DEDICATED TO PROFESSOR V. PAPATHANASIOU

Integrated Pearson family and orthogonality of the Rodrigues polynomials: A review including new results and an alternative classification of the Pearson system*

G. Afendras† and N. Papadatos‡

Department of Mathematics, Section of Statistics and O.R., University of Athens, Panepistemiopolis, 157 84 Athens, Greece.

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 *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 to decide whether a given Pearson density belongs to the integrated system, or not. Recurrences between the derivatives of the corresponding orthonormal polynomial systems are also given.

MSC: Primary 62E15, 60E05; Secondary 62-00.

Key words and phrases: Integrated Pearson Family of distributions; Derivatives of orthogonal polynomials; Rodrigues polynomials.

1 Introduction

Karl Pearson (1895), in the context of fitting curves to real data, 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 Pearson distributions. The original classification given by Pearson contains twelve types (I–XII), although this 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

*Work partially supported by the University of Athens Research Grant 70/4/5637

†e-mail: g_afendras@math.uoa.gr

‡*Corresponding author.* e-mail: npapadat@math.uoa.gr, url: users.uoa.gr/∼npapadat/

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, the polynomials that 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 to 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 (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

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

This fact will be denoted by $X \sim \mathrm{IP}(\mu; q)$ or $f \sim \mathrm{IP}(\mu; q)$ or, more explicitly, 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 that are worth to be highlighted. The integrated Pearson system satisfies many interesting properties, like recurrences on moments and on Rodrigues polynomials, covariance identities, closeness of each type under particularly useful transformations etc.; such properties are by far more complicated (if they are, at all, true) for distributions outside the Integrated Pearson system. 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. In other words, the Rodrigues polynomials and, consequently, the ordinary Pearson densities, 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 enough, there exist 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 given by Corollary 5.4. It provides an explicit relation (in terms of μ and q) between the m-th derivative of an orthonormal polynomial of degree $k \geqslant m$ 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 $f_m \propto q^m f$.

In the sequel and elsewhere in this article, $X \sim \mathrm{IP}(\mu;\delta,\beta,\gamma)$ means that X has finite mean μ, and that X admits a density f (w.r.t. Lebesgue measure on $\mathbb{R}$) such that (1.1) is fulfilled. Define the open (bounded or unbounded) interval

$$J = J(X) := (\operatorname{ess\,inf}(X), \operatorname{ess\,sup}(X)). \tag{1.2}$$

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 (α, ω); thus, $f(\rho)$ may assume any value in this case. However, in order to be specific, we can redefine $f(\rho) = 0$ at such points ρ, if any, without any loss of generality. Therefore, we shall use this convention through the whole article without any further reference to it.

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 $\delta \leqslant 0$) 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 up to order $1 + \frac{1}{\delta} - \varepsilon$ (for any $\varepsilon > 0$) 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 $J = (\alpha, \omega) = (\operatorname{ess\,inf}(X), \operatorname{ess\,sup}(X))$. 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 $x \leqslant \alpha$ (if any) and for all $x \geqslant \omega$ (if any). Also, it must be strictly positive for $x \in J$. Indeed, if $x \in (\mu, \omega)$ then $q(x)f(x) = \int_x^{\infty}(t-\mu)f(t)\mathrm{d}t \geqslant (x-\mu)(1-F(x)) > 0$; if $x \in (\alpha, \mu)$ then $q(x)f(x) = \int_{-\infty}^{x}(\mu - t)f(t)\mathrm{d}t \geqslant (\mu - x)F(x) > 0$; 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, $f : (\alpha, \omega) \to (0, \infty)$ is $C^{\infty}(J)$. Indeed, writing $p_1(x) = \mu - x - q'(x)$ (a polynomial of degree at most one) we see from (1.1) that $f : J \to (0, \infty)$ is continuous and 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 left neighborhood of ω, it follows that $q(\omega) \geqslant 0$. Assume now that $q(\omega) > 0$ and define

$$\lambda_1 := \inf_{x \in [\mu, \omega]} \{q(x)\} > 0, \quad \lambda_2 := \sup_{x \in [\mu, \omega]} |\mu - x - q'(x)| < \infty.$$

Then, for all $x \in [\mu, \omega)$,

$$\left| \int_{\mu}^{x} \frac{\mu - t - q'(t)}{q(t)} \mathrm{d}t \right| \leqslant \int_{\mu}^{x} \frac{|\mu - t - q'(t)|}{q(t)} \mathrm{d}t \leqslant \int_{\mu}^{\omega} \frac{|\mu - t - q'(t)|}{q(t)} \mathrm{d}t \leqslant (\omega - \mu)\frac{\lambda_2}{\lambda_1} < \infty.$$

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

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

we have

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

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

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

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 mean $\widetilde{\mu} = -\mu$ and support $J(\widetilde{X}) = (\widetilde{\alpha}, \widetilde{\omega}) = (-\omega, -\alpha)$. According to (vi), its density $\widetilde{f}$ satisfies (1.1) with quadratic polynomial $\widetilde{q}(x) = q(-x)$. 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 to 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 $\widetilde{J} \subseteq S^+(q)$ it is true that either $\widetilde{J} \subseteq J$ or $\widetilde{J} \cap J = \varnothing$.

In other words, the support J of X can be taken to be an open interval that coincides to some 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 connected 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, however, $\mathbb{E}X = \mu \in J$, 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 $\mu \in \{x : q(x) > 0\}$; thus, $\{x : q(x) > 0\} \neq \varnothing$ whenever q is admissible. In the sequel we shall show that for any admissible choice of q and for 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 $\delta = 0$

We have to further distinguish between the cases $\beta = 0$ 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 get

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

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 a number $\mu \in \{x : q(x) > 0\}$; that is, $q(\mu) = \beta\mu + \gamma > 0$. According to Proposition 2.1(vi) we may further assume that $\beta > 0$, $\gamma = 0$ and $\mu > 0$; otherwise, it suffices to consider the random variable $\widetilde{X} = \frac{\beta}{|\beta|}(X + \frac{\gamma}{\beta})$ with $\widetilde{q}(x) = |\beta| x$ and $\mathbb{E}\widetilde{X} = \widetilde{\mu} = \frac{\beta}{|\beta|}(\mu + \frac{\gamma}{\beta}) = \frac{q(\mu)}{|\beta|} > 0$ since $q(\mu) > 0$. Now, since $q(x) = \beta x$ with $\beta > 0$ we must have $J(X) = (0, \infty)$. Fixing $\mu > 0$ and solving the differential equation (2.1) we get

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

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

2.2 The case $\delta < 0$

Since $\delta < 0$ and $\{x : q(x) > 0\}$ must contain some interval it follows that the discriminant $\beta^2 - 4\delta\gamma$ of q must be strictly positive. If $\rho_1 < \rho_2$ are the real roots of q we can 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 exist 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 $0 < \mu < 1$; the general case is reduced 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 get

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

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

2.3 The case $\delta > 0$

We have to further distinguish between the cases 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 any 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 $c = 0$; otherwise we can consider the random variable $\widetilde{X} = X - c$. Fixing $\mu \in \mathbb{R}$, $q(x) = \delta x^2 + \theta$ and solving (2.1) one finds that

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

The normalizing constant $C = C_\mu(\delta, \theta)$ can be calculated explicitly when $\mu = 0$:

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

Therefore, the quadratic polynomial $q(x) = \delta(x-c)^2 + \theta$ with $\delta > 0$ and $\theta > 0$ corresponds to *Student-type* distributions centered at c, provided that $\mu = c$; otherwise, i.e., when $\mu \neq c$, it corresponds to some asymmetric, say *skew Student-type*, distributions.

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 (ρ, ∞), according to $\mu < \rho$ or $\mu > \rho$, respectively. Without loss of generality we may assume that $q(x) = \delta x^2$ with $\delta > 0$ and $\mu > 0$; otherwise, 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 eq. (2.1) we get the solution

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

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 *Reciprocal Gamma-type* distributions.

