^{(2)} = \widetilde{p}_1^{(1)}$ and $\widetilde{p}_2^{(2)} = \widetilde{p}_2^{(1)}$.

STEP 3. If $\omega < \infty$ and $\widetilde{p}_2^{(2)}(\omega) \neq 0$ then multiply both $\widetilde{p}_1^{(2)}$ and $\widetilde{p}_2^{(2)}$ by $\omega - x$ and name the resulting polynomials $p_1$ and $p_2$; otherwise (i.e. if either $\omega = \infty$, or $\omega < \infty$ and $\widetilde{p}_2^{(2)}(\omega) = 0$) set $p_1 = \widetilde{p}_1^{(2)}$ and $p_2 = \widetilde{p}_2^{(2)}$.

STEP 4. If $\deg(p_1) \leq 1$ and $\deg(p_2) \leq 2$, then $p_2$ is a correct choice and $f \sim \mathrm{IP}(\mu; q)$ with $q(x) = \theta p_2(x)$ for some $\theta \neq 0$; otherwise the given density $f$ does not belong to the Integrated Pearson Family.

It is clear that the above procedure starts with the equation $f'/f = \widetilde{p}_1/\widetilde{p}_2$ and, at Step 3, it produces two new (real) polynomials $p_1, p_2$, of degree at most three and four, respectively, such that $f'/f = p_1/p_2$. Moreover, $p_2(\alpha) = 0$ if $\alpha > -\infty$, $p_2(\omega) = 0$ if $\omega < \infty$ and $p_2(x) \neq 0$ for all $x \in (\alpha, \omega)$. Furthermore, because of Step 1, the polynomials $p_1(z)$ and $p_2(z)$ cannot have any common zeros in $\mathbb{C} \smallsetminus \{\alpha, \omega\}$.

The algorithm guarantees that a correct $p_2$ is chosen when such a $p_2$ exists. For example, the standard exponential density,

$$f(x) = e^{-x}, \qquad x > 0,$$

satisfies (3.1) when $(p_1, p_2) = (-1, 1)$, when $(p_1, p_2) = (-x, x)$ and when $(p_1, p_2) = (-x - 1, x + 1)$, but the correct choice is the second one. The standard uniform density,

$$f(x) = 1, \qquad 0 < x < 1,$$

satisfies (3.1) for $p_1 \equiv 0$ and for any $p_2$ (with no roots in $(0, 1)$), and the correct choice is $p_2 = x(1-x)$. The power density,

$$f(x) = 2x, \qquad 0 < x < 1,$$

satisfies (3.1) with $(p_1, p_2) = (2-x, x(2-x))$. A correct choice arises when we multiply both polynomials by $(1-x)/(2-x)$, that is, $(p_1, p_2) = (1-x, x(1-x))$. The Pareto density,

$$f(x) = \frac{2}{(x+1)^3}, \qquad x > 0,$$

satisfies (3.1) when $(p_1, p_2) = (-3, x+1)$, when $(p_1, p_2) = (-3x, x(x+1))$ and when $(p_1, p_2) = (-3(x+1), (x+1)^2)$, but the correct choice is the second one. The half-Normal density,

$$f(x) = \sqrt{\frac{2}{\pi}} e^{-x^2/2}, \qquad x > 0,$$

satisfies (3.1) in its interval support $(\alpha, \omega) = (0, \infty)$, although it does not satisfy (1.1). Here, there does not exist a correct choice for $p_2$. A more natural example is as follows: Consider the density

$$f(x) = \frac{C}{\sqrt{1+x^2}}, \qquad \alpha < x < \omega,$$

where $C = C(\alpha, \omega) > 0$ is the normalizing constant. This density satisfies the Pearson differential equation (3.1) in any finite interval $(\alpha, \omega)$, with $p_1 = -x$, $p_2 = 1 + x^2$, while its integral over unbounded intervals diverges. This density does not fulfill (1.1) so that it does not belong to the Integrated Pearson Family. Again there does not exist a correct choice for $p_2$.

The algorithm is justified by the following propositions.

Proposition 3.1. *Let $X \sim f$ and assume that the density $f$ satisfies the assumptions of Step 0. If $X \sim \mathrm{IP}(\mu; q)$ then the polynomials $p_1$ and $p_2$ of Step 3 are of degree at most one and two, respectively, and $q(x) = \theta p_2(x)$ for some $\theta \neq 0$.*

*Proof.* Since $X$ is integrated Pearson, so is $Y = \lambda X + c$ for all $\lambda \neq 0$ and $c \in \mathbb{R}$; see Proposition 2.1(vi). Also, its density $f_Y(x) = \frac{1}{|\lambda|} f\left(\frac{x-c}{\lambda}\right)$ satisfies, by assumption, the differential equation

$$\frac{f_Y'(x)}{f_Y(x)} = \frac{\widetilde{p}_1^Y(x)}{\widetilde{p}_2^Y(x)}, \quad x \in (\widetilde{\alpha}, \widetilde{\omega}),$$
$$\text{with } \widetilde{p}_1^Y(x) = \lambda \widetilde{p}_1\left(\frac{x-c}{\lambda}\right), \ \widetilde{p}_2^Y(x) = \lambda^2 \widetilde{p}_2\left(\frac{x-c}{\lambda}\right),$$

where $(\widetilde{\alpha}, \widetilde{\omega}) = (\lambda\alpha + c, \lambda\omega + c)$ or $(\lambda\omega + c, \lambda\alpha + c)$, according to whether $\lambda > 0$ or $\lambda < 0$. It is easily shown that the new polynomials $p_1, p_2$ (those that the algorithm produces at Step 3 for $f$) are related to the corresponding polynomials $p_1^Y$, $p_2^Y$ (those that the algorithm produces at Step 3 for $f_Y$) by

$$p_1^Y(x) = \lambda^i p_1\Big(\frac{x-c}{\lambda}\Big), \qquad p_2^Y(x) = \lambda^{i+1} p_2\Big(\frac{x-c}{\lambda}\Big),$$

for some $i \in \{1,2,3\}$. Therefore, it suffices to show that $\deg(p_i^Y) \le i$, $i = 1,2$, and that the quadratic polynomial $q_Y(x) = \lambda^2 q\big(\frac{x-c}{\lambda}\big)$ is related to $p_2^Y$ through $q_Y(x) = \theta p_2^Y(x)$ for some $\theta \neq 0$. Thus, without any loss of generality we may assume that $f$ is one of the densities given in Table 2.1.

Now observe that $(\widetilde{p}_1, \widetilde{p}_2)$ is always irreducible for types $1,4,5$ (Normal-type, Student-type, Reciprocal Gamma-type) with $\deg(\widetilde{p}_1) = 1$ for all types $1,4,5$, while $\deg(\widetilde{p}_2) = 0$ for type 1 and $\deg(\widetilde{p}_2) = 2$ for types 4 and 5. Since the corresponding supports are $\mathbb{R}$, $\mathbb{R}$ and $(0,\infty)$, respectively, and since in type 5, $\widetilde{p}_2(x) = \theta x^2$ for some $\theta \neq 0$, it follows that $(p_1,p_2) = (\widetilde{p}_1,\widetilde{p}_2)$, $q = \theta p_2$ for some $\theta \neq 0$, and the assertion follows.

For types $2,3$ and 6 (Gamma-type, Beta-type and Snedecor-type) the irreducibility of $\widetilde{p}_1$ and $\widetilde{p}_2$ depends on the parameters. Let us consider these cases in detail.

If $f \sim \Gamma(a,\lambda)$ with $a \neq 1$ ($a > 0$, $\lambda > 0$) then $\widetilde{p}_1 = \theta(a-1-\lambda x)$ and $\widetilde{p}_2 = \theta x$ for some $\theta \neq 0$, so that $\widetilde{p}_1, \widetilde{p}_2$ are irreducible with degree one. It follows that $p_i = \widetilde{p}_i$, $\deg(p_i) = 1$ ($i = 1,2$) and

$$q(x) = \frac{x}{\lambda} = \frac{p_2(x)}{\theta\lambda}.$$

If $f \sim \Gamma(1,\lambda)$ ($\lambda > 0$) then all possible choices for $(\widetilde{p}_1,\widetilde{p}_2)$ are given by $\widetilde{p}_1 = -\lambda\theta(x+c)$ and $\widetilde{p}_2 = \theta(x+c)$ for $\theta \neq 0$, $c \in \mathbb{R}$. Therefore, Step 3 yields $(p_1,p_2) = (-\lambda\theta x, \theta x)$, and thus $\deg(p_i) = 1$ ($i = 1,2$) with

$$q(x) = \frac{x}{\lambda} = \frac{p_2(x)}{\lambda\theta}.$$

If $f$ is of type 6 and $b \neq 1$ then

$$(\widetilde{p}_1(x), \widetilde{p}_2(x)) = \big(c((b-1)-(a+1)x), cx(x+\theta)\big) \qquad \text{for some } c \neq 0.$$

Here the parameters are $a > 1$, $b > 0$ and $\theta > 0$. It follows that $(p_1,p_2) = (\widetilde{p}_1,\widetilde{p}_2)$, $\deg(p_i) = i$ ($i = 1,2$) and

$$q(x) = \frac{x(x+\theta)}{a-1} = \frac{p_2(x)}{(a-1)c}.$$

If $f$ is of type 6 with $b = 1$ then all possible choices for $(\widetilde{p}_1,\widetilde{p}_2)$ are given by

$$\widetilde{p}_1(x) = -c(a+1)(x+\gamma),\ \widetilde{p}_2(x) = c(x+\theta)(x+\gamma) \text{ for some } c \neq 0,\ \gamma \in \mathbb{R}.$$

Therefore, Step 3 yields $(p_1,p_2) = (-c(a+1)x, cx(x+\theta))$. Thus, $\deg(p_i) = i$ ($i = 1,2$) and

$$q(x) = \frac{x(x+\theta)}{a-1} = \frac{p_2(x)}{(a-1)c}.$$

Finally, let $f$ be of type 3 (Beta-type), that is, $f \sim B(a,b)$ with $a,b>0$. If $a \neq 1$ and $b \neq 1$, it is easily shown that the polynomials

$$(\widetilde{p}_1(x), \widetilde{p}_2(x)) = \big(\theta(a-1-(a+b-2)x), \theta x(1-x)\big) \qquad (\theta \neq 0)$$

are irreducible, so that $(p_1,p_2) = (\widetilde{p}_1, \widetilde{p}_2)$, $\deg(p_i) = i$ $(i=1,2)$ and

$$q(x) = \frac{x(1-x)}{a+b} = \frac{p_2(x)}{(a+b)\theta}.$$

If $a=1$, $b \neq 1$, the most general form of $(\widetilde{p}_1, \widetilde{p}_2)$ is given by

$$(\widetilde{p}_1(x), \widetilde{p}_2(x)) = \big(-(b-1)\theta(x+c), \theta(1-x)(x+c)\big), \qquad \text{where } \theta \neq 0,\ c \in \mathbb{R}.$$

Therefore, Step 3 yields $(p_1,p_2) = (-(b-1)\theta x, \theta x(1-x))$. Thus, $\deg(p_i) = i$ $(i=1,2)$ and

$$q(x) = \frac{x(1-x)}{b+1} = \frac{p_2(x)}{(b+1)\theta}.$$

If $a \neq 1$, $b=1$, the most general form of $(\widetilde{p}_1, \widetilde{p}_2)$ is given by

$$(\widetilde{p}_1(x), \widetilde{p}_2(x)) = ((a-1)\theta(x+c), \theta x(x+c)), \quad \text{where } \theta \neq 0,\ c \in \mathbb{R}.$$