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 (ρ_2, ∞), according to $\mu < \rho_1$ or $\mu > \rho_2$, respectively. By considering the random variable $\widetilde{X} = -(X - \rho_1)$ when $\mu < \rho_1$ and the random variable $\widetilde{X} = X - \rho_2$ 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 $\delta > 0$ and $\theta = \rho_2 - \rho_1 > 0$. Thus, there is no loss of generality in assuming $\mu > 0$, $J(X) = (0, \infty)$ and $q(x) = \delta x(x + \theta)$ with $\delta > 0$ and $\theta > 0$. Then, (2.1) yields

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

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

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

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

type / **usual notation**	**density $f(x)$**	**support**	**$q(x)$**	**parameters**	**mean μ**	**classification rule**
1. Normal-type $X \sim N(\mu,\sigma^2)$	$\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$\mathbb{R}$	σ^2	$\gamma=\sigma^2>0$	$\mu\in\mathbb{R}$	$\delta=\beta=0$
2. Gamma-type $X\sim\Gamma(a,\lambda)$	$\frac{\lambda^a}{\Gamma(a)}x^{a-1}e^{-\lambda x}$	$(0,+\infty)$	$\dfrac{x}{\lambda}$	$a,\lambda>0$	$\dfrac{a}{\lambda}>0$	$\delta=0,\ \beta\neq 0$
3. Beta-type $X\sim B(a,b)$	$\frac{x^{a-1}(1-x)^{b-1}}{B(a,b)}$	$(0,1)$	$\dfrac{x(1-x)}{a+b}$	$a,b>0$	$\dfrac{a}{a+b}>0$	$\delta=\dfrac{-1}{a+b}<0$
4. Student-type [1]	$\dfrac{C\exp\left(\frac{\mu\tan^{-1}(x\sqrt{\delta/\gamma})}{\sqrt{\delta\gamma}}\right)}{(\delta x^2+\gamma)^{1+\frac{1}{2\delta}}}$ [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	$\frac{\lambda^a}{\Gamma(a)}x^{-a-1}e^{-\frac{\lambda}{x}}$	$(0,+\infty)$	$\dfrac{x^2}{a-1}$	$a>1,\ \lambda>0$	$\dfrac{\lambda}{a-1}>0$	$\delta=\frac{1}{a-1}>0$ $\beta^2=4\delta\gamma$ $\frac{1}{X}\sim\Gamma(a,\lambda)$
6. Snedecor-type [3]	$\frac{\theta^a}{B(a,b)}x^{b-1}(x+\theta)^{-a-b}$	$(0,+\infty)$	$\dfrac{x(x+\theta)}{a-1}$	$a>1$, $b,\ \theta>0$	$\dfrac{b\theta}{a-1}>0$	$\delta=\frac{1}{a-1}>0$ $\beta^2>4\delta\gamma$ $\frac{\theta}{X+\theta}\sim B(a,b)$

* 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 $\widetilde{X}=c_1X+c_2$ is contained in the table.

[1] For $n>1$ and if $\mu=0$ and $\delta=\frac{1}{n-1}=\frac{\gamma}{n}$ then $X\sim t_n$.

[2] $C=C_\mu(\delta,\gamma)>0$, with $C_0(\delta,\gamma)=\Gamma\left(1+\frac{1}{2\delta}\right)\sqrt{\delta\gamma^{1+\frac{1}{\delta}}}\Big/\Gamma\left(\frac{1}{2}+\frac{1}{2\delta}\right)\sqrt{\pi}$.

[3] For $n>0$, $m>2$ and if $a=\frac{m}{2}$, $b=\frac{n}{2}$, $\theta=\frac{m}{n}$ then $X\sim F_{n,m}$.

REMARK 2.1. Since

$$(\mu-x)\frac{\exp\left(\frac{\mu}{\sqrt{\delta\gamma}}\tan^{-1}(x\sqrt{\delta/\gamma})\right)}{(\delta x^2+\gamma)^{1+\frac{1}{2\delta}}}=\frac{\mathrm{d}}{\mathrm{d}x}\frac{\exp\left(\frac{\mu}{\sqrt{\delta\gamma}}\tan^{-1}(x\sqrt{\delta/\gamma})\right)}{(\delta x^2+\gamma)^{\frac{1}{2\delta}}},$$

it follows that $\mathbb{E}X=\mu$ for the Student-type densities (type 4), while for all other cases it is evident to check 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 μ 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 $\delta\leqslant 0$ 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 $n\geqslant 1$ (that is, $\delta<\frac{1}{n-1}$) then

$$\lim_{x\nearrow\omega}x^kq(x)f(x)=\lim_{x\searrow\alpha}x^kq(x)f(x)=0,\quad k=0,1,\ldots,n-1,\tag{2.2}$$

and, in general, for any $c\in\mathbb{R}$,

$$\lim_{x\nearrow\omega}(x-c)^kq(x)f(x)=\lim_{x\searrow\alpha}(x-c)^kq(x)f(x)=0,\quad k=0,1,\ldots,n-1.\tag{2.3}$$

Proof. Since $x^kq(x)f(x)=x^k\int_{\alpha}^{x}(\mu-t)f(t)\mathrm{d}t$, $\alpha<x<\omega$, the second limit in (2.2) is trivial whenever $\alpha>-\infty$ and the first one is trivial whenever $\omega<\infty$. If $\omega=\infty$ it suffices to verify the first limit in (2.2) only when $k=n-1$ and $n\geqslant 2$ (because the case $k=0$ is obvious); then, since $q(x)f(x)$ is eventually decreasing we have that for large enough $x>0$,

$$\begin{aligned}x^{n-1}q(x)f(x)&=q(x)f(x)\frac{(n-1)2^{n-1}}{2^{n-1}-1}\int_{x/2}^{x}t^{n-2}\mathrm{d}t\\&\leqslant\frac{(n-1)2^{n-1}}{2^{n-1}-1}\int_{x/2}^{x}t^{n-2}q(t)f(t)\mathrm{d}t\\&\leqslant\frac{(n-1)2^{n-1}}{2^{n-1}-1}\int_{x/2}^{\infty}t^{n-2}q(t)f(t)\mathrm{d}t\to 0,\ \text{ as } x\to\infty,\end{aligned}$$

because $\deg(q) \leqslant 2$ and, by assumption, $\mathbb{E}q(X)|X|^{n-2} < \infty$. The case $\alpha = -\infty$ is translated to the previous one by considering the random variable $\widetilde{X} = -X$ with density $\widetilde{f}(x) = f(-x)$, quadratic polynomial $\widetilde{q}(x) = q(-x)$ and support $J(\widetilde{X}) = (\widetilde{\alpha}, \widetilde{\omega}) = (-\omega, -\alpha) = (-\omega, \infty)$. Then $\mathbb{E}|\widetilde{X}|^n = \mathbb{E}|X|^n < \infty$ and

$$\lim_{x\to-\infty} x^k q(x) f(x) = (-1)^k \lim_{x\to\infty} x^k q(-x) f(-x) = (-1)^k \lim_{x\to\infty} x^k \widetilde{q}(x)\widetilde{f}(x) = 0$$

for all $k \in \{0, 1, \ldots, n-1\}$. Now it suffices to observe that all limits in (2.3) are linear combinations of limits in (2.2). Indeed, the first limit in (2.3) is

$$\lim_{x\nearrow\omega} (x-c)^k q(x) f(x) = \textstyle\sum_{i=0}^k \binom{k}{i} (-c)^{k-i} \lim_{x\nearrow\omega} x^i q(x) f(x) = 0$$

and, similarly, the second limit in (2.3) is

$$\lim_{x\searrow\alpha} (x-c)^k q(x) f(x) = \textstyle\sum_{i=0}^k \binom{k}{i} (-c)^{k-i} \lim_{x\searrow\alpha} x^i q(x) f(x) = 0. \qquad \square$$

LEMMA 2.2. If $X \sim \mathrm{IP}(\mu; \delta, \beta, \gamma)$ and $\mathbb{E}|X|^n < \infty$ for some $n \geqslant 2$ (that is, $\delta < \frac{1}{n-1}$) then for any $c \in \mathbb{R}$, the central moments about c satisfy the recurrence

$$\mathbb{E}(X-c)^{k+1} = \frac{(\mu - c + kq'(c))\mathbb{E}(X-c)^k + kq(c)\mathbb{E}(X-c)^{k-1}}{1-k\delta}, \quad k = 1, 2, \ldots, n-1, \tag{2.4}$$

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

$$\mathbb{E}X^{k+1} = \frac{(\mu + k\beta)\mathbb{E}X^k + k\gamma\mathbb{E}X^{k-1}}{1-k\delta}, \quad k = 1, 2, \ldots, n-1, \tag{2.5}$$

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

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

$$\mathbb{E}(X-\mu)^{k+1} = \frac{kq'(\mu)\mathbb{E}(X-\mu)^k + kq(\mu)\mathbb{E}(X-\mu)^{k-1}}{1-k\delta}, \quad k = 1, 2, \ldots, n-1, \tag{2.6}$$

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

Proof. If $J(X) = (\alpha, \omega)$ is the support of X and $k \in \{1, 2, \ldots, n-1\}$ then we have

$$\begin{aligned}\mathbb{E}(X-c)^{k+1} &= \mathbb{E}[((\mu - c) - (\mu - X))(X-c)^k] \\ &= (\mu - c)\mathbb{E}(X-c)^k - \int_\alpha^\omega (x-c)^k (\mu - x) f(x)\mathrm{d}x.\end{aligned}$$

Using (2.3) and the fact that $q(X) = \delta(X-c)^2 + q'(c)(X-c) + q(c)$ we see that

$$\begin{aligned}
-\int_\alpha^\omega (x-c)^k(\mu - x)f(x)\mathrm{d}x &= -\int_\alpha^\omega (x-c)^k(q(x)f(x))'\mathrm{d}x \\
&= -(x-c)^k q(x)f(x)\big|_\alpha^\omega + k\mathbb{E}q(X)(X-c)^{k-1} \\
&= k\delta\mathbb{E}(X-c)^{k+1} + kq'(c)\mathbb{E}(X-c)^k + kq(c)\mathbb{E}(X-c)^{k-1}.
\end{aligned}$$

Therefore,

$$(1-k\delta)\mathbb{E}(X-c)^{k+1} = (\mu - c + kq'(c))\mathbb{E}(X-c)^k + kq(c)\mathbb{E}(X-c)^{k-1}, \qquad k = 1,2,\ldots,n-1,$$

and, since the initial conditions are obvious, (2.4) follows. □

3 Comparison with the ordinary Pearson system

The ordinary Pearson family consists of absolutely continuous random variables X supported in some (open) interval (α,ω), such that their density f, which is assumed strictly positive and differentiable in (α,ω), satisfies the *Pearson differential equation*

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

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 nominator 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 non-constant, 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, above. 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 (α,ω). Also, it may happen that p_1 and p_2 have common zeros outside the interval (α,ω); this is 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 the ‘correct’ choice of p_2 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 decides 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 $\widetilde{p}_2^{(1)}$, have become 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 the resulting polynomials p_1 and p_2 satisfy the conditions $\deg(p_1)\leqslant 1$ and $\deg(p_2)\leqslant 2$ 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 system.

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, the polynomial p_2 satisfies the relations $p_2(\alpha)=0$ 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 we have chosen a correct p_2 in each case where such a p_2 exists. For example, the standard exponential density,

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

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

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

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,

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

satisfies (3.1) with $(p_1, p_2) = (2-x, x(2-x))$ and the 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,

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

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

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

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

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

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

The algorithm is justified because of 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, $Y = \lambda X + c$ is also Integrated Pearson for all $\lambda \neq 0$ and $c \in \mathbb{R}$; see Proposition 2.1(vi). Also, its density $f_Y(x) = \frac{1}{|\lambda|} f(\frac{x-c}{\lambda})$ satisfies, by assumption, the differential equation

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

where $(\widetilde{\alpha}, \widetilde{\omega}) = (\lambda\alpha + c, \lambda\omega + c)$ or $(\lambda\omega + c, \lambda\alpha + c)$, according to $\lambda > 0$ or $\lambda < 0$, respectively. 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 the relationships

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

for some $i \in \{1, 2, 3\}$. Therefore, it suffices to show that $\deg(p_i^Y) \leqslant i$, $i = 1, 2$, and that the quadratic polynomial $q_Y(x) = \lambda^2 q(\frac{x-c}{\lambda})$ of Y 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 see these cases separately.

If $f\sim\Gamma(a,\lambda)$ with $a\neq 1$ ($\alpha>0$, $\lambda>0$) 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, $\deg(p_i)=1$ ($i=1,2$) and

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

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

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

here the parameters are a,b,θ with $a>1$, $b>0$ 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

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

Therefore, Step 3 yields $(p_1,p_2)=(-c(a+1)x,cx(x+\theta))$ and, 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

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

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

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

Therefore, Step 3 yields $(p_1,p_2)=(-(b-1)\theta x,\theta x(1-x))$ and, 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

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

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

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

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 subsumes all cases and 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 (α,ω) and has finite (unknown) mean. Then, the following 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. Assume first that (B) holds. Since $f\sim \mathrm{IP}(\mu;q)$, (2.1) shows that $f'/f=\widetilde{p}_1/\widetilde{p}_2$ where $\widetilde{p}_1=\mu-x-q'$ and $\widetilde{p}_2=q$. Putting the polynomials $\widetilde{p}_1=\mu-x-q'$ and $\widetilde{p}_2=q$ in Step 0 of the above algorithm and using Proposition 3.1 we conclude that the resulting polynomials p_1 and p_2 (of Step 3) satisfy the requirements of (A); also, $q(x)=\theta p_2(x)$ for some $\theta\neq 0$.

Assume now that (A) holds. Using a suitable mapping $Y=\lambda X+c$, $\lambda\neq 0$, $c\in\mathbb{R}$, we can transform the interval (α,ω) into $(\widetilde{\alpha},\widetilde{\omega})$, where $(\widetilde{\alpha},\widetilde{\omega})$ is one of the intervals $(0,1)$, $(0,\infty)$ or $(-\infty,\infty)$. The polynomials p_1 and p_2 are transformed to $p_1^Y(x)=\lambda p_1(\frac{x-c}{\lambda})$ and

$p_2^Y(x)=\lambda^2 p_2(\frac{x-c}{\lambda})$, and the differential equation (3.1) yields $f_Y'(x)/f_Y(x)=p_1^Y(x)/p_2^Y(x)$, $\widetilde{\alpha}<x<\widetilde{\omega}$, where f_Y is the density of Y and $(\widetilde{\alpha},\widetilde{\omega})$ its support. Moreover, it is easy to see that p_1^Y and p_2^Y satisfy the requirements of (A), i.e., $p_2^Y(\widetilde{\alpha})=0$ if $\widetilde{\alpha}>-\infty$, $p_2^Y(\widetilde{\omega})=0$ if $\widetilde{\omega}<\infty$ and $p_2^Y(x)\neq 0$ for all $x\in(\widetilde{\alpha},\widetilde{\omega})$. Clearly, in view of Proposition 2.1(vi), it suffices to verify that Y is Integrated Pearson. Thus, from now on (and without any loss of generality) we shall assume that (α,ω) is one of the intervals $(0,1)$, $(0,\infty)$ or $\mathbb{R}$.

If $(\alpha,\omega)=(0,1)$ then the assumptions (A) show that $p_2(x)=\theta x(1-x)$ for some $\theta\neq 0$. Let $p_1(x)=a_0+a_1x$. Solving (3.1) we get

$$f(x)=Cx^{a_0/\theta}(1-x)^{-(a_0+a_1)/\theta},\ \ 0<x<1,$$

where, necessarily, $1+a_0/\theta>0$ and $1-(a_0+a_1)/\theta>0$. Thus,

$$(1+a_0/\theta)+(1-(a_0+a_1)/\theta)=(2\theta-a_1)/\theta>0,$$

so that $2\theta-a_1\neq 0$. It follows that $f\sim B(a,b)$ with $a=1+a_0/\theta$, $b=1-(a_0+a_1)/\theta$ and, therefore,

$$q(x)=\frac{x(1-x)}{a+b}=\frac{x(1-x)}{2-a_1/\theta}=\frac{p_2(x)}{2\theta-a_1}.$$

Assume that $(\alpha,\omega)=(0,\infty)$. Then, assumptions (A) show that the possible forms of p_2 are either $p_2=\theta x$ or $p_2=\theta x^2$ or $p_2=\theta x(x+c)$ for some $\theta\neq 0$ and $c>0$. If $p_2=\theta x$ set $p_1=a_0+a_1x$ and solve (3.1) to obtain

$$f(x)=Cx^{a_0/\theta}\exp(a_1x/\theta),\ \ x>0,$$

where, necessarily, $a_0/\theta>-1$ and $a_1/\theta<0$; thus, $X\sim\Gamma(a,\lambda)$ with $a=\frac{a_0}{\theta}-1>0$ and $\lambda=-\frac{a_1}{\theta}>0$. Therefore, $a_1\neq 0$ and

$$q(x)=\frac{x}{\lambda}=\frac{p_2(x)}{-a_1}.$$

If $p_2=\theta x^2$, set $p_1=a_0+a_1x$ and solve (3.1) to obtain

$$f(x)=Cx^{a_1/\theta}\exp(-a_0/(\theta x)),\ \ x>0,$$

where, necessarily, $a_0/\theta>0$ and $a_1/\theta<-2$; these conditions are necessary and sufficient for $\int_0^\infty f(x)\mathrm{d}x$ and $\int_0^\infty xf(x)\mathrm{d}x$ to be finite. Therefore, $f(x)=Cx^{-a-1}e^{-\lambda/x}$, $x>0$, where $a=-1-\frac{a_1}{\theta}>1$ and $\lambda=\frac{a_0}{\theta}>0$. Observe now that f is of Reciprocal Gamma type (type 5) and $q(x)=\frac{x^2}{a-1}$. Since $a=-1-\frac{a_1}{\theta}>1$ it follows that $\frac{-a_1-2\theta}{\theta}>0$ and, finally, $a_1+2\theta\neq 0$. Thus,

$$q(x)=\frac{x^2}{a-1}=\frac{\theta x^2}{-a_1-2\theta}=\frac{p_2(x)}{-a_1-2\theta}.$$