Therefore, Step 3 yields $(p_1,p_2) = ((a-1)\theta(1-x), \theta x(1-x))$, and thus $\deg(p_i) = i$ $(i=1,2)$,

$$q(x) = \frac{x(1-x)}{a+1} = \frac{p_2(x)}{(a+1)\theta}.$$

Finally, if $a=b=1$ (standard uniform density, $U(0,1) \equiv B(1,1)$) then $\widetilde{p}_1 \equiv 0$ so that $(p_1,p_2) = (0, x(1-x))$, $\deg(p_1) < 0$, $\deg(p_2) = 2$ and

$$q(x) = \frac{x(1-x)}{2} = \frac{p_2(x)}{2}.$$

This completes the proof. ∎

PROPOSITION 3.2. *Assume that $X \sim f$ where the density $f$ is differentiable with derivative $f'$ in its* (*known*) *interval support* $(\alpha,\omega)$ *and has finite* (*unknown*) *mean. Then the following statements are equivalent:*

- (A) *$f$ satisfies* (3.1) *for some* (*real*) *polynomials $p_1$* (*of degree at most one*) *and $p_2 \not\equiv 0$* (*of degree at most two*) *with $p_2(\alpha) = 0$ if $\alpha > -\infty$, $p_2(\omega) = 0$ if $\omega < \infty$ and $p_2(x) \neq 0$ for all $x \in (\alpha,\omega)$.*
- (B) *$X \sim \mathrm{IP}(\mu;q)$ for some $q(x) = \delta x^2 + \beta x + \gamma$ with $\{x : q(x) > 0\} = (\alpha,\omega)$ and some $\mu \in (\alpha,\omega)$.*

*Moreover, if* (A) *and* (B) *hold, then there exists a constant $\theta \neq 0$ such that $q(x) = \theta p_2(x)$, $x \in \mathbb{R}$.*

*Proof.* See arXiv:1205.2903v2, pp. 15–17. ∎

Eventually, Proposition 3.2 says that for a particular choice of $p_2$ to be correct it is necessary and sufficient that $p_2$ remains nonzero in $(\alpha,\omega)$ and vanishes at every finite endpoint of $(\alpha,\omega)$ (if any).

If the mean $\mu$ is known, then another simple criterion for an ordinary Pearson variate to belong to the Integrated Pearson Family is provided by the following proposition.

PROPOSITION 3.3. *Let $X$ be a random variable with density $f$ and finite mean $\mu$. Assume that the set $\{x : f(x) > 0\}$ is the (bounded or unbounded) interval $J(X) = (\alpha, \omega)$ and that $f$ is differentiable in $(\alpha, \omega)$, with derivative $f'(x)$, $\alpha < x < \omega$. Then the following are equivalent:*

(A) $X \sim \mathrm{IP}(\mu; q)$.
(B) *The density $f$ satisfies (3.1) and the polynomials $p_1$ ($p_1 \equiv 0$ is allowed) and $p_2$ can be chosen in such a way that:*
  (i) *there exists a constant $\theta \neq 0$ such that $p_1(x)+p_2'(x) = (\mu-x)/\theta$ for all $x \in \mathbb{R}$,*
  (ii) *either $\lim_{x\searrow\alpha} p_2(x)f(x) = 0$ or $\lim_{x\nearrow\omega} p_2(x)f(x) = 0$.*

*If (i) and (ii) are true then the polynomials $p_2$ and $q$ are related through $q(x) = \theta p_2(x)$ where $\theta \neq 0$ is as in (i). Moreover, if (3.1) is satisfied in an* unbounded *interval $(\alpha, \omega)$ then (ii) is unnecessary since it is implied by (i).*

*Proof.* If $X \sim \mathrm{IP}(\mu; q)$, then we see from (2.1) that (3.1) is satisfied for the polynomials $p_1(x) = \mu - x - q'(x)$ and $p_2(x) = q(x)$. With this choice of $p_1$, $p_2$, Proposition 2.1 shows that (i) (with $\theta = 1$) is valid. Also, (ii) reduces to $p_2(x)f(x) = q(x)f(x) \to 0$ as $x \nearrow \omega$ or $x \searrow \alpha$; this follows by an obvious application of dominated convergence since the mean exists and, by assumption, $p_2(x)f(x) = q(x)f(x) = \int_\alpha^x (\mu - t)f(t)\,dt$—see (1.1).

Conversely, (3.1) and (i) imply that $[\theta p_2(t)f(t)]' = (\mu-t)f(t)$, $\alpha < t < \omega$. Integrating this equation over the interval $[x, y] \subset (\alpha, \omega)$, we obtain

$$\int_x^y (\mu - t)f(t)\,dt = \theta p_2(y)f(y) - \theta p_2(x)f(x), \qquad \alpha < x < y < \omega. \tag{3.2}$$

Taking limits as $x \searrow \alpha$ in (3.2), using the dominated convergence theorem for the l.h.s. and the first assumption in (ii) for the r.h.s. we obtain

$$\int_\alpha^y (\mu - t)f(t)\,dt = \theta p_2(y)f(y), \qquad \alpha < y < \omega.$$

That is, $X \sim \mathrm{IP}(\mu; q)$ with $q(x) = \theta p_2(x)$. We arrive at the same conclusion if we make use of the second assumption in (ii) and evaluate the limits as $y \nearrow \omega$ in (3.2). Indeed, in this case we obtain

$$\int_x^\omega (t - \mu)f(t)\,dt = \theta p_2(x)f(x) = q(x)f(x), \qquad \alpha < x < \omega,$$

which is equivalent to (1.1), since $\int_\alpha^\omega (\mu - t)f(t)\,dt = 0$.

It is clear that, in the presence of (i), both assumptions in (ii) are equivalent. In fact, (3.2) shows that both limits $\lim_{y\nearrow\omega} p_2(y)f(y)$ and $\lim_{x\searrow\alpha} p_2(x)f(x)$ exist (in $\mathbb{R}$) and are equal. Indeed,

$$\theta p_2(y)f(y) = \theta p_2(x)f(x) + \int_x^y (\mu - t)f(t)\,dt, \qquad \alpha < x < y < \omega,$$

and the existence of the first moment implies that, as $y \nearrow \omega$, the r.h.s. has the well-defined finite limit $C(x) = \theta p_2(x)f(x) + \int_x^\omega (\mu - t)f(t)\,dt$. The l.h.s, however, is independent of $x$ and, certainly, the same is true for its limit, so that $C(x) \equiv C$. Hence,

$$\theta p_2(x)f(x) = C + \int_x^\omega (t-\mu)f(t)\,dt, \qquad \alpha < x < \omega.$$

Since $\lim_{x\searrow\alpha} \int_x^\omega (t-\mu)f(t)\,dt = \int_\alpha^\omega (t-\mu)f(t)\,dt = 0$, we conclude that

$$\lim_{x\searrow\alpha} p_2(x)f(x) = \lim_{y\nearrow\omega} p_2(y)f(y) = \frac{C}{\theta} \in \mathbb{R}.$$

It remains to verify that if (3.1) holds in an unbounded interval $(\alpha,\omega)$ and $X$ has a finite first moment then (i) implies (ii). To this end assume that $\omega = \infty$, so that $J(X) = (\alpha,\infty)$ with $\alpha \in [-\infty,\infty)$. It follows that $f'(x) = p_1(x)f(x)/p_2(x)$ does not change sign for large enough $x$, and thus $f'(x) < 0$ for $x > x_0$. Therefore, for $x > \max\{2x_0, 0\}$,

$$0 < x^2 f(x) = \frac{8}{3} f(x) \int_{x/2}^x t\,dt < \frac{8}{3} \int_{x/2}^x tf(t)\,dt < \frac{8}{3} \int_{x/2}^\infty tf(t)\,dt \to 0$$

as $x \to \infty$, i.e., $f(x) = o(x^{-2})$ as $x \to \infty$. Thus, $p_2(x)f(x) \to 0$ as $x \to \infty$. The case $\alpha = -\infty$ is similar and the proof is complete. ∎

**4. Are the Rodrigues-type polynomials orthogonal in the ordinary Pearson system?** Associated with any Pearson density $f$ is a (unique) sequence of polynomials, defined by a Rodrigues-type formula. As we shall see, these polynomials are given in terms of the pair $(p_1, p_2)$ that appears in the differential equation (3.1). That is, they do not depend at all either on $f$ or on the interval $(\alpha,\omega)$.

These considerations will become clear if we slightly relax the form of differential equation (3.1) and permit more solutions, as follows:

Definition 4.1. Let $\emptyset \neq (\alpha,\omega) \subseteq \mathbb{R}$, and consider a pair of real polynomials $(p_1,p_2) = (a_0+a_1x, b_0+b_1x+b_2x^2)$ such that $p_2 \not\equiv 0$ (i.e., $|b_0|+|b_1|+|b_2| > 0$). The pair $(p_1,p_2)$ is called *Pearson-compatible* in $(\alpha,\omega)$, or simply *compatible*, if there exists a differentiable function $f : (\alpha,\omega) \to \mathbb{R}$, $f \not\equiv 0$ ($f$ is not assumed to be nonnegative or integrable), such that the following *generalized*

*Pearson differential equation* is fulfilled:

$$p_2(x)f'(x) = p_1(x)f(x), \qquad \alpha < x < \omega. \tag{4.1}$$

In other words, $(p_1, p_2)$ is compatible if (4.1) has a nontrivial solution $f$.

It is easily seen that $(p_1, p_2)$ is compatible whenever $p_2$ has no roots in $(\alpha, \omega)$. In this case, the general solution $f$ is $C^\infty(\alpha, \omega)$ and can be chosen to be strictly positive in $(\alpha, \omega)$. The presence of a zero of $p_2$ in $(\alpha, \omega)$, however, may result in incompatibility. For example, in the interval $(\alpha, \omega) = (-2, 2)$ the pair $(p_1, p_2) = (4x, x^2-1)$ is compatible, in contrast to the pair $(p_1, p_2) = (x, x^2 - 1)$.

If $(p_1, p_2)$ is compatible in $(\alpha, \omega)$ then we can find the general solution as follows: First we solve (4.1) separately in any open subinterval of $(\alpha, \omega) \cap \{x : p_2(x) \neq 0\}$. Clearly, there are at most three subintervals and the general solutions for the distinct intervals $(J_1, J_2, J_3) = ((\alpha, \rho_1), (\rho_1, \rho_2), (\rho_2, \omega))$ will be of the form $f_i = C_i e^{g_i}$ for some $g_i \in C^\infty(J_i)$, $i = 1, 2, 3$, with $C_i$ being arbitrary constants. Next, we match the solutions and their first derivatives at the common endpoints of any two $J_i$; any such point is, necessarily, a zero of $p_2$. The compatibility of $(p_1, p_2)$ guarantees that this procedure will success in producing some solution $f \not\equiv 0$ (in which case, $|f| \geq 0$ will be also a nontrivial solution), but it may happen that $f_i \equiv 0$ in some $J_i$. The following proposition describes all possible cases for the support of $f$.

Proposition 4.1. *Assume that the function $f : (\alpha, \omega) \to \mathbb{R}$, $f \not\equiv 0$ (not necessarily positive or integrable) is differentiable in $(\alpha, \omega)$ and satisfies the differentiable equation* (4.1) *for some real polynomials $p_1(x) = a_0 + a_1 x$ and $p_2(x) = b_0 + b_1 x + b_2 x^2$ with $|b_0| + |b_1| + |b_2| > 0$. Then the support of $f$, $S(f) := \{x \in (\alpha, \omega) : f(x) \neq 0\}$, is either of the form $(\widetilde{\alpha}, \widetilde{\omega}) \subseteq (\alpha, \omega)$ with $\alpha \leq \widetilde{\alpha} < \widetilde{\omega} \leq \omega$, or of the form $(\widetilde{\alpha}, \rho_1) \cup (\rho_2, \widetilde{\omega}) \subseteq (\alpha, \omega)$ with $\alpha \leq \widetilde{\alpha} < \rho_1 \leq \rho_2 < \widetilde{\omega} \leq \omega$, or finally of the form $(\alpha, \rho_1) \cup (\rho_1, \rho_2) \cup (\rho_2, \omega)$, with $\alpha < \rho_1 < \rho_2 < \omega$. Moreover, the boundary of $S(f)$ is contained in the set $\{\alpha, \omega\} \cup \{x \in (\alpha, \omega) : p_2(x) = 0\}$, that is, $\partial S(f) \subseteq \{\alpha, \omega\} \cup \{x \in (\alpha, \omega) : p_2(x) = 0\}$. Finally, for any solution $f$, $f(\rho) = 0$ (that is, $\rho \notin S(f)$) for all $\rho$ that satisfy $p_2(\rho) = 0$ and $p_1(\rho) \neq 0$.*

Corollary 4.1. *The differential equation* (4.1) *has a nontrivial and nonnegative solution if and only if the pair $(p_1, p_2)$ is compatible in $(\alpha, \omega)$. Moreover, assuming that $(p_1, p_2)$ is compatible in $(\alpha, \omega)$, it follows that:*

(a) *any nonnegative solution is of the form $|f|$ for some solution $f$;*
(b) *the support $S(f) = \{x \in (\alpha, \omega) : f(x) \neq 0\}$ of any nontrivial solution $f$ of* (4.1) *is a union of one, two or three disjoint open intervals of positive length, and the same is true for any nonnegative and nontrivial solution;*

(c) *the boundary points of* $S(f) = S(|f|)$ *of any nontrivial solution* $f$ *of* (4.1) *are either roots of* $p_2$ *or boundary points of* $(\alpha, \omega)$.

We now turn to the corresponding Rodrigues polynomials. It is well-known that the (generalized) Pearson differential equation (4.1) produces a sequence of polynomials $\{h_k, k = 0, 1, 2, \ldots\}$, defined by a *Rodrigues-type formula.*

THEOREM 4.1 (Hildebrandt, 1931, p. 401; Beale, 1941, pp. 99–100; Diaconis and Zabell, 1991, p. 295). *Assume that a function* $f : (\alpha, \omega) \to \mathbb{R}$ (*not necessarily positive or integrable*) *does not vanish identically in* $(\alpha, \omega)$ *and satisfies the differential equation* (4.1) *for some polynomials* $p_1(x) = a_0 + a_1 x$ *and* $p_2(x) = b_0 + b_1 x + b_2 x^2$, *with* $|b_0| + |b_1| + |b_2| > 0$. *Then the set* $\{x \in (\alpha, \omega) : f(x) \neq 0\}$ *contains some interval of positive length, and the function*

$$h_k(x) := \frac{1}{f(x)} \frac{d^k}{dx^k}[p_2^k(x) f(x)], \quad x \in (\alpha, \omega) \smallsetminus \{x : f(x) = 0\}, \quad k = 0, 1, 2, \ldots, \tag{4.2}$$

*is a polynomial* (*more precisely,* $h_k$ *is the restriction to* $(\alpha, \omega) \smallsetminus \{x : f(x) = 0\}$ *of a polynomial* $\widetilde{h}_k : \mathbb{R} \to \mathbb{R}$) *with*

$$\deg(h_k) \le k \quad \textit{and} \quad \mathrm{lead}(h_k) = \prod_{j=k+1}^{2k} (a_1 + j b_2), \quad k = 0, 1, 2, \ldots, \tag{4.3}$$

*where* $\mathrm{lead}(h_k) := \lim_{x\to\infty} \widetilde{h}_k(x)/x^k$ *denotes the coefficient of* $x^k$ *in* $h_k(x)$.

Hildebrandt (1931) actually showed that the relation $p_2 f' = p_1 f$ implies that $D^k[p_2^k f] = \widetilde{h}_k f$, $k = 0, 1, 2, \ldots$, where the polynomials $\widetilde{h}_k$ (with $\deg(\widetilde{h}_k) \le k$) are defined inductively. Each polynomial $\widetilde{h}_k$ can be viewed as the value of a functional $\mathcal{R}_k$ that maps the (arbitrary) pair $(p_1, p_2)$ to a real polynomial of degree at most $k$. The form of this functional is

$$(p_1, p_2) \mapsto \mathcal{R}_k(p_1, p_2) := \widetilde{h}_k = \sum_{r,i,j} C_{k;rij}^{a_1,b_2} (p_1)^r (p_2')^i (p_2)^j,$$

where the sum ranges over all integers $r, i, j \ge 0$ with $r + i + 2j \le k$, and the constant $C_{k;rij}^{a_1,b_2}$ depends only on $k, r, i, j, p_1' = a_1$ and $p_2'' = 2b_2$. Clearly, given an arbitrary pair $(p_1, p_2)$ with $p_2 \not\equiv 0$, we can fix an interval $(\alpha, \omega)$ containing no roots of $p_2$. With the help of a positive solution $f$ of the differential equation (4.1) one can determine $h_k(x)$, $\alpha < x < \omega$, using the Rodrigues-type formula (4.2). Obviously, this $h_k$ extends uniquely to $\widetilde{h}_k$.

To give an idea about the nature of the polynomials in (4.2) we give the first four:

$$
\begin{aligned}
h_0 &= 1;\\
h_1 &= p_1 + p_2' = (a_1+2b_2)x + (a_0+b_1);\\
h_2 &= p_1^2 + 3p_1p_2' + p_1'p_2 + 2p_2p_2'' + 2(p_2')^2\\
&= (a_1+3b_2)(a_1+4b_2)x^2 + 2(a_0+2b_1)(a_1+3b_2)x\\
&\quad + (a_0+b_1)(a_0+2b_1) + b_0(a_1+4b_2);\\
h_3 &= p_1^3 + 6p_1^2p_2' + 3p_1p_1'p_2 + 8p_1p_2p_2'' + 11p_1(p_2')^2\\
&\quad + 7p_1'p_2p_2' + 18p_2p_2'p_2'' + 6(p_2')^3\\
&= (a_1+4b_2)(a_1+5b_2)(a_1+6b_2)x^3 + 3(a_0+3b_1)(a_1+4b_2)(a_1+5b_2)x^2\\
&\quad + 3(a_1+4b_2)[(a_0+2b_1)(a_0+3b_1) + b_0(a_1+6b_2)]x\\
&\quad + a_0^3 + 6a_0^2b_1 + a_0[11b_1^2 + b_0(3a_1+16b_2)] + b_1[6b_1^2 + b_0(7a_1+36b_2)].
\end{aligned}
$$

Provided that the solution $f$ of (4.1) is a probability density in $(\alpha,\omega)$, the polynomials $h_k$ are candidates to form an orthogonal system for $f$. Indeed, Hildebrandt (1931, pp. 404–405) showed that each $h_k$ satisfies a specific second order differential equation in $(\alpha,\omega)$. Using this differential equation, Diaconis and Zabell (1991) proved that the $h_k$ are eigenfunctions of a particular self-adjoint, second order Sturm–Liouville differential equation; thus, their orthogonality with respect to the density $f$ is a consequence of Sturm–Liouville theory. Specifically, it is shown in Diaconis and Zabell (1991, Theorem 1, p. 295) that each polynomial $h_k$ satisfies the equation

$$
[f(x)p_2(x)h_k'(x)]' = k(a_1+(k+1)b_2)f(x)h_k(x), \qquad \alpha < x < \omega,\ k = 0,1,2,\ldots. \tag{4.4}
$$

An adaption of the Diaconis–Zabell approach to the present general case reveals that the orthogonality is valid only when a number of regularity conditions is satisfied. It will be proved here that these regularity conditions give an equivalent definition of the integrated Pearson system. In fact, it will be shown that the Rodrigues polynomials (4.2) are orthogonal with respect to the corresponding density $f$ if and only if it belongs to Integrated Pearson Family, provided that we have chosen a correct $p_2$ in the differential equation (4.1), i.e., provided that $p_2 = q/\theta$ for some $\theta \neq 0$. Note that even for integrated Pearson densities, an incorrect choice of $p_2$ results in nonorthogonality of the Rodrigues polynomials; see, e.g., the polynomials $h_k = P_k^2$ given in Diaconis and Zabell (1991, p. 297) for the Beta-type density $f(x) = Cx^N$, $0 < x < x_0$. In light of Proposition 3.2 (and Table 2.1), a correct choice is $p_2 = x(x_0 - x)$.

In order to discuss the orthogonality of $h_k$ we first show the following lemma.

LEMMA 4.1. *Let $f$ be a density satisfying* (4.1) *and for fixed $k, m \in \{0, 1, \ldots\}$, $k \neq m$, consider the polynomials $h_k$ and $h_m$ given by* (4.2). *Assume that*

(a) *$f$ has a suitable number of moments so that $\int_\alpha^\omega |h_k(t)h_m(t)|f(t)\,dt < \infty$;*
(b) *$a_1 + (k+m+1)b_2 \neq 0$;*
(c) *$\lim_{x\nearrow\omega}\{p_2(x)f(x)[h_k'(x)h_m(x) - h_k(x)h_m'(x)]\}$ $= \lim_{x\searrow\alpha}\{p_2(x)f(x)[h_k'(x)h_m(x) - h_k(x)h_m'(x)]\}$.*

*Then*

$$\int_\alpha^\omega h_k(x)h_m(x)f(x)\,dx = 0.$$

[We shall show that, under (a) and (b), both limits in (c) exist (in $\mathbb{R}$), but it is not guaranteed that they are equal. In fact, their difference is $(k-m)(a_1+(k+m+1)b_2) \times \int_\alpha^\omega h_k(t)h_m(t)f(t)\,dt$.]

*Proof.* Multiply (4.4) by $h_m$, interchange the roles of $k$ and $m$ and subtract the resulting equations to get

$$\begin{aligned}(4.5)\qquad &\lambda h_k(t)h_m(t)f(t)\\ &\quad = h_m(t)[f(t)p_2(t)h_k'(t)]' - h_k(t)[f(t)p_2(t)h_m'(t)]', \qquad \alpha < t < \omega,\end{aligned}$$

where $\lambda = (k-m)(a_1+(k+m+1)b_2) \neq 0$, by (b). Now, it is easy to verify the *Lagrange identity*

$$\begin{aligned}(4.6)\qquad &\{[f(t)p_2(t)h_k'(t)]h_m(t) - [f(t)p_2(t)h_m'(t)]h_k(t)\}'\\ &\qquad\qquad = h_m(t)[f(t)p_2(t)h_k'(t)]' - h_k(t)[f(t)p_2(t)h_m'(t)]'.\end{aligned}$$

Integrating (4.5) over $[x, y] \subseteq (\alpha, \omega)$, and in view of (4.6), we conclude that

$$\begin{aligned}\int_x^y h_k(t)h_m(t)f(t)\,dt &= \frac{1}{\lambda}p_2(y)f(y)[h_k'(y)h_m(y) - h_k(y)h_m'(y)]\\ &\quad - \frac{1}{\lambda}p_2(x)f(x)[h_k'(x)h_m(x) - h_k(x)h_m'(x)].\end{aligned}$$

Taking limits as $x \searrow \alpha$ and $y \nearrow \omega$ and using (a) and (c) we obtain the desired result. Working as in the proof of Proposition 3.3, it is easily seen that both limits in (c) exist in $\mathbb{R}$ whenever (a) and (b) hold. In fact, under (a),

$$\begin{aligned}(4.7)\qquad &(k-m)(a_1+(k+m+1)b_2)\int_\alpha^\omega h_k(t)h_m(t)f(t)\,dt\\ &\quad = \lim_{y\nearrow\omega} p_2(y)f(y)[h_k'(y)h_m(y) - h_k(y)h_m'(y)]\\ &\qquad - \lim_{x\searrow\alpha} p_2(x)f(x)[h_k'(x)h_m(x) - h_k(x)h_m'(x)]. \quad \blacksquare\end{aligned}$$

The following result is an immediate consequence of Lemma 4.1.