Assume now that $p_2=\theta x(x+c)$, $\theta\neq 0$, $c>0$ and let $p_1=a_0+a_1x$. Solving (3.1) we obtain

$$f(x)=Cx^{\frac{a_0}{c\theta}}(x+c)^{\frac{a_1c-a_0}{c\theta}},\ \ x>0,$$

where, necessarily, $\frac{a_0}{c\theta} > -1$ and $\frac{a_1}{\theta} < -2$; these conditions are necessary and sufficient for $\int_0^\infty f(x)\mathrm{d}x$ and $\int_0^\infty xf(x)\mathrm{d}x$ to be finite. Now observe that $f(x) = Cx^{b-1}(x+c)^{-a-b}$ $(x>0)$ is of Snedecor-type (type 6) with $a = -\frac{a_1}{\theta} - 1 > 1$ and $b = 1 + \frac{a_0}{c\theta} > 0$. From $\frac{a_1}{\theta} < -2$ we get $a_1 + 2\theta \neq 0$ and, thus, we conclude that (see Table 2.1)

$$q(x) = \frac{x(x+c)}{a-1} = \frac{x(x+c)}{-2-a_1/\theta} = \frac{\theta x(x+c)}{-a_1-2\theta} = \frac{p_2(x)}{-a_1-2\theta}.$$

Finally, assume that $(\alpha,\omega) = \mathbb{R}$. In this case assumptions (A) imply that either $p_2 \equiv \theta \neq 0$ or $p_2 = \pm(\theta(x-c)^2+\lambda)$ with $\theta > 0$, $\lambda > 0$ and $c \in \mathbb{R}$. Assume first that $p_2 \equiv \theta \neq 0$ and let $p_1 = a_0 + a_1 x$. Then, it is easily seen from (3.1) that

$$f(x) = C\exp\left(\frac{a_1}{2\theta}x^2 + \frac{a_0}{\theta}x\right), \quad x \in \mathbb{R}.$$

This can represents a density if and only if $\frac{a_1}{2\theta} < 0$; in this case it is easily seen that $f \sim N(\mu,\sigma^2)$ with $\mu = \frac{-a_0}{a_1}$, $\sigma = \sqrt{-\frac{\theta}{a_1}}$, and thus,

$$q(x) \equiv \sigma^2 = -\frac{\theta}{a_1} = \frac{p_2(x)}{-a_1}.$$

For the last remaining case it suffices to consider

$$p_2(x) = \theta(x-c)^2+\lambda \ \text{ and } \ p_1(x) = a_0 + a_1(x-c) \ \text{ where } \theta > 0,\ \lambda > 0 \text{ and } a_0, a_1, c \in \mathbb{R}.$$

Also, using the transformation $X \mapsto X - c$, the general case is simplified to $p_2 = \theta x^2 + \lambda$ and $p_1 = a_0 + a_1 x$. Now, the differential equation (3.1) has the general solution

$$f(x) = C(\theta x^2+\lambda)^{\frac{a_1}{2\theta}} \exp\left[\frac{a_0}{\sqrt{\theta\lambda}}\tan^{-1}(x\sqrt{\theta/\lambda})\right], \quad x \in \mathbb{R}.$$

The necessary and sufficient condition for this f to represent a density with finite mean is $-\frac{a_1}{2\theta} - 1 > 0$ or, equivalently, $a_1 + 2\theta < 0$. Therefore, setting

$$\delta = \frac{\theta}{-a_1-2\theta} > 0, \quad \gamma = \frac{\lambda}{-a_1-2\theta} > 0 \ \text{ and } \ \mu = \frac{a_0}{-a_1-2\theta} \in \mathbb{R}$$

we see that this is a Student-type density (type 4); see Table 2.1. Consequently,

$$q(x) = \delta x^2 + \gamma = \frac{\theta x^2+\lambda}{-a_1-2\theta} = \frac{p_2(x)}{-a_1-2\theta},$$

and the proof is complete. □

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 (α,ω) and vanishes at all (if any) finite endpoints of (α,ω).

If the mean μ 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 μ. Assume that the set $\{x : f(x) > 0\}$ is the (bounded or unbounded) interval $J(X) = (\alpha,\omega)$ and that f is differentiable in (α,ω) 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) and (ii), below, hold:

(i) there exist a constant $\theta \neq 0$ such that $p_1(x) + p_2'(x) = (\mu - x)/\theta$, $x \in \mathbb{R}$, and

(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 (α,ω) 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, $p_2(x)f(x) = q(x)f(x) = \int_\alpha^x (\mu - t)f(t)\mathrm{d}t$ – 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 get

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

Now, let us take into account the first assumption in (ii), $\lim_{x\searrow\alpha} p_2(x)f(x) = 0$. Taking limits in (3.2) and using dominated convergence for the l.h.s. we conclude that

$$\int_\alpha^y (\mu - t)f(t)\mathrm{d}t = \theta p_2(y)f(y), \quad \alpha < y < \omega;$$

that is, $X \sim \mathrm{IP}(\mu;q)$ with $q(x) = \theta p_2(x)$. Clearly we get the same conclusion if we use the second assumption in (ii), $\lim_{y\nearrow\omega} p_2(y)f(y) = 0$, and evaluate the limits as $y \nearrow \omega$ in (3.2); in this case we get the identity $\int_x^\omega (t - \mu)f(t)\mathrm{d}t = \theta p_2(x)f(x) = q(x)f(x)$, $\alpha < x < \omega$, which is equivalent to (1.1), since $\int_\alpha^\omega (\mu - t)f(t)\mathrm{d}t = 0$.

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,

$$\theta p_2(y)f(y) = \theta p_2(x)f(x) + \int_x^y (\mu - t)f(t)\mathrm{d}t, \quad \alpha < x < y < \omega,$$

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

$$\theta p_2(x)f(x) = C + \int_x^\omega (t - \mu)f(t)\mathrm{d}t, \quad \alpha < x < \omega,$$

and since $\lim_{x\searrow\alpha}\int_x^{\omega}(t-\mu)f(t)\mathrm{d}t=\int_{\alpha}^{\omega}(t-\mu)f(t)\mathrm{d}t=0$ 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 (α,ω) and X has 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 $x>\max\{2x_0,0\}$,

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

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. Actually, these polynomials are by-products of the pair (p_1,p_2) that appears in the nominator and the denominator of the differential equation (3.1); that is, they have nothing to do either with f or with the interval (α,ω).

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

DEFINITION 4.1. Let $\varnothing\neq(\alpha,\omega)\subseteq\mathbb{R}$, 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 (α,ω), or simply *compatible*, if there exists a differentiable function $f:(\alpha,\omega)\to\mathbb{R}$, $f\not\equiv 0$ (f is not assumed nonnegative or integrable), such that the following *generalized Pearson differential equation* is fulfilled:

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

In other words, (p_1,p_2) is compatible if (4.1) has non-trivial solutions for f.

It is easily seen that (p_1,p_2) is compatible whenever p_2 has no roots in (α,ω); in this case, the general solution f is $C^{\infty}(\alpha,\omega)$ and can be chosen to be strictly positive in (α,ω). The presence of a zero of p_2 in (α,ω), however, may results in incompatibility; e.g., 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 (α,ω) 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, in the worst case, the three 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 $f_i=C_ie^{g_i}$

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, $|f| \geqslant 0$ will be also a non-trivial 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 (α, ω) 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 $\alpha \leqslant \widetilde{\alpha} < \widetilde{\omega} \leqslant \omega$, or of the form $(\widetilde{\alpha}, \rho_1) \cup (\rho_2, \widetilde{\omega}) \subseteq (\alpha, \omega)$ with $\alpha \leqslant \widetilde{\alpha} < \rho_1 \leqslant \rho_2 < \widetilde{\omega} \leqslant \omega$, 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, $\rho \notin S(f)$) whenever ρ is a zero of p_2 which is not a zero of p_1.