THEOREM 4.2. *Let $f$ be a density in $(\alpha,\omega)$ which satisfies* (4.1). *For some* (*fixed*) $n\in\{1,2,\dots\}$ *consider the set* $\mathcal{H}_n:=\{h_0,h_1,\dots,h_n\}$, *formed by the first $n+1$ polynomials in* (4.2). *Then $\mathcal{H}_n$ is an orthogonal system* (*containing only nonzero elements*) *with respect to $f$ if and only if the following conditions are satisfied:*

(i) *The density $f$ has $2n-1$ finite moments;*
(ii) $\prod_{j=2}^{2n}(a_1+jb_2)\neq 0$;
(iii) $\lim_{x\nearrow\omega} x^j p_2(x)f(x)=\lim_{x\searrow\alpha} x^j p_2(x)f(x)$ *for each* $j\in\{0,1,\dots,2n-2\}$.

*Proof.* Let $X\sim f$ and assume first that (i)–(iii) are satisfied. Condition (ii) shows, in view of (4.3), that $\deg(h_k)=k$ for all $k\in\{0,1,\dots,n\}$. Fix $k,m\in\{0,1,\dots,n\}$ with $m\neq k$. Since $\mathbb{E}\,|X|^{2n-1}<\infty$ by (i), it follows that $\mathbb{E}\,|h_k(X)h_m(X)|<\infty$, i.e. the integral $\int_\alpha^\omega h_k(x)h_m(x)f(x)\,dx$ is (well-defined and) finite. Finally, since $h_k'h_m-h_kh_m'$ is a polynomial of degree $k+m-1$ (observe that $\operatorname{lead}(h_k'h_m-h_kh_m')=(k-m)\operatorname{lead}(h_k)\operatorname{lead}(h_m)\neq 0$), (iii) ensures that assumption (c) of Lemma 4.1 is also fulfilled, and hence

$$\int_\alpha^\omega h_k(x)h_m(x)f(x)\,dx=0.$$

Conversely, assume that the set $\mathcal{H}_n=\{h_0,h_1,\dots,h_n\}$ is orthogonal with respect to $f$. That is, $\mathbb{E}\,|h_k(X)h_m(X)|=\int_\alpha^\omega|h_k(x)h_m(x)|f(x)\,dx<\infty$ for all $k,m\in\{0,1,\dots,n\}$ with $m\neq k$, and $\int_\alpha^\omega h_k(x)h_m(x)f(x)\,dx=0$. It follows that $\deg(h_k)=k$ for all $k=1,\dots,n$. To see this, let $k$ be the smallest integer in $\{1,\dots,n\}$ for which $\operatorname{lead}(h_k)=0$. Then we can write $h_k(x)=\sum_{j=0}^{k-1}c_jh_j(x)$ for some constants $c_j$, and this implies that

$$h_k^2(x)f(x)=\Big|\sum_{j=0}^{k-1}c_jh_j(x)h_k(x)\Big|f(x)\le\sum_{j=0}^{k-1}|c_j|\,|h_j(x)h_k(x)|f(x).$$

Subsequently, the inequality

$$\int_\alpha^\omega h_k^2(x)f(x)\,dx\le\sum_{j=0}^{k-1}|c_j|\int_\alpha^\omega|h_k(x)h_j(x)|f(x)\,dx<\infty$$

shows that $h_k\in L_f^2(\alpha,\omega)$, and finally

$$\int_\alpha^\omega h_k^2(x)f(x)\,dx=\sum_{j=0}^{k-1}c_j\int_\alpha^\omega h_k(x)h_j(x)f(x)\,dx=0,$$

by the orthogonality assumption. Since $h_k$ is continuous (a polynomial) and $f$ is positive in a subinterval of $(\alpha,\omega)$ of positive length, it follows that

$h_k \equiv 0$, which contradicts the assumption that $\mathcal{H}_n$ contains only nonzero elements. Therefore, $\prod_{k=0}^{n} \text{lead}(h_k) \neq 0$, and (4.3) yields (ii). Obviously, $\mathbb{E}\,|h_n(X)h_{n-1}(X)| < \infty$ is equivalent to $\mathbb{E}\,|X|^{2n-1} < \infty$ and (i) follows. Since $g_{k,m} = h_k' h_m - h_k h_m'$ is a polynomial of degree exactly $k+m-1$ (for $k \neq m$), one can form a linearly independent set

$$\{g_0, g_1, \ldots, g_{2n-2}\} \subseteq \{g_{k,m} : k, m = 0, 1, \ldots, n,\ k \neq m\}$$

with $\deg(g_j) = j$ for each $j$. Applying (4.7) inductively to $g_0, g_1, \ldots, g_{2n-2}$, we get (iii). ∎

Example 4.1. It may happen that $h_k \equiv 0$ for all $k \geq 1$. For instance consider the density $f(x) = C/x$, $1 < x < 2$. This density satisfies (4.1) with $(p_1, p_2) = (-1, x)$. Although $\int_1^2 h_k h_m f = 0$ for $m \neq k$, the trivial system $\mathcal{H}_n = \{1, 0, \ldots, 0\}$ is not considered as orthogonal in this case. Condition (ii) of Theorem 4.2 eliminates such trivial cases.

Example 4.2. The density $f(x) = \frac{3}{2}x^2$, $-1 < x < 1$, satisfies (4.1) in $(\alpha, \omega) = (-1, 1)$. The choice $(p_1, p_2) = (2, x)$ leads to constant polynomials, $h_k \equiv (k+2)!/2$. A set $\{h_k, h_m\}$ can never be orthogonal; this shows that the condition (b) of Lemma 4.1 is necessary. On the other hand, the choice $(p_1, p_2) = (2x, x^2)$ yields the polynomials $h_k = c_k x^k$ with $c_k = (2k+2)!/(k+2)!$. The limits in Lemma 4.1(c) are $\frac{3}{2}c_k c_m(k-m)$ and $\frac{3}{2}c_k c_m(k-m)(-1)^{k+m+1}$. They are equal if and only if $k+m$ is odd, in which case $h_k$ and $h_m$ are, obviously, orthogonal. Any set $\mathcal{H}$ containing $\{h_k, h_m, h_s\}$ $(k \neq m \neq s \neq k)$ cannot be an orthogonal set, because at least one of $k+m$, $k+s$, $m+s$ is even.

Remark 4.1. While the density $f$ of Example 4.2 satisfies the (generalized) Pearson differential equation (4.1) and has finite moments of any order, the system $\{h_0, h_1, h_2\}$ fails to be orthogonal. The same is true for the Pearson density

$$f(x) = \frac{C}{\sqrt{1+x^2}}, \quad -\infty < \alpha < x < \omega < \infty.$$

Now $(p_1, p_2) = (-x, 1+x^2)$ and $\{h_0, h_1, h_2\} = \{1, x, 3+6x^2\}$, so that $h_0 h_2 \geq 3$ and the system $\{h_0, h_1, h_2\}$ cannot be orthogonal (with respect to any measure). Does this happen because these $f$ lie outside the Integrated Pearson Family? Therefore, it is natural to pose the following question:

> Suppose that a given density $f$ has finite moments up to order $2n-1$ (for some fixed $n \geq 2$) and satisfies (4.1). If the system $\{h_0, h_1, \ldots, h_n\}$ of the first $n+1$ Rodrigues polynomials is orthogonal with respect to $f$, does this imply that $f$ belongs to the Integrated Pearson Family?

The answer is 'yes'. In particular, the following (stronger) result holds.

THEOREM 4.3. *Assume that a differentiable density $f$ with $S(f)=\{x : f(x)>0\}\subseteq(\alpha,\omega)$ has finite third moment and satisfies* (4.1). *Let $h_0\equiv 1$, $h_1,h_2$ be the first three Rodrigues polynomials given by* (4.2). *Consider the system $\mathcal{H}_2=\{h_0,h_1,h_2\}$ and assume that $\mathcal{H}_2$ is nontrivial, i.e., $h_1\not\equiv 0$ and $h_2\not\equiv 0$. If $\mathcal{H}_2$ is orthogonal with respect to $f$, then there exists a subinterval $(\alpha',\omega')\subseteq(\alpha,\omega)$, a quadratic polynomial*

$$q(x)=\delta x^2+\beta x+\gamma \quad \textit{with} \quad \{x : q(x)>0\}=(\alpha',\omega'),$$

*and $\mu\in(\alpha',\omega')$ such that $f\sim \mathrm{IP}(\mu;q)\equiv \mathrm{IP}(\mu;\delta,\beta,\gamma)$. Moreover, there exists a constant $\theta\neq 0$ such that $q(x)=\theta p_2(x)$, $x\in\mathbb{R}$.*

*Proof.* In view of Theorem 4.2 and the fact that $f$ has finite third moment, the orthogonality assumption is equivalent to

$$(a_1+2b_2)(a_1+3b_2)(a_1+4b_2)\neq 0 \tag{4.8}$$

and

$$L_j(\alpha)=L_j(\omega), \qquad j=0,1,2, \tag{4.9}$$

where

$$L_j(\alpha):=\lim_{x\searrow\alpha} x^j p_2(x)f(x), \qquad L_j(\omega):=\lim_{x\nearrow\omega} x^j p_2(x)f(x).$$

To simplify, we can apply an affine transformation $x\mapsto \lambda x+c$ ($\lambda\neq 0$, $c\in\mathbb{R}$) to $f$. By considering $\widetilde{f}(x)=\frac{1}{|\lambda|}f\big(\frac{x-c}{\lambda}\big)$ in place of $f$ it is easily seen that (4.1) is satisfied in the translated interval $(\widetilde{\alpha},\widetilde{\omega})$ for $\widetilde{p}_1(x)=\lambda p_1\big(\frac{x-c}{\lambda}\big)$ and $\widetilde{p}_2(x)=\lambda^2 p_2\big(\frac{x-c}{\lambda}\big)$; since $\widetilde{a}_1=a_1$ and $\widetilde{b}_2=b_2$, (4.8) remains unchanged. Obviously $f$ has finite third moment if and only if $\widetilde{f}$ does. Moreover, it is easily seen from (4.2) that the translated polynomials $\widetilde{h}_k$ are related to $h_k$ by $\widetilde{h}_k(x)=\lambda^k h_k\big(\frac{x-c}{\lambda}\big)$; thus, $\mathrm{lead}(\widetilde{h}_k)=\mathrm{lead}(h_k)$, and in particular the system $\mathcal{H}_2$ is nontrivial if and only if the same is true for $\widetilde{\mathcal{H}}_2:=\{\widetilde{h}_0,\widetilde{h}_1,\widetilde{h}_2\}$. The orthogonality of $\widetilde{\mathcal{H}}_2$ with respect to $\widetilde{f}$ is equivalent to the orthogonality of $\mathcal{H}_2$ with respect to $f$. Indeed,

$$\int_{\widetilde{\alpha}}^{\widetilde{\omega}} \widetilde{h}_k(x)\widetilde{h}_m(x)\widetilde{f}(x)\,dx=\lambda^{k+m}\int_{\alpha}^{\omega} h_k(x)h_m(x)f(x)\,dx.$$

It remains to verify that (4.9) are equivalent to $\widetilde{L}_j(\widetilde{\alpha})=\widetilde{L}_j(\widetilde{\omega})$ ($j=0,1,2$), where $\widetilde{L}_j(\widetilde{\alpha}):=\lim_{x\searrow\widetilde{\alpha}} x^j\widetilde{p}_2(x)\widetilde{f}(x)$ and $\widetilde{L}_j(\widetilde{\omega}):=\lim_{x\nearrow\widetilde{\omega}} x^j\widetilde{p}_2(x)\widetilde{f}(x)$. To this end, it suffices to observe the relations

$$\sum_{i=0}^{j}\binom{j}{i}\lambda^{i+1}c^{j-i}L_i(\alpha)=\begin{cases}\widetilde{L}_j(\widetilde{\alpha}) & \text{if } \lambda>0,\\ -\widetilde{L}_j(\widetilde{\omega}) & \text{if } \lambda<0,\end{cases}$$

$$\sum_{i=0}^{j}\binom{j}{i}\lambda^{i+1}c^{j-i}L_i(\omega)=\begin{cases}\widetilde{L}_j(\widetilde{\omega}) & \text{if } \lambda>0,\\ -\widetilde{L}_j(\widetilde{\alpha}) & \text{if } \lambda<0.\end{cases}$$

Thus, it is easily seen that $L_j(\alpha)=L_j(\omega)$ ($j=0,1,2$) if and only if $\widetilde{L}_j(\widetilde{\alpha})=\widetilde{L}_j(\widetilde{\omega})$ ($j=0,1,2$).