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 (α, ω). Moreover, assuming that (p_1, p_2) is compatible in (α, ω), 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 (α, ω).

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 $\{h_k, k = 1,2,\ldots\}$, defined by a *Rodrigues-type formula*, as follows:

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 (α, ω) 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

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

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

$$\deg(h_k)\leqslant k \quad \text{and} \quad \mathrm{lead}\,(h_k)=\prod_{j=k+1}^{2k}(a_1+jb_2), \quad k=0,1,2,\ldots, \tag{4.3}$$

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 $\deg(\widetilde{h}_k)\leqslant k$) are defined inductively. Each polynomial $\widetilde{h}_k$ can be viewed as the value of a functional $\mathscr{R}_k$ that maps any pair (p_1,p_2) to a real polynomial of degree at most k. The form of this functional is

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

where the sum ranges over all integers $r,i,j\geqslant 0$ with $r+i+2j\leqslant k$, and the constant $C_{k;rij}^{a_1,b_2}$ depends only on $k,r,i,j,p_1'=a_1$ and $p_2''=2b_2$. On the other hand it is clear that, given an arbitrary pair (p_1,p_2) with $p_2\not\equiv 0$, we can fix an interval (α,ω) containing no roots of p_2. With the help of a positive solution f of the differential equation (4.1) we 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 expand the first four:

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

Provided that the solution f of (4.1) is a probability density in (α,ω), the polynomials h_k are candidate 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 (α,ω). 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 the Sturm-Liouville theory. Specifically, it is shown in Theorem 1 of [9] (see p. 295) that each polynomial h_k satisfies the equation

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

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 consist of 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 this f 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 $\theta \neq 0$. We mention here that, even for Integrated Pearson densities, a wrong choice of p_2 results in non-orthogonality of the Rodrigues polynomials; see, e.g., the polynomials $h_k = P_k^2$ given in [9], 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 for this density is given by $p_2 = x(x_0 - x)$.

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) The density f process a suitable number of moments so that $\int_\alpha^\omega |h_k(t)h_m(t)|f(t)\mathrm{d}t < \infty$;

(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,

$$\int_\alpha^\omega h_k(x)h_m(x)f(x)\mathrm{d}x = 0.$$

[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 equals to $(k-m)(a_1 + (k+m+1)b_2) \times \int_\alpha^\omega h_k(t)h_m(t)f(t)\mathrm{d}t$.]

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

$$\lambda h_k(t)h_m(t)f(t) = h_m(t)[f(t)p_2(t)h_k'(t)]' - h_k(t)[f(t)p_2(t)h_m'(t)]', \quad \alpha < t < \omega, \tag{4.5}$$

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

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

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

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

Therefore, taking limits as $x \searrow \alpha$ and $y \nearrow \omega$ and using (a) and (c) we get the 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, it is true that under (a),

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

□

The following result is an immediate consequence of Lemma 4.1.

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

(i) The density f process $2n-1$ 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 $j\in\{0,1,\ldots,2n-2\}$.

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 $k\in\{0,1,\ldots,n\}$. Fix $k,m\in\{0,1,\ldots,n\}$ 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 $\int_\alpha^\omega h_k(x)h_m(x)f(x)\mathrm{d}x$ is (well-defined and) finite. Finally, since $h_k'h_m-h_kh_m'$ is a polynomial of degree $k+m-1$ (observe that $\mathrm{lead}(h_k'h_m-h_kh_m')=(k-m)\mathrm{lead}(h_k)\mathrm{lead}(h_m)\neq 0$), (iii) ensures that assumption (c) of Lemma 4.1 is also fulfilled and, hence,

$$\int_\alpha^\omega h_k(x)h_m(x)f(x)\mathrm{d}x=0.$$

Conversely, assume that the set $\mathscr{H}_n=\{h_0,h_1,\ldots,h_n\}$ is orthogonal with respect to f; that is, $\mathbb{E}|h_k(X)h_m(X)|=\int_\alpha^\omega|h_k(x)h_m(x)|f(x)\mathrm{d}x<\infty$ for all $k,m\in\{0,1,\ldots,n\}$ with $m\neq k$, and $\int_\alpha^\omega h_k(x)h_m(x)f(x)\mathrm{d}x=0$. It follows that, necessarily, $\deg(h_k)=k$ for all $k=1,2,\ldots,n$; for if k is the smallest integer in $\{1,2,\ldots,n\}$ for which $\mathrm{lead}(h_k)=0$ 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

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

Subsequently, the inequality

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

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

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

by the orthogonality assumption. Since h_k is continuous (a polynomial) and f is positive in a subinterval of (α,ω) with positive length, it follows that $h_k \equiv 0$, which contradicts the assumption that $\mathscr{H}_n$ contains only non-zero elements. Therefore, $\prod_{k=0}^n \mathrm{lead}(h_k) \neq 0$, 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 $k \neq m$), we 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 $k \geqslant 1$. 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 $\mathscr{H}_n = \{1,0,\ldots,0\}$ 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 explains that 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. Clearly, any set containing three (or more) polynomials cannot be an orthogonal set.

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 $h_0h_2 \geqslant 3$ 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? In other words, it is natural to state the following question:

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

The answer is in the affirmative. In particular, the following 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 $\mathscr{H}_2=\{h_0,h_1,h_2\}$ and assume that $\mathscr{H}_2$ is non-trivial, i.e., $h_1 \not\equiv 0$ and $h_2 \not\equiv 0$. If the system $\mathscr{H}_2$ is orthogonal with respect to f then there exists a subinterval $(\alpha',\omega') \subseteq (\alpha,\omega)$, a quadratic polynomial

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

and a number $\mu \in (\alpha',\omega')$ 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

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

where

$$L_j(\alpha):=\lim_{x\searrow\alpha} x^j p_2(x) f(x), \quad L_j(\omega):=\lim_{x\nearrow\omega} x^j p_2(x) f(x). \tag{4.10}$$

To simplify cases we can apply an affine transformation $x \mapsto \lambda x+c$ ($\lambda \neq 0$, $c \in \mathbb{R}$) to f. By considering $\widetilde{f}(x)=\frac{1}{|\lambda|} f(\frac{x-c}{\lambda})$ in place of f it is easily seen that (4.1) is satisfied in the translated interval $(\widetilde{\alpha},\widetilde{\omega})$ for $\widetilde{p}_1(x)=\lambda p_1(\frac{x-c}{\lambda})$ and $\widetilde{p}_2(x)=\lambda^2 p_2(\frac{x-c}{\lambda})$; 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 $\widetilde{h}_k(x)=\lambda^k h_k(\frac{x-c}{\lambda})$; thus, $\mathrm{lead}(\widetilde{h}_k)=\mathrm{lead}(h_k)$ and, in particular, the system $\mathscr{H}_2$ is non-trivial if and only if the same is true for the system $\widetilde{\mathscr{H}_2}:=\{\widetilde{h}_0,\widetilde{h}_1,\widetilde{h}_2\}$. The orthogonality of the system $\widetilde{\mathscr{H}_2}$ with respect to $\widetilde{f}$ is equivalent to the orthogonality of the system $\mathscr{H}_2$ with respect to f; indeed,

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

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 $\widetilde{L}_j(\widetilde{\alpha}):=\lim_{x\searrow\widetilde{\alpha}} x^j \widetilde{p}_2(x)\widetilde{f}(x)$, $\widetilde{L}_j(\widetilde{\omega}):=\lim_{x\nearrow\widetilde{\omega}} x^j \widetilde{p}_2(x)\widetilde{f}(x)$. 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, and in view of Proposition 2.1(vi), that we can freely apply any affine transformation, either to the polynomial p_2 or to the density f and its support (α,ω); 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 non-equivalent cases by solving the differential equation (4.1) in each case. We shall try to give a somewhat complete approach as follows:

Assume first that $\deg(p_2)=2$ and that its discriminant, Δ, is strictly negative. Applying an affine transformation and dividing both p_1 and p_2 by $\mathrm{lead}(p_2)\neq 0$ we may assume that $p_2=x^2+b_0$ for some $b_0>0$. If $p_1\equiv 0$ then, necessarily, (α,ω) is finite and $f\sim U(\alpha,\omega)$; but in this case, $h_2(x)=6x^2+4b_0\geqslant 4b_0>0$ cannot be orthogonal to $h_0\equiv 1$. If $\deg(p_1)=0$, that is, $p_1\equiv a_0\neq 0$, then the density

$$f(x)=C\exp\left(\frac{a_0}{b_0}\tan^{-1}\left(\frac{x}{\sqrt{b_0}}\right)\right)$$

is bounded away from zero, so that (α,ω) must be again finite. Then, the assumed orthogonality of $\mathscr{H}_2$ fails because (4.9) shows that $\alpha=\omega$. Finally, assume that $\deg(p_1)=1$ i.e. $p_1=a_0+a_1x$ with $a_1\neq 0$. In this case, $a_1\notin\{-2,-3,-4\}$ because of (4.8). Since

$$f(x)=C(x^2+b_0)^{\frac{a_1}{2}}\exp\left(\frac{a_0}{b_0}\tan^{-1}\left(\frac{x}{\sqrt{b_0}}\right)\right),$$

it follows that either (α,ω) is finite or, otherwise, $a_1<-4$ (for the third moment to exists). If (α,ω) is finite, the assumed orthogonality fails because (4.9) shows that $\alpha=\omega$. If $\alpha>-\infty$, $\omega=\infty$ then the assumed orthogonality fails again from (4.9) since $L_0(\alpha)>0$, $L_0(\infty)=0$. The case $\alpha=-\infty$, $\omega<\infty$ is similar to the previous one (we can also make the transformation $x\mapsto -x$). Therefore, the unique case where $\mathscr{H}_2$ is indeed orthogonal is when $(\alpha,\omega)=\mathbb{R}$. Then,

$$\mathbb{E}h_1(X)=(a_1+2b_2)\mu+(a_0+2b_1)=0 \text{ implies that } \mu=\frac{a_0}{-2-a_1}$$

(note that $b_2=1$, $b_1=0$) and, hence, $p_1+p_2'=(-2-a_1)(\mu-x)$. In view of Proposition 3.3 we see that

$$f\sim \mathrm{IP}(\mu;q) \text{ with } \mu=\frac{a_0}{-2-a_1} \text{ and } q(x)=\frac{x^2+b_0}{-2-a_1}=\frac{p_2(x)}{-2-a_1}.$$

Next, assume that $\deg(p_2)=2$ and $\Delta=0$. Applying an affine transformation and dividing both p_1 and p_2 by $\mathrm{lead}(p_2)\neq 0$ we may further assume that $p_2=x^2$. If $p_1\equiv 0$ then, necessarily, (α,ω) is finite and $f\sim U(\alpha,\omega)$; but in this case, $h_2(x)=12x^2\geqslant 0$ cannot be orthogonal to $h_0\equiv 1$. Let $\deg(p_1)=0$, that is, $p_1\equiv a_0\neq 0$. With the map $x\mapsto -x$, if necessary, we may further translate the density to have either the form

$$f(x)=Ce^{-\frac{a_0}{x}},\ \ 0\leqslant\alpha<x<\omega<\infty,$$

or the form

$$f(x) = \begin{cases} C_1 e^{-\frac{a_0}{x}}, & 0 < x < \omega < \infty, \\ 0, & -\infty \leqslant \alpha < x \leqslant 0, \end{cases}$$

where, necessarily, $a_0 > 0$ in the second case. In both cases the assumed orthogonality fails because of (4.9). Finally, assume that $\deg(p_1) = 1$ i.e. $p_1 = a_0 + a_1 x$ with $a_1 \neq 0$. In this case, $a_1 \notin \{-2, -3, -4\}$ because of (4.8). With the map $x \mapsto -x$, if necessary, we may further translate the density to have either the form

$$f(x) = C x^{a_1} e^{-\frac{a_0}{x}}, \ 0 \leqslant \alpha < x < \omega \leqslant \infty,$$

or the form

$$f(x) = \begin{cases} C_1 x^{a_1} e^{-\frac{a_0}{x}}, & 0 < x < \omega \leqslant \infty, \\ 0, & -\infty \leqslant \alpha < x \leqslant 0, \end{cases}$$

where, necessarily, $a_0 > 0$ in the second case. If $\omega < \infty$ then, due to (4.9), the assumed orthogonality fails for both cases. If $\omega = \infty$ and $\alpha > 0$ then we must take $a_1 < -4$ for the finiteness of the third moment (note that in this case, $a_0 \in \mathbb{R}$ can be arbitrary since $\alpha > 0$), but the orthogonality fails because of (4.9), since $L_0(\alpha) > 0$, $L_0(\infty) = 0$. In the last case where $\alpha \leqslant 0$ and $\omega = \infty$ (thus, $a_0 > 0$ and $a_1 < -4$) the orthogonality is indeed satisfied. This is so because it is easy to verify both (4.9) and (4.8). On the other hand, since we have assumed that $\mathbb{E}h_1(X) = (a_1 + 2b_2)\mu + (a_0 + b_1) = 0$, it follows that $\mu = \frac{a_0}{-a_2-2}$ (note that $b_2 = 1$, $b_1 = 0$) and $p_1 + p_2' = (-2 - a_1)(\mu - x)$. In view of Proposition 3.3, this density belongs to the Integrated Pearson system with

$$\mu = \frac{a_0}{-a_2 - 2} \quad \text{and} \quad q(x) = \frac{p_2(x)}{-2 - a_1} = \frac{x^2}{-2 - a_1}.$$

Moreover, observe that its support, $(\alpha', \omega) = (0, \infty) \subsetneq (\alpha, \omega)$, is different than (α, ω), whenever $\alpha < 0$.

Next, assume that $\deg(p_2) = 2$ and $\Delta > 0$. Applying an affine transformation and dividing both p_1 and p_2 by $\mathrm{lead}(p_2) \neq 0$ we may further assume that $p_2 = x(1 - x)$. Solving the differential equation (4.1) for arbitrary p_1 and for all $x \in \mathbb{R} \smallsetminus \{0, 1\}$ we see that the general solution has the form

$$f(x) = \begin{cases} C_1 (-x)^A (1 - x)^B, & \text{if } x < 0, \\ C_2 x^A (1 - x)^B, & \text{if } 0 < x < 1, \\ C_3 x^A (x - 1)^B, & \text{if } x > 1, \end{cases}$$

where A and B are arbitrary parameters and $C_1, C_2, C_3 \geqslant 0$ are arbitrary constants, not all zero. The restrictions on A and B depend on the interval (α, ω) that we consider and the positivity or vanishing of each branch; they have to be chosen in such a way that the resulting function is differentiable and integrable in (α, ω). For example, if $[0, 1] \subseteq (\alpha, \omega)$ and $C_1, C_2, C_3 > 0$ then, in order that f is (continuous and) differentiable at the points 0 and 1, we must take $A > 1$ and $B > 1$; but then it is necessary for (α, ω) to be bounded, since, otherwise, the resulting f could not be integrable. The several possibilities can be classified according to the number of roots of p_2 that fall into (α, ω), as follows:

(1) Let $\{0,1\}\cap(\alpha,\omega)=\varnothing$. Then, either $(\alpha,\omega)\subseteq(0,1)$ or $(\alpha,\omega)\subseteq(-\infty,0)$ or $(\alpha,\omega)\subseteq(1,\infty)$. For the first case we observe that (4.9) fails whenever $\alpha>0$ or $\omega<1$; if $(\alpha,\omega)=(0,1)$ then $A>-1$, $B>-1$, $p_1=A-(A+B)x$ and the orthogonality assumption yields $\mathbb{E}h_1(X)=(a_1+2b_2)\mu+(a_0+b_1)=-(A+B+2)\mu+(A+1)=0$. Hence, $p_1+p_2'=A+1-(A+B+2)x=(A+B+2)(\mu-x)$ and $p_2(x)f(x)=Cx^{A+1}(1-x)^{B+1}\to 0$ as $x\nearrow 1$; thus, Proposition 3.3 shows that $f\sim\mathrm{IP}(\mu;q)$ with

$$\mu=\frac{A+1}{A+B+2}\quad\text{and}\quad q(x)=\frac{p_2(x)}{A+B+2}.$$

Using the map $x\mapsto -x$ for the second case and the map $x\mapsto x-1$ for the third case it is seen that both cases are reduced to $(\alpha,\omega)\subseteq(0,\infty)$ and translate p_2 to $p_2=-x(x+1)$; equivalently, we can take $p_2=x(x+1)$. Moreover, the general solution in this case takes the form

$$f(x)=Cx^{\theta}(x+1)^{\lambda},\quad 0\leqslant\alpha<x<\omega\leqslant\infty.$$