It is clear from the above considerations that, in view of Proposition 2.1(vi), we can apply any affine transformation, either to the polynomial $p_2$ or to the density $f$ and its support $(\alpha,\omega)$. Under such transformations, the conclusions, as well as the assumptions of our theorem, remain unchanged.

The rest of the proof is easy but tedious since we just have to examine all possible nonequivalent cases by solving the differential equation (4.1) in each case. The details are given in the arXiv version of the paper, arXiv:1205.2903v2, pp. 26–31. ■

**5. Orthogonality of the Rodrigues-type polynomials and of their derivatives within the Integrated Pearson Family.** Assume that $f$ is the density of a random variable $X\sim \mathrm{IP}(\mu;q)\equiv\mathrm{IP}(\mu;\delta,\beta,\gamma)$ with support $(\alpha,\omega)$. From Theorem 4.1 it follows that

(5.1) $$P_k(x):=\frac{(-1)^k}{f(x)}\frac{d^k}{dx^k}[q^k(x)f(x)],\qquad \alpha<x<\omega,\ k=0,1,2,\ldots,$$

is a polynomial with

(5.2) $$\deg(P_k)\le k\quad\text{and}\quad \mathrm{lead}(P_k)=\prod_{j=k-1}^{2k-2}(1-j\delta):=c_k(\delta),\qquad k=0,1,2,\ldots.$$

Obviously $c_0(\delta):=1$, i.e. an empty product equals one.

The polynomials $P_k$ are special cases of the polynomials $h_k$ defined by (4.2); in fact, $P_k=(-1)^k h_k$. They are particularly important because under natural moment conditions they are, indeed, orthogonal with respect to the density $f$; see, e.g., Diaconis and Zabell (1991, pp. 295–296); Johnson (1993); Papathanasiou (1995); Afendras et al. (2011). Moreover, the polynomials $P_k$ and their derivatives satisfy a number of useful properties that will be reviewed here. The first three are

$$P_0(x)=1,$$
$$P_1(x)=x-\mu,$$
$$P_2(x)=(1-\delta)(1-2\delta)x^2-2(1-\delta)(\mu+\beta)x+\mu^2+\beta\mu-(1-2\delta)\gamma.$$

An alternative simple proof of the orthogonality of the polynomials defined by (5.1) can be derived by means of the following covariance identity, which extends Stein's identity for the Normal distribution and has independent interest in itself.

Theorem 5.1 (Afendras et al., 2011, pp. 515–516). *Let $X \sim \mathrm{IP}(\mu;\delta,\beta,\gamma) \equiv \mathrm{IP}(\mu;q)$ with density $f$ and support $(\alpha,\omega)$. Assume that $X$ has $2k$ finite moments for some fixed $k\in\{1,2,\ldots\}$. Let $g:(\alpha,\omega)\to\mathbb{R}$ be any function such that $g\in C^{k-1}(\alpha,\omega)$, and assume that the function*

$$g^{(k-1)}(x) := \frac{d^{k-1}}{dx^{k-1}}g(x)$$

*is absolutely continuous in $(\alpha,\omega)$ with a.s. derivative $g^{(k)}$. If $\mathbb{E}\,q^k(X)|g^{(k)}(X)| < \infty$ then $\mathbb{E}\,|P_k(X)g(X)|<\infty$ and the following covariance identity holds:*

$$\mathbb{E}\,P_k(X)g(X) = \mathbb{E}\,q^k(X)g^{(k)}(X). \tag{5.3}$$

Here, the claim that $h:(\alpha,\omega)\to\mathbb{R}$ is an absolutely continuous function with a.s. derivative $h'$ means that there exists a Borel measurable function $h':(\alpha,\omega)\to\mathbb{R}$ such that $h'$ is integrable in every finite subinterval $[x,y]$ of $(\alpha,\omega)$ and

$$\int_x^y h'(t)\,dt = h(y)-h(x) \qquad \text{for all } [x,y]\subseteq(\alpha,\omega).$$

Corollary 5.1 (Afendras et al., 2011, p. 516). *Let $X\sim \mathrm{IP}(\mu;\delta,\beta,\gamma)\equiv \mathrm{IP}(\mu;q)$. Assume that for some $n\in\{1,2,\ldots\}$, $\mathbb{E}\,|X|^{2n}<\infty$, or equivalently $\delta<1/(2n-1)$. Then*

$$\begin{aligned}\mathbb{E}[P_k(X)P_m(X)] &= \delta_{k,m}k!\,\mathbb{E}\,q^k(X)\prod_{j=k-1}^{2k-2}(1-j\delta)\\ &= \delta_{k,m}k!c_k(\delta)\,\mathbb{E}\,q^k(X), \qquad k,m\in\{0,1,\ldots,n\},\end{aligned} \tag{5.4}$$

*where $\delta_{k,m}$ is Kronecker's delta.*

It should be noted that the orthogonality of $P_k$ and $P_m$, $k\neq m$, $k,m\in\{0,1,\ldots,n\}$, remains valid even if $\delta\in\left[\frac{1}{2n-1},\frac{1}{2n-2}\right)$; in this case, however, $P_n\notin L^2(\mathbb{R},X)$ since $\mathrm{lead}(P_n)>0$ and $\mathbb{E}\,|X|^{2n}=\infty$. On the other hand, in view of Corollary 2.2, the assumption $\mathbb{E}\,|X|^{2n}<\infty$ is equivalent to the condition $\delta<1/(2n-1)$. Therefore, for each $k\in\{0,1,\ldots,n\}$ and for all $j\in\{k-1,\ldots,2k-2\}$, we have $1-j\delta>0$, since $\{k-1,\ldots,2k-2\}\subseteq\{0,1,\ldots,2n-2\}$. Thus, $c_k(\delta)>0$. Since $\mathbb{P}[q(X)>0]=1$, $\deg(q)\le 2$ and $\mathbb{E}\,|X|^{2n}<\infty$, we conclude that $0<\mathbb{E}\,q^k(X)<\infty$ for all $k\in\{0,1,\ldots,n\}$. It follows that $\{\phi_0,\phi_1,\ldots,\phi_n\}\subset L^2(\mathbb{R},X)$, where

$$\begin{aligned}\phi_k(x) &:= \frac{P_k(x)}{(k!c_k(\delta)\,\mathbb{E}\,q^k(X))^{1/2}}\\ &= \frac{\frac{(-1)^k}{f(x)}\frac{d^k}{dx^k}[q^k(x)f(x)]}{(k!\,\mathbb{E}\,q^k(X)\prod_{j=k-1}^{2k-2}(1-j\delta))^{1/2}}, \qquad k=0,1,\ldots,n,\end{aligned} \tag{5.5}$$

is an orthonormal basis of all polynomials with degree at most $n$. Moreover, (5.2) shows that the leading coefficient is given by

$$\text{lead}(\phi_k) := d_k(\mu; q) = \left(\frac{\prod_{j=k-1}^{2k-2}(1 - j\delta)}{k!\,\mathbb{E}\,q^k(X)}\right)^{1/2} = \left(\frac{c_k(\delta)}{k!\,\mathbb{E}\,q^k(X)}\right)^{1/2} > 0, \qquad k = 0, 1, \ldots, n. \tag{5.6}$$

Let $X$ be any random variable with $\mathbb{E}\,|X|^{2n} < \infty$ and assume that the support of $X$ is not concentrated on a finite subset of $\mathbb{R}$. It is well known that we can always construct an orthonormal set of real polynomials up to order $n$. This construction is based on the first $2n$ moments of $X$, and is a by-product of the Gram–Schmidt orthonormalization process, applied to the linearly independent system $\{1, x, x^2, \ldots, x^n\} \subset L^2(\mathbb{R}, X)$. The orthonormal polynomials are then uniquely defined, apart from the fact that we can multiply each polynomial by $\pm 1$. It follows that the standardized Rodrigues polynomials $\phi_k$ of (5.5) are the unique orthonormal polynomials that can be defined for a density $f \sim \text{IP}(\mu; \delta, \beta, \gamma)$, provided $\text{lead}(\phi_k) > 0$. Therefore, it is useful to express the $L^2$-norm of each $P_k$ in terms of the parameters $\delta, \beta, \gamma$ and $\mu$ and, in view of (5.4) and (5.5), it remains to obtain an expression for $\mathbb{E}\,q^k(X)$. To this end, we first recall a definition from Papadatos and Papathanasiou (2001); cf. Goldstein and Reinert (1997).

Definition 5.1. Let $X \sim f$ and assume that $X$ has support $J(X) = (\alpha, \omega)$ and belongs to the Integrated Pearson Family, that is, $f \sim \text{IP}(\mu; q) \equiv \text{IP}(\mu; \delta, \beta, \gamma)$. Furthermore, assume that $\mathbb{E}\,X^2 < \infty$ (i.e. $\delta < 1$). Then we define $X^*$ to be the random variable with density $f^*$ given by

$$f^*(x) := \frac{q(x) f(x)}{\mathbb{E}\,q(X)}, \qquad \alpha < x < \omega.$$

Since $P_1 = x - \mu$, setting $k = 1$ in the covariance identity (5.3), we get (see Cacoullos and Papathanasiou, 1989; Papadatos and Papathanasiou, 2001)

$$\mathbb{E}[(X - \mu) g(X)] = \text{Cov}[X, g(X)] = \mathbb{E}[q(X) g'(X)]. \tag{5.7}$$

This identity is valid for all absolutely continuous functions $g : (\alpha, \omega) \to \mathbb{R}$ with a.s. derivative $g'$ such that $\mathbb{E}\,q(X)|g'(X)| < \infty$. Thus, applying (5.7) to the identity function $g(x) = x$, it is easily seen that $\mathbb{E}\,q(X) = \text{Var}\,X = \sigma^2$, so that (cf. Goldstein and Reinert, 1997)

$$X^* \sim f^*(x) = \frac{1}{\sigma^2} q(x) f(x), \qquad \alpha < x < \omega.$$

The following lemma shows that $X^*$ is integrated Pearson whenever $X$ is integrated Pearson and has finite third moment.