If $\omega<\infty$ or $\alpha>0$ it is easily seen that (4.9) fails. In the remaining case where $(\alpha,\omega)=(0,\infty)$ we must have $\theta>-1$ (for integrability close to zero) and $\theta+\lambda<-4$ (for finiteness of the third moment). Since $p_1=a_0+a_1x=\theta+(\theta+\lambda)x$, $p_2=b_0+b_1x+b_2x^2=x+x^2$ and $h_1=(a_1+2b_2)x+(a_0+b_1)=(\theta+\lambda+2)x+(\theta+1)$, the assumed orthogonality yields $\mathbb{E}h_1(X)=(\theta+\lambda+2)\mu+\theta+1=0$; thus, $p_1+p_2'=(\theta+\lambda+2)x+(\theta+1)=-(\theta+\lambda+2)(\mu-x)$ and Proposition 3.3 shows that

$$f\sim\mathrm{IP}(\mu;q)\ \text{ with }\ \mu=\frac{\theta+1}{-(\theta+\lambda+2)}\ \text{ and }\ q(x)=\frac{p_2(x)}{-(\theta+\lambda+2)}.$$

(2) Let $\{0,1\}\cap(\alpha,\omega)=\{1\}$ or $\{0,1\}\cap(\alpha,\omega)=\{0\}$, that is, $0\leqslant\alpha<1<\omega\leqslant\infty$ or $-\infty\leqslant\alpha<0<\omega\leqslant 1$. Clearly the map $x\mapsto 1-x$ translates the second case to the first one and leaves p_2 unchanged; thus, it suffices to consider only the first case. If $0<\alpha<1<\omega<\infty$ it is easily seen that (4.9) fails for all choices of $(C_2,C_3)\in\{(+,+),(+,0),(0,+)\}$, where $(C_2,C_3)=(+,0)$ means $C_2>0$, $C_3=0$, etc. If $\alpha=0$ and $1<\omega<\infty$ then (4.9) fails for all choices of $(C_2,C_3)\in\{(+,+),(0,+)\}$, while it is satisfied when $C_2>0$ and $C_3=0$. Similarly, if $0<\alpha<1$ and $\omega=\infty$ then (4.9) fails for all choices of $(C_2,C_3)\in\{(+,+),(+,0)\}$, while it is satisfied when $C_2=0$ and $C_3>0$. Finally, if $\alpha=0$ and $\omega=\infty$ then (4.9) is satisfied for all choices of $(C_2,C_3)\in\{(0,+),(+,0)\}$, while $C_2>0$, $C_3>0$ is not a permissible choice because f is not integrable. Therefore, the two distinct situations where orthogonality can be verified are given by

$$f_1(x)=\begin{cases}C_2x^A(1-x)^B, & 0<x<1,\\ 0, & 1\leqslant x<\omega,\end{cases}\qquad\text{and}\qquad f_2(x)=\begin{cases}0, & \alpha<x\leqslant 1,\\ C_3x^A(x-1)^B, & 1<x<\infty,\end{cases}$$

where $C_2>0$, $A>-1$, $B>1$ and $1<\omega\leqslant\infty$ for f_1; $C_3>0$, $B>1$, $A+B<-4$ and $0\leqslant\alpha<1$ for f_2. Now it is easily seen that both f_1 and f_2 belong to the Integrated Pearson family. Specifically, Proposition 3.3 shows that $f_1\sim\mathrm{IP}(\mu;q)$ with

$$\mu=\frac{A+1}{A+B+2}\quad\text{and}\quad q(x)=\frac{p_2(x)}{A+B+2}=\frac{x(1-x)}{A+B+2},\tag{4.11}$$

while $f_2 \sim \mathrm{IP}(\mu;q)$ with

$$\mu = \frac{A+1}{A+B+2} = 1 + \frac{B+1}{-A-B-2} \quad \text{and} \quad q(x) = \frac{p_2(x)}{A+B+2} = \frac{x(x-1)}{-A-B-2}. \tag{4.12}$$

(3) Let $\{0,1\} \subseteq (\alpha,\omega)$, that is, $-\infty \leqslant \alpha < 0 < 1 < \omega \leqslant \infty$. We have to study the following cases: (3a): $\alpha = -\infty$, $\omega = \infty$; (3b): $-\infty < \alpha < 0$, $\omega = \infty$; (3b$'$): $\alpha = -\infty$, $1 < \omega < \infty$; (3c): $-\infty < \alpha < 0$, $1 < \omega < \infty$. Clearly the map $x \mapsto 1-x$ translates the case (3b$'$) to (3b) and leaves p_2 unchanged; thus, it suffices to consider only the cases (3a), (3b) and (3c).

Assume first (3a). If $(C_1,C_2,C_3) \in \{(+,+,+),(+,0,+),(0,+,+)\}$ (where, e.g., $(C_1,C_2,C_3) = (+,0,+)$ means $C_1 > 0$, $C_2 = 0$, $C_3 > 0$ etc.) it follows that $A > 1$ and $B > 1$ and, thus, f fails to be integrable (at a neighborhood of $+\infty$). The case $(C_1,C_2,C_3) = (+,+,0)$ is equivalent to $(C_1,C_2,C_3) = (0,+,+)$ (by the map $x \mapsto 1-x$) and, again, f fails to be integrable. By the same map, the cases $(+,0,0)$ and $(0,0,+)$ are also equivalent. Assuming, e.g., $(C_1,C_2,C_3) = (0,0,+)$ it is easily seen that $B > 1$, $A+B < -4$ are necessary and sufficient for f being integrable, differentiable at 0 and 1 and with finite third moment. In this case both (4.9) and (4.8) are satisfied so that the system $\{h_0,h_1,h_2\}$ is indeed orthogonal. Finally, if we assume that $(C_1,C_2,C_3) = (0,+,0)$ then, necessarily, $A > 1$, $B > 1$ (for differentiability of f at 0 and 1) and it follows that the system $\{h_0,h_1,h_2\}$ is indeed orthogonal, since both (4.9) and (4.8) are satisfied.

Next, assume (3b). If $(C_1,C_2,C_3) \in \{(+,+,+),(+,0,+),(0,+,+)\}$ it follows that $A > 1$ and $B > 1$ and, thus, f fails to be integrable. If $(C_1,C_2,C_3) = (+,+,0)$ then $A > 1$, $B > 1$ and (4.9) fails. Also, if $(C_1,C_2,C_3) = (+,0,0)$ then $B > 1$ and (4.9) again fails. Assuming $(C_1,C_2,C_3) = (0,0,+)$ it is easily seen that $B > 1$, $A+B < -4$ are necessary and sufficient for f being integrable, differentiable at 0 and 1 and with finite third moment. In this case both (4.9) and (4.8) are satisfied so that the system $\{h_0,h_1,h_2\}$ is indeed orthogonal. Finally, if we assume that $(C_1,C_2,C_3) = (0,+,0)$ then, necessarily, $A > 1$, $B > 1$ (for differentiability of f at 0 and 1) and it follows that the system $\{h_0,h_1,h_2\}$ is indeed orthogonal, since both (4.9) and (4.8) are satisfied.

Finally, assume (3c). If $(C_1,C_2,C_3) \in \{(+,+,+),(+,0,+),(0,+,+),(+,+,0)\}$ it follows that $A > 1$ and $B > 1$ and (4.9) fails. By the map $x \mapsto 1-x$ it is easily seen that the cases $(+,0,0)$ and $(0,0,+)$ are equivalent. Assuming, e.g., $(C_1,C_2,C_3) = (0,0,+)$ it is easily seen that $B > 1$ is necessary and sufficient for f being integrable, differentiable at 0 and 1 and with finite third moment; but then, (4.9) fails. Finally, if we assume that $(C_1,C_2,C_3) = (0,+,0)$ then, necessarily, $A > 1$, $B > 1$ (for differentiability of f at 0 and 1) and it follows that the system $\{h_0,h_1,h_2\}$ is indeed orthogonal, since both (4.9) and (4.8) are satisfied.

Therefore, the two distinct situations where orthogonality can be verified are given by

$$f_1(x) = \begin{cases} 0, & \alpha < x \leqslant 0, \\ C_2 x^A (1-x)^B, & 0 < x < 1, \\ 0, & 1 \leqslant x < \omega, \end{cases} \quad \text{and} \quad f_2(x) = \begin{cases} 0, & \alpha < x \leqslant 1, \\ C_3 x^A (x-1)^B, & 1 < x < \infty, \end{cases}$$

where $C_2 > 0$, $A > 1$, $B > 1$ and $-\infty \leqslant \alpha < 0$, $1 < \omega \leqslant \infty$ for f_1; $C_3 > 0$, $B > 1$, $A+B < -4$ and $-\infty \leqslant \alpha < 0$ for f_2. Now it is easily seen that both f_1 and f_2 belong to the Integrated

Pearson family. Specifically, Proposition 3.3 shows that $f_1 \sim \mathrm{IP}(\mu;q)$ with μ and q as in (4.11), while $f_2 \sim \mathrm{IP}(\mu;q)$ with μ and q as in (4.12).