Lemma 5.1. *If* $X \sim \mathrm{IP}(\mu;\delta,\beta,\gamma) \equiv \mathrm{IP}(\mu;q)$ *with support* $J(X)=(\alpha,\omega)$ *and* $\mathbb{E}|X|^3<\infty$ *then* $X^* \sim \mathrm{IP}(\mu^*;q^*)$ *with the same support* $J(X^*)=J(X)=(\alpha,\omega)$,

$$\mu^*=\frac{\mu+\beta}{1-2\delta} \quad \textit{and} \quad q^*(x)=\frac{q(x)}{1-2\delta}, \qquad \alpha<x<\omega.$$

*Proof.* See arXiv:1205.2903v2, pp. 33–35. ∎

Theorem 5.2. *Let* $X$ *be a random variable with density* $f \sim \mathrm{IP}(\mu;q) \equiv \mathrm{IP}(\mu;\delta,\beta,\gamma)$, *supported in* $J(X)=(\alpha,\omega)$. *Furthermore, assume that* $\mathbb{E}|X|^{2n+1}<\infty$ *(i.e.* $\delta<1/(2n)$*) for some fixed* $n\in\{0,1,\ldots\}$. *Define the random variable* $X_k$ *with density* $f_k$ *given by*

$$f_k(x):=\frac{q^k(x)f(x)}{\mathbb{E}\,q^k(X)}, \qquad \alpha<x<\omega,\ k=0,1,\ldots,n. \tag{5.8}$$

*Then,* $f_k \sim \mathrm{IP}(\mu_k;q_k)$ *with (the same) support* $J(X_k)=J(X)=(\alpha,\omega)$,

$$\mu_k=\frac{\mu+k\beta}{1-2k\delta} \quad \textit{and} \quad q_k(x)=\frac{q(x)}{1-2k\delta}, \qquad \alpha<x<\omega,\ k=0,1,\ldots,n.$$

*Moreover,* $X_0=X$, $X_1=X_0^*=X^*$, $X_2=X_1^*$ *and, in general,* $X_k=X_{k-1}^*$ *for* $k\in\{1,\ldots,n\}$.

*Proof.* For $k=0$ the assertion is obvious, while for $k=1$ (and thus, $n\ge 1$) it follows from Lemma 5.1 since $\mathbb{E}|X|^3<\infty$ and, by definition, $f_1=f^*$, $\mu_1=\mu^*$ and $q_1=q^*$. Assume now that the assertion has been proved for some $k\in\{1,\ldots,n-1\}$. Then

$$\mathbb{E}|X_k|^3=\frac{\mathbb{E}\,q^k(X)|X|^3}{\mathbb{E}\,q^k(X)}<\infty,$$

because $\mathbb{E}|X|^{2k+3}<\infty$ since $k\le n-1$. Therefore, we can apply Lemma 5.1 to the random variable $X_k\sim\mathrm{IP}(\mu_k;q_k)\equiv\mathrm{IP}(\mu_k;\delta_k,\beta_k,\gamma_k)$, obtaining $X_k^*\sim\mathrm{IP}(\mu_k^*;q_k^*)\equiv\mathrm{IP}(\mu_k^*;\delta_k^*,\beta_k^*,\gamma_k^*)$ where

$$\mu_k^*=\frac{\mu_k+\beta_k}{1-2\delta_k}=\frac{\frac{\mu+k\beta}{1-2k\delta}+\frac{\beta}{1-2k\delta}}{1-2\frac{\delta}{1-2k\delta}}=\frac{\mu+(k+1)\beta}{1-2(k+1)\delta}=\mu_{k+1}$$

and

$$q_k^*(x)=\frac{q_k(x)}{1-2\delta_k}=\frac{\frac{q(x)}{1-2k\delta}}{1-2\frac{\delta}{1-2k\delta}}=\frac{q(x)}{1-2(k+1)\delta}=q_{k+1}(x), \qquad \alpha<x<\omega.$$

On the other hand, since $\mathbb{E}\, q(X_k) = \frac{\mathbb{E}\, q^{k+1}(X)}{\mathbb{E}\, q^k(X)}$ and $X_k^* \sim f_k^*$ we get

$$f_k^*(x) = \frac{q_k(x) f_k(x)}{\mathbb{E}\, q_k(X_k)} = \frac{\frac{q(x)}{1-2k\delta} \frac{q^k(x)f(x)}{\mathbb{E}\, q^k(X)}}{\frac{\mathbb{E}\, q(X_k)}{1-2k\delta}} = \frac{\frac{q^{k+1}(x)f(x)}{\mathbb{E}\, q^k(X)}}{\frac{\mathbb{E}\, q^{k+1}(X)}{\mathbb{E}\, q^k(X)}}$$
$$= \frac{q^{k+1}(x)f(x)}{\mathbb{E}\, q^{k+1}(X)} = f_{k+1}(x), \qquad \alpha < x < \omega,$$

that is, $X_k^* = X_{k+1} \sim f_{k+1} \sim \mathrm{IP}(\mu_{k+1}; q_{k+1})$, and the proof is complete. ∎

Corollary 5.2. *If $X \sim \mathrm{IP}(\mu; q)$ and $\mathbb{E}\,|X|^{2n+2} < \infty$ (equivalently, if $\delta < 1/(2n+1)$), then for each $k \in \{0, 1, \ldots, n\}$,*

$$\sigma_k^2 := \operatorname{Var} X_k = \mathbb{E}\, q_k(X_k) = \frac{q\big(\frac{\mu+k\beta}{1-2k\delta}\big)}{1-(2k+1)\delta}, \tag{5.9}$$

*where $q_k(x) = \delta_k x^2 + \beta_k x + \gamma_k$ and $X_k$ are as in Theorem 5.2. In particular, if $\delta < 1$, then*

$$\sigma^2 := \operatorname{Var} X = \mathbb{E}\, q(X) = \frac{q(\mu)}{1-\delta}. \tag{5.10}$$

*Proof.* First observe that for any $k \in \{0, 1, \ldots, n\}$, $\mathbb{E}\,|X_k|^2 < \infty$ (and thus, $\mathbb{E}\, q^k(X_k) < \infty$) since $\delta_k = \delta/(1-2k\delta) < 1$ because $\delta < 1/(2n+1)$ $\le 1/(2k+1)$. Note that it suffices to show only (5.10). Indeed, since $X_k \sim \mathrm{IP}(\mu_k; q_k)$ it follows from (5.7) (applied to the random variable $X_k$ and to the function $g(x) = x$) that $\sigma_k^2 = \operatorname{Var} X_k = \mathbb{E}\, q_k(X_k)$. On the other hand, if it is shown that $\operatorname{Var} X = q(\mu)/(1-\delta)$ for any $X \sim \mathrm{IP}(\mu; q)$ with $\delta < 1$ then, by (5.10) applied to $X_k$, we get

$$\operatorname{Var} X_k = \frac{q_k(\mu_k)}{1-\delta_k}.$$

Since

$$\mu_k = \frac{\mu + k\beta}{1-2k\delta}, \qquad q_k(x) = \frac{q(x)}{1-2k\delta} \quad \text{and} \quad \delta_k = \frac{\delta}{1-2k\delta} < 1,$$

(5.10) yields the identity (5.9) as follows:

$$\mathbb{E}\, q_k(X_k) = \operatorname{Var} X_k = \frac{q_k(\mu_k)}{1-\delta_k} = \frac{\frac{q(\mu_k)}{1-2k\delta}}{1-\frac{\delta}{1-2k\delta}} = \frac{q(\mu_k)}{1-(2k+1)\delta}$$
$$= \frac{q\big(\frac{\mu+k\beta}{1-2k\delta}\big)}{1-(2k+1)\delta}.$$

It remains to verify that $\operatorname{Var} X = \sigma^2 = q(\mu)/(1-\delta)$ whenever $X \sim \mathrm{IP}(\mu; q)$ and $\delta < 1$. To this end, write

$$q(X) = q(\mu) + q'(\mu)(X-\mu) + \delta(X-\mu)^2$$

and take expectations to get $\sigma^2 = q(\mu) + \delta\sigma^2$, which is equivalent to (5.10). ∎

Corollary 5.3. *If* $X \sim \mathrm{IP}(\mu;q)$ *and* $\mathbb{E}\,|X|^{2n} < \infty$ *for some* $n \ge 1$ *(i.e.* $\delta < 1/(2n-1)$*), then for each* $k \in \{1,\ldots,n\}$,

$$A_k = A_k(\mu;q) := \mathbb{E}\,q^k(X) = \frac{\prod_{j=0}^{k-1}(1-2j\delta)}{\prod_{j=0}^{k-1}(1-(2j+1)\delta)} \prod_{j=0}^{k-1} q\Big(\frac{\mu+j\beta}{1-2j\delta}\Big). \tag{5.11}$$

*Proof.* Observe that

$$(1-2j\delta)\,\mathbb{E}\,q_j(X_j) = \mathbb{E}\,q(X_j) = \frac{A_{j+1}}{A_j}, \qquad j=0,1,\ldots,n-1,$$

where $A_0 := 1$, $q_0 = q$, $X_0 = X$. Multiplying these relations for $j = 0,1,\ldots,k-1$ and using (5.9) we get (5.11). ∎

Remark 5.1. (a) It is important to note that the identity (5.3) enables a convenient calculation of the Fourier coefficients of any smooth enough function $g$ with $\mathrm{Var}\,g(X) < \infty$ (i.e., $g \in L^2(\mathbb{R},X)$). Indeed, if $X \sim \mathrm{IP}(\mu;\delta,\beta,\gamma) \equiv \mathrm{IP}(\mu;q)$ and $\mathbb{E}\,|X|^{2n} < \infty$, then the Fourier coefficients $\alpha_k = \mathbb{E}\,\phi_k(X)g(X)$ are given by $\alpha_0 = \mathbb{E}\,g(X)$ and

$$\alpha_k = \frac{\mathbb{E}\,q^k(X)g^{(k)}(X)}{(k!c_k(\delta)A_k(\mu;q))^{1/2}}, \qquad k=1,\ldots,n,$$

where $c_k(\delta)$ and $A_k(\mu;q)$ are given by (5.2) and (5.11), respectively, provided that $g$ is smooth enough so that $\mathbb{E}\,q^k(X)|g^{(k)}(X)| < \infty$ for $k \in \{1,2,\ldots,n\}$.

(b) Obviously, if $X \sim \mathrm{IP}(\mu;\delta,\beta,\gamma)$ and $\delta \le 0$ (i.e. if $X$ is of Normal, Gamma or Beta-type) then $\mathbb{E}\,|X|^n < \infty$ for all $n$. Moreover, since there exists an $\epsilon > 0$ such that $\mathbb{E}\,e^{tX} < \infty$ for $|t| < \epsilon$, it follows that the corresponding polynomials $\{\phi_k\}_{k=0}^{\infty}$, given by (5.5), form a complete orthonormal system in $L^2(\mathbb{R};X)$; see, e.g., Riesz (1923); Berg and Christensen (1981); Afendras et al. (2011). Therefore, for smooth enough $g$ with $\mathrm{Var}\,g(X) < \infty$ and $\mathbb{E}\,q^k(X)|g^{(k)}(X)| < \infty$ for all $k \ge 1$, the Fourier coefficients are given by

$$\alpha_k = \mathbb{E}\,\phi_k(X)g(X) = \frac{\mathbb{E}\,q^k(X)g^{(k)}(X)}{(k!c_k(\delta)A_k(\mu;q))^{1/2}}, \qquad k=0,1,2,\ldots,$$

and the variance of $g$ can be calculated by *Parseval's identity* (see Afendras et al., 2011, Theorem 5.1, pp. 522–523):

$$\mathrm{Var}\,g(X) = \sum_{k=1}^{\infty} \frac{\mathbb{E}^2\,q^k(X)g^{(k)}(X)}{k!c_k(\delta)A_k(\mu;q)}. \tag{5.12}$$

Furthermore, the completeness of the Rodrigues polynomials (when $X \sim \mathrm{IP}(\mu;\delta,\beta,\gamma)$ and $\delta \le 0$) enables one to write (Afendras et al., 2011, Theorem 5.2, p. 523)

$$\text{(5.13)} \qquad \operatorname{Cov}[g_1(X), g_2(X)] = \sum_{k=1}^{\infty} \frac{\mathbb{E}[q^k(X) g_1^{(k)}(X)]\,\mathbb{E}[q^k(X) g_2^{(k)}(X)]}{k!\,c_k(\delta) A_k(\mu; q)},$$

provided that for $i = 1, 2$, $g_i \in L^2(\mathbb{R}, X)$ and $\mathbb{E}\, q^k(X)|g_i^{(k)}(X)| < \infty$ for all $k \geq 1$. The important thing in (5.12) and (5.13) is that we do not need explicit forms for the polynomials; in view of (5.2) and (5.11), everything is calculated from the four numbers $(\mu; \delta, \beta, \gamma)$ and the derivatives of $g$ or $g_i$ $(i = 1, 2)$. In particular, for the first three types of Table 2.1, (5.12) yields the formulae

$$\operatorname{Var} g(X) = \sum_{k=1}^{\infty} \frac{\sigma^{2k}}{k!}\, \mathbb{E}^2\, g^{(k)}(X) \qquad \text{if } X \sim N(\mu, \sigma^2);$$

$$\operatorname{Var} g(X) = \sum_{k=1}^{\infty} \frac{\Gamma(a)}{k!\,\Gamma(a+k)}\, \mathbb{E}^2\, X^k g^{(k)}(X) \qquad \text{if } X \sim \Gamma(a, \lambda);$$

$$\operatorname{Var} g(X) = \sum_{k=1}^{\infty} \frac{(a+b+2k-1)\Gamma(a)\Gamma(b)\Gamma(a+b+k-1)}{k!\,\Gamma(a+b)\Gamma(a+k)\Gamma(b+k)}\, \mathbb{E}^2\, X^k (1-X)^k g^{(k)}(X) \qquad \text{if } X \sim B(a, b).$$

Turn now to the orthogonal polynomial system $\{P_k : k = 0, 1, \ldots, n\}$, of (5.1), obtained for a random variable $X \sim \mathrm{IP}(\mu; \delta, \beta, \gamma)$ with support $J(X) = (\alpha, \omega)$ and $\mathbb{E}\,|X|^{2n} < \infty$ for some $n \geq 2$, i.e. with $\delta < 1/(2n-1)$. By Lemma 5.1 the random variable $X^* = X_1 \sim \mathrm{IP}(\mu_1; q_1) \equiv \mathrm{IP}(\mu_1; \delta_1, \beta_1, \gamma_1)$ with

$$\mu_1 = \frac{\mu + \beta}{1 - 2\delta} \quad \text{and} \quad q_1(x) = \frac{q(x)}{1 - 2\delta}$$

is supported by $(\alpha, \omega)$. Since $\delta < 1/(2n-1)$ is equivalent to $\delta_1 = \delta/(1-2\delta) < 1/(2n-3)$, we conclude that $\mathbb{E}\,|X_1|^{2n-2} < \infty$, in particular $\operatorname{Var} X_1 < \infty$. Therefore, we can define the orthogonal polynomial system $\{P_{k,1} : k = 0, 1, \ldots, n-1\}$ by applying (5.1) to the density $f_1$ and to the quadratic polynomial $q_1$ of $X_1$, that is (recall that $f_1(x) = q(x) f(x)/\mathbb{E}\, q(X)$),

$$\text{(5.14)} \qquad P_{k,1}(x) := \frac{(-1)^k}{f_1(x)} \frac{d^k}{dx^k}[q_1^k(x) f_1(x)] = \frac{(-1)^k}{(1-2\delta)^k q(x) f(x)} \frac{d^k}{dx^k}[q^{k+1}(x) f(x)],$$
$$\alpha < x < \omega,\ k = 0, 1, \ldots, n-1.$$

Clearly the system $\{P_{k,1} : k = 0, 1, \ldots, n-1\}$ is orthogonal with respect to $X_1$, but the important observation is that we can also obtain it by differentiating the polynomials $P_k$ (which are orthogonal with respect to $X$). In fact, the following lemma holds.

LEMMA 5.2. *If $X \sim \mathrm{IP}(\mu;q)$ and $\mathbb{E}\,|X|^{2n} < \infty$ for some $n \ge 1$, then the polynomials $P_k$ of (5.1) and $P_{k,1}$ of (5.14) are related through*

$$P'_{k+1}(x) = C_k(\delta)P_{k,1}(x), \qquad k = 0,1,\dots,n-1, \tag{5.15}$$

*where $C_k(\delta) := (k+1)(1-k\delta)(1-2\delta)^k$.*

*Proof.* First we show that the polynomials $P'_{k+1}$ are orthogonal with respect to $X_1$. Indeed, $\deg(P'_{k+1}) = k$ (for $k = 0,1,\dots,n-1$) and for $k,m \in \{0,1,\dots,n-1\}$ with $k < m$, we have

$$\begin{aligned}\mathbb{E}\,P'_{k+1}(X_1)P'_{m+1}(X_1) &= \frac{1}{\sigma^2}\int_\alpha^\omega P'_{m+1}(x)P'_{k+1}(x)q(x)f(x)\,dx \\ &= \frac{1}{\sigma^2}\Big\{P_{m+1}(x)P'_{k+1}(x)q(x)f(x)\Big|_\alpha^\omega - \int_\alpha^\omega P_{m+1}(x)[P'_{k+1}(x)q(x)f(x)]'\,dx\Big\}.\end{aligned}$$

Now observe that, in view of Lemma 2.1,

$$P_{m+1}(x)P'_{k+1}(x)q(x)f(x)\Big|_\alpha^\omega = 0,$$

because $P_{m+1}P'_{k+1}$ is a polynomial of degree $m+k+1 \le 2n-2$ and $\mathbb{E}\,|X|^{2n} < \infty$. Moreover,

$$[P'_{k+1}(x)q(x)f(x)]' = P''_{k+1}(x)q(x)f(x) + P'_{k+1}(x)(\mu-x)f(x) = H_{k+1}(x)f(x),$$

where $H_{k+1}(x) = P''_{k+1}(x)q(x) + (\mu - x)P'_{k+1}(x)$ is a polynomial in $x$ of degree at most $k+1 < m+1$. Therefore,

$$\mathbb{E}\,P'_{k+1}(X_1)P'_{m+1}(X_1) = -\frac{1}{\sigma^2}\,\mathbb{E}\,P_{m+1}(X)H_{k+1}(X) = 0,$$

since $P_{m+1}$ is orthogonal (with respect to $X$) to any polynomial of degree lower than $m+1$. Note that the same orthogonality conditions are also valid for $\{P_{k,1}\}_{k=0}^{n-1}$, that is,

$$\mathbb{E}\,P_{k,1}(X_1)P_{m,1}(X_1) = 0 \quad \text{ for } k,m \in \{0,1,\dots,n-1\} \text{ with } k \ne m.$$

Since $\deg(P'_{k+1}) = \deg(P_{k,1}) = k$, $k = 0,1,\dots,n-1$, the uniqueness of the orthogonal polynomial system implies that there exist constants $C_k \ne 0$ such that $P'_{k+1}(x) = C_kP_{k,1}(x)$. Equating the leading coefficients, we obtain

$$\begin{aligned}C_k &= \frac{\mathrm{lead}(P'_{k+1})}{\mathrm{lead}(P_{k,1})} = \frac{(k+1)\,\mathrm{lead}(P_{k+1})}{\mathrm{lead}(P_{k,1})} = \frac{(k+1)c_{k+1}(\delta)}{c_k(\delta_1)} \\ &= (k+1)(1-k\delta)(1-2\delta)^k. \ \blacksquare\end{aligned}$$

REMARK 5.2. We note that the recurrence (5.15) is contained in Beale (1937, eq. (2), p. 207). Actually, Beale's recurrence (which is stated in a much different notation) is valid for the polynomials $h_k$ of (4.2) and for all $k \ge 0$.

Hence, orthogonality is not required. For more details see arXiv:1205.2903v2, pp. 39–40.

Applying Lemma 5.2 inductively it is easy to verify the following result.

THEOREM 5.3. *If $X \sim \mathrm{IP}(\mu;\delta,\beta,\gamma)$ with support $J(X) = (\alpha,\omega)$ and $\mathbb{E}\,|X|^{2n} < \infty$ for some $n \ge 1$ (i.e. $\delta < \frac{1}{2n-1}$) then*

$$P_{k+m}^{(m)}(x) = C_k^{(m)}(\delta) P_{k,m}(x), \qquad m = 1,\ldots,n,\ k = 0,1,\ldots,n-m, \tag{5.16}$$

*where*

$$C_k^{(m)}(\delta) := \frac{(k+m)!}{k!}(1-2m\delta)^k \prod_{j=k+m-1}^{k+2m-2} (1-j\delta).$$

*Here, $P_k$ are the polynomials given by (5.1) associated with $f$, and $P_{k,m}$ are the corresponding Rodrigues polynomials of (5.1), associated with the density $f_m(x) = \frac{q^m(x)f(x)}{\mathbb{E}\,q^m(X)}$, $\alpha < x < \omega$, of the random variable $X_m \sim \mathrm{IP}(\mu;\delta,\beta,\gamma)(\mu_m;q_m)$ of Theorem 5.2, i.e.,*

$$\begin{aligned} P_{k,m}(x) &:= \frac{(-1)^k}{f_m(x)} \frac{d^k}{dx^k}[q_m^k(x) f_m(x)] \\ &= \frac{(-1)^k}{(1-2m\delta)^k q^m(x) f(x)} \frac{d^k}{dx^k}[q^{k+m}(x) f(x)], \\ &\qquad\qquad \alpha < x < \omega,\ k = 0,1,\ldots,n-m. \end{aligned}$$

REMARK 5.3. (a) An alternative calculation of the constant $C_k = C_k^{(m)}(\delta)$ can be given as follows. Lemma 5.2 guarantees that $P_{k+m}^{(m)}(x) = C_k P_{k,m}(x)$ for some $C_k$. Arguing as in the proof of Lemma 5.2 we see that $C_k$ can be derived from the corresponding leading coefficients.

(b) We note that the recurrence (5.16) is also contained in Beale (1937, eq. (4), p. 207), although it is stated in a quite different notation there. Specifically, it can be shown that (5.16) holds for all $k \in \{0,1,\ldots\}$; see arXiv:1205.2903v2, pp. 40–42, for a more detailed discussion.

(c) Krall (1936, 1941) characterizes the Pearson system by the fact that the derivatives of orthogonal polynomials are orthogonal polynomials.

We can now adapt the preceding results to the corresponding orthonormal polynomial systems. Notice that the following corollary is our main result regarding Fourier expansions within the Pearson family and, to our knowledge, it is not stated elsewhere in the present simple, unified, explicit form.

COROLLARY 5.4. *Let $X \sim \mathrm{IP}(\mu;\delta,\beta,\gamma) \equiv \mathrm{IP}(\mu;q)$ with support $(\alpha,\omega)$, and assume that $\mathbb{E}\,|X|^{2n} < \infty$ for some fixed $n \ge 1$ (equivalently, $\delta < 1/(2n-1)$). Let $\{\phi_k\}_{k=0}^n$ be the orthonormal polynomials associated with $X$*

*(with* $\mathrm{lead}(\phi_k) > 0$ *for all* $k$*; see* (5.5), (5.6)), *fix* $m \in \{0,1,\ldots,n\}$, *and consider the corresponding orthonormal polynomials* $\{\phi_{k,m}\}_{k=0}^{n-m}$, *with* $\mathrm{lead}(\phi_{k,m}) > 0$, *associated with*

$$X_m \sim f_m(x) = \frac{q^m(x) f(x)}{\mathbb{E}\, q^m(X)}, \qquad \alpha < x < \omega.$$

*Then there exist constants* $\nu_k^{(m)} = \nu_k^{(m)}(\mu; q) > 0$ *such that*

$$\phi_{k+m}^{(m)}(x) = \nu_k^{(m)} \phi_{k,m}(x), \qquad \alpha < x < \omega,\ k = 0,1,\ldots,n-m.$$

*Specifically, the constants* $\nu_k^{(m)}$ *have the explicit form*

$$\nu_k^{(m)} = \nu_k^{(m)}(\mu; q) := \left\{ \frac{\frac{(k+m)!}{k!} \prod_{j=k+m-1}^{k+2m-2} (1 - j\delta)}{A_m(\mu; q)} \right\}^{1/2},$$

*where* $A_m(\mu; q) = \mathbb{E}\, q^m(X)$ *is given by* (5.11). *In particular, setting* $\sigma^2 = \mathrm{Var}\, X$ *we have*

$$\begin{aligned}\phi'_{k+1}(x) &= \frac{\sqrt{(k+1)(1-k\delta)}}{\sigma} \phi_{k,1}(x) \\ &= \sqrt{\frac{(k+1)(1-\delta)(1-k\delta)}{q(\mu)}} \phi_{k,1}(x), \qquad k = 0,1,\ldots,n-1.\end{aligned}$$

*Proof.* Observe that

$$\phi_{k+m}(x) = \frac{P_{k+m}(x)}{\sqrt{\mathbb{E}\,|P_{k+m}(X)|^2}} \quad \text{and} \quad \phi_{k,m}(x) = \frac{P_{k,m}(x)}{\sqrt{\mathbb{E}\,|P_{k,m}(X_m)|^2}}, \qquad \alpha < x < \omega,$$

where $P_{k+m}$ and $P_{k,m}$ are as in Theorem 5.3. Since

$$P_{k+m}^{(m)}(x) = C_k^{(m)}(\delta) P_{k,m}(x), \qquad \alpha < x < \omega,$$

we conclude that there exists a constant $\nu_k^{(m)}$ such that $\phi_{k+m}^{(m)}(x) = \nu_k^{(m)} \phi_{k,m}(x)$. Hence,

$$\begin{aligned}\nu_k^{(m)} &= \frac{\mathrm{lead}(\phi_{k+m}^{(m)})}{\mathrm{lead}(\phi_{k,m})} = \frac{\frac{(k+m)!}{k!}\,\mathrm{lead}(\phi_{k+m})}{\mathrm{lead}(\phi_{k,m})} = \frac{\frac{(k+m)!}{k!} \frac{\mathrm{lead}(P_{k+m})}{\sqrt{\mathbb{E}\,|P_{k+m}(X)|^2}}}{\frac{\mathrm{lead}(P_{k,m})}{\sqrt{\mathbb{E}\,|P_{k,m}(X_m)|^2}}} \\ &= \frac{(k+m)!\,\mathrm{lead}(P_{k+m}) \sqrt{\mathbb{E}\,|P_{k,m}(X_m)|^2}}{k!\,\mathrm{lead}(P_{k,m}) \sqrt{\mathbb{E}\,|P_{k+m}(X)|^2}} \\ &= \frac{(k+m)! c_{k+m}(\delta) \sqrt{\mathbb{E}\,|P_{k,m}(X_m)|^2}}{k! c_k(\delta_m) \sqrt{\mathbb{E}\,|P_{k+m}(X)|^2}},\end{aligned}$$

where, by (5.2), $c_{k+m}(\delta)=\prod_{j=k+m-1}^{2k+2m-2}(1-j\delta)$ and

$$
\begin{aligned}
c_k(\delta_m) &= \prod_{j=k-1}^{2k-2}(1-j\delta_m) = \prod_{j=k-1}^{2k-2}\left(1-j\frac{\delta}{1-2m\delta}\right)\\
&= \frac{\prod_{j=k-1}^{2k-2}(1-(2m+j)\delta)}{(1-2m\delta)^k} = \frac{\prod_{j=k+2m-1}^{2k+2m-2}(1-j\delta)}{(1-2m\delta)^k}.
\end{aligned}
$$

From (5.4) we see that $\mathbb{E}\,|P_{k+m}(X)|^2=(k+m)!c_{k+m}(\delta)\,\mathbb{E}\,q^{k+m}(X)$ and

$$
\begin{aligned}
\mathbb{E}\,|P_{k,m}(X_m)|^2 &= k!c_k(\delta_m)\,\mathbb{E}\,q_m^k(X_m)\\
&= k!c_k(\delta_m)\frac{\mathbb{E}\,q_m^k(X)q^m(X)}{\mathbb{E}\,q^m(X)} = \frac{k!c_k(\delta_m)\,\mathbb{E}\,q^{k+m}(X)}{(1-2m\delta)^k\,\mathbb{E}\,q^m(X)}.
\end{aligned}
$$

Combining the preceding relations we obtain

$$
\begin{aligned}
\nu_k^{(m)} &= \frac{(k+m)!c_{k+m}(\delta)\sqrt{\mathbb{E}\,|P_{k,m}(X_m)|^2}}{k!c_k(\delta_m)\sqrt{\mathbb{E}\,|P_{k+m}(X)|^2}}\\
&= \frac{(k+m)!c_{k+m}(\delta)\sqrt{\frac{k!c_k(\delta_m)\,\mathbb{E}\,q^{k+m}(X)}{(1-2m\delta)^k\,\mathbb{E}\,q^m(X)}}}{k!c_k(\delta_m)\sqrt{(k+m)!c_{k+m}(\delta)\,\mathbb{E}\,q^{k+m}(X)}}\\
&= \frac{(k+m)!c_{k+m}(\delta)\sqrt{k!c_k(\delta_m)\,\mathbb{E}\,q^{k+m}(X)}}{k!c_k(\delta_m)\sqrt{(k+m)!c_{k+m}(\delta)\,\mathbb{E}\,q^{k+m}(X)(1-2m\delta)^k\,\mathbb{E}\,q^m(X)}}\\
&= \frac{\sqrt{(k+m)!c_{k+m}(\delta)}}{\sqrt{k!c_k(\delta_m)(1-2m\delta)^k\,\mathbb{E}\,q^m(X)}}\\
&= \sqrt{\frac{(k+m)!}{k!\,\mathbb{E}\,q^m(X)}}\sqrt{\frac{c_{k+m}(\delta)}{c_k(\delta_m)(1-2m\delta)^k}}\\
&= \sqrt{\frac{(k+m)!}{k!\,\mathbb{E}\,q^m(X)}}\sqrt{\frac{\prod_{j=k+m-1}^{2k+2m-2}(1-j\delta)}{\frac{\prod_{j=k+2m-1}^{2k+2m-2}(1-j\delta)}{(1-2m\delta)^k}(1-2m\delta)^k}}\\
&= \sqrt{\frac{(k+m)!}{k!\,\mathbb{E}\,q^m(X)}\prod_{j=k+m-1}^{k+2m-2}(1-j\delta)},
\end{aligned}
$$

and the proof is complete. ∎

**Acknowledgements.** We would like to thank M.C. Jones, H. Papageorgiou, M. Akritas and an anonymous referee for a number of suggestions and helpful comments that improved the presentation.

This research has been co-financed by the European Union (European Social Fund–ESF) and Greek national funds through the Operational Program “Education and Lifelong Learning” of the National Strategic Reference

Framework (NSRF)—Research Funding Program: ARISTEIA, Grant No. 4357. Also, this work is partially supported by the University of Athens Research Grant 70/4/5637 and by internal funds, Department of Biostatistics, SUNY Buffalo.

## References

G. Afendras and N. Papadatos (2011), *On matrix variance inequalities*, J. Statist. Plann. Inference 141, 3628–3631.

G. Afendras, N. Papadatos and V. Papathanasiou (2007), *The discrete Mohr and Noll inequality with applications to variance bounds*, Sankhyā 69, 162–189.

G. Afendras, N. Papadatos and V. Papathanasiou (2011), *An extended Stein-type covariance identity for the Pearson family, with applications to lower variance bounds*, Bernoulli 17, 507–529.

F. S. Beale (1937), *On the polynomials related to Pearson's differential equation*, Ann. Math. Statist. 8, 206–223.

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

C. Berg and J. P. R. Christensen (1981), *Density questions in the classical theory of moments*, Ann. Inst. Fourier (Grenoble) 31, no. 3, 99–114.

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

C. C. Craig (1936), *A new exposition and chart for the Pearson system of frequency curves*, Ann. Math. Statist. 7, 16–28.

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

L. Goldstein and G. Reinert (1997), *Stein's method and the zero-bias transformation with application to simple random sampling*, Ann. Appl. Probab. 7, 935–952.

L. Goldstein and G. Reinert (2005), *Distributional transformations, orthogonal polynomials and Stein characterizations*, J. Theoret. Probab. 18, 237–260.

E. H. Hildebrandt (1931), *Systems of polynomials connected with the Charlier expansions and the Pearson differential and difference equations*, Ann. Math. Statist. 2, 379–439.

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

R. W. Johnson (1993), *A note on variance bounds for a function of a Pearson variate*, Statist. Decisions 11, 273–278.

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

H. L. Krall (1936), *On derivatives of orthogonal polynomials*, Bull. Amer. Math. Soc. 42, 423–428.

H. L. Krall (1941), *On derivatives of orthogonal polynomials, II*, Bull. Amer. Math. Soc. 47, 261–264.

Y. Nagahara (1999), *The PDF and CF of Pearson type IV distributions and the ML estimation of the parameters*, Statist. Probab. Lett. 43, 251–264.

N. Nielsen (2005), *Die Gammafunktion*, Chelsea, New York.

J. K. Ord (1967), *On a system of discrete distributions*, Biometrika 54, 649–656.

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

N. Papadatos and V. Papathanasiou (2001), *Unified variance bounds and a Stein-type identity*, in: Probability and Statistical Models with Applications, Ch. A. Charalambides et al. (eds.), Chapman & Hall/CRC, New York, 87–100.

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

K. Pearson (1895), *Contributions to the mathematical theory of evolution*, Philos. Trans. Roy. Soc. London Ser. A 186, 343–414.

K. Pearson (1924), *On the mean error of frequency distributions*, Biometrika 16, 198–200.

M. Riesz (1923), *Sur le problème des moments et le théorème de Parseval*, Acta Litt. Acad. Sci. (Szeged) 1, 209–225.

G. Afendras
Department of Biostatistics
University at Buffalo
Buffalo, NY 14214, U.S.A.
E-mail: gafendra@buffalo.edu

N. Papadatos (corresponding author)
Department of Mathematics
Section of Statistics and O.R.
University of Athens
Panepistemiopolis
157 84 Athens, Greece
E-mail: npapadat@math.uoa.gr
url: users.uoa.gr/∼npapadat/

*Received on 29.4.2013;*
*revised version on 30.6.2015* (2182)