Next, assume that $\deg(p_2)=1$ and, without loss of generality (by using an affine map) we shall further assume that $p_2=x$. If $p_1=a_0+a_1x$, the general solution of (4.1) is

$$f(x)=\begin{cases} C_1x^{a_0}e^{a_1x} & \text{if } x<0,\\ C_2x^{a_0}e^{a_1x} & \text{if } x>0,\end{cases}$$

where a_0 and a_1 are arbitrary parameters and $C_1,C_2\geqslant 0$ are arbitrary constants, not both zero. The restrictions on a_0 and a_1 depend on the interval (α,ω) that we consider and the positivity or vanishing of each branch; they have to be chosen in such a way that the resulting function is differentiable and integrable in (α,ω). Assuming that $0<\alpha<\omega<\infty$ we readily see that any values of $a_0,a_1\in\mathbb{R}$ are admissible but (4.9) fails. If $0<\alpha<\omega=\infty$ then either $a_1=0$ and $a_0<-3$ (for finiteness of the third moment) or $a_1<0$ and $a_0\in\mathbb{R}$. In the first case both (4.9) and (4.8) are violated: the limits are unequal although

$$\int_\alpha^\infty h_k(x)h_m(x)f(x)\mathrm{d}x=0 \ \text{ for } \ k\neq m,\ \ k,m\in\{0,1,2\},$$

because $h_1=h_2\equiv 0$. In the second case, (4.9) fails. If $\alpha=0<\omega<\infty$ then $a_0>-1$ and $a_1\in\mathbb{R}$; it follows that (4.9) fails. Finally, if $\alpha=0$ and $\omega=\infty$ then $a_0>-1$ and $a_1<0$. In this case both (4.9) and (4.8) are satisfied and the system $\{h_0,h_1,h_2\}$ is, indeed, orthogonal. Also we see that $\mathbb{E}h_1(X)=a_1\mu+a_0+1=0$ so that $p_1+p_2'=a_1x+a_0+1=-a_1(\mu-x)$. Now, from Proposition 3.3 it follows that

$$f\sim\mathrm{IP}(\mu;q) \ \text{ with } \ \mu=\frac{a_0+1}{-a_1} \ \text{ and } \ q(x)=\frac{x}{-a_1}=\frac{p_2(x)}{-a_1}. \tag{4.13}$$

By the map $x\mapsto -x$ we can transform the cases $-\infty\leqslant\alpha<\omega\leqslant 0$ to the previous ones, since $p_2=x$ is transformed to $p_2=-x$. It remains to investigate the cases $-\infty\leqslant\alpha<0<\omega\leqslant\infty$; then, necessarily, $a_0>1$. Assuming that $-\infty<\alpha<0<\omega<\infty$ it is easily seen that (4.9) fails for all choices of $(C_1,C_2)\in\{(+,+),(+,0),(0,+)\}$. Assuming that $\alpha=-\infty$, $\omega=\infty$ we see that for f to be integrable it is necessary and sufficient that $a_1<0$ if $C_2>0$ and $a_1>0$ if $C_1>0$; therefore, if $(C_1,C_2)=(+,+)$ then f is not integrable. The case $(C_1,C_2)=(+,0)$ is transformed (by $x\mapsto -x$) to $(C_1,C_2)=(0,+)$. In the last case we can see that $a_0>1$ and $a_1<0$ are necessary and sufficient for f to be differentiable (in $(\alpha,\omega)=\mathbb{R}$) and to have finite third moment. As before we can easily check that both (4.9) and (4.8) are satisfied, that $\{h_0,h_1,h_2\}$ is orthogonal and that $f\sim\mathrm{IP}(\mu;q)$ with μ and q as in (4.13). The map $x\mapsto -x$ shows that the last two cases, $\alpha=-\infty$, $0<\omega<\infty$, and $-\infty<\alpha<0$, $\omega=\infty$, are equivalent. By considering the second one we see that $a_0>1$ and $a_1<0$ are necessary and sufficient for f to be differentiable (in (α,∞)) and to have finite third moment. However, if $(C_1,C_2)\in\{(+,+),(+,0)\}$ it is easily seen that (4.9) is violated because the limits as $x\searrow\alpha$ are nonzero. In the remaining case $(C_1,C_2)=(0,+)$ we can easily check, as before, that both (4.9) and (4.8) are satisfied, that $\{h_0,h_1,h_2\}$ is orthogonal and that $f\sim\mathrm{IP}(\mu;q)$ with μ and q as in (4.13).

Finally, assume that $\deg(p_2)=0$ or, equivalently, $p_2\equiv 1$. Then, if $p_1=a_0+a_1x$, it follows that
$$f(x)=C\exp(a_0x+a_1x^2/2),\quad \alpha<x<\omega.$$
If the support (α,ω) is bounded then it is easily seen that (4.9) fails. The cases $-\infty<\alpha<\omega=\infty$ and $-\infty=\alpha<\omega<\infty$ are, obviously, equivalent (by the map $x\mapsto -x$, which leaves p_2 unchanged). Assuming that $-\infty<\alpha<\omega=\infty$ we see that either $a_1=0$, $a_0<0$ or $a_1<0$, $a_0\in\mathbb{R}$; in the first case both (4.9) and (4.8) fail, while (4.9) fails in the second one. Finally, in the last remaining case where $(\alpha,\omega)=\mathbb{R}$ we see that, necessarily, $a_1<0$. Then, for any value of $a_0\in\mathbb{R}$ we check that both (4.9) and (4.8) are satisfied so that $\{h_0,h_1,h_2\}$ is, indeed, orthogonal. Observe that, by assumption, $\mathbb{E}h_1(X)=a_1\mu+a_0=0$; thus, $p_1+p_2'=a_0+a_1x=-a_1(\mu-x)$. Proposition 3.3 shows that $f\sim \mathrm{IP}(\mu;q)$ with
$$\mu=\frac{a_0}{-a_1}\ \text{ and }\ q(x)=\frac{p_2(x)}{-a_1}=\frac{1}{-a_1};\ \text{ in fact, }\ f\sim N\left(a_0/(-a_1),(1/\sqrt{-a_1})^2\right).$$
This subsumes all possible cases and completes the proof. □

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 (α,ω). From Theorem 4.1 it follows that the function
$$P_k(x):=\frac{(-1)^k}{f(x)}\frac{\mathrm{d}^k}{\mathrm{d}x^k}[q^k(x)f(x)],\quad \alpha<x<\omega,\quad k=0,1,2,\ldots \tag{5.1}$$
is a polynomial with
$$\deg(P_k)\leqslant k\ \text{ and }\ \mathrm{lead}(P_k)=\prod_{j=k-1}^{2k-2}(1-j\delta):=c_k(\delta),\ \ k=0,1,2,\ldots\ . \tag{5.2}$$
Obviously $c_0(\delta):=1$, i.e. an empty product should be treated as one.

The polynomials P_k are special cases of the polynomials h_k defined by (4.2); in fact, $P_k=(-1)^kh_k$. They are particularly important because under natural moment conditions they are, indeed, orthogonal with respect to the density f; see, e.g., [9] (pp. 295–296), [14], [21], [3]. The orthogonality follows immediately from Theorems 4.2 and 4.3. Moreover, the polynomials P_k and their derivatives satisfy a number of useful properties that will be reviewed here. The first three are
$$\begin{aligned}&P_0(x)=1,\\&P_1(x)=x-\mu,\\&P_2(x)=(1-\delta)(1-2\delta)x^2-2(1-\delta)(\mu+\beta)x+\mu^2+\beta\mu-(1-2\delta)\gamma.\end{aligned} \tag{5.3}$$

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 ([3], pp. 515–516). Let $X \sim \mathrm{IP}(\mu;\delta,\beta,\gamma) \equiv \mathrm{IP}(\mu;q)$ with density f and support (α,ω). 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

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

is absolutely continuous in (α,ω) 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:

$$\mathbb{E}P_k(X)g(X) = \mathbb{E}q^k(X)g^{(k)}(X). \tag{5.4}$$

It should be noted that when we claim that $h:(\alpha,\omega) \to \mathbb{R}$ is an absolutely continuous function with a.s. derivative h' we mean 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 (α,ω) such that

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

COROLLARY 5.1 ([3], 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, $\delta < \frac{1}{2n-1}$. Then

$$\begin{aligned}\mathbb{E}[P_k(X)P_m(X)] = \delta_{k,m} k! \mathbb{E}q^k(X) \prod_{j=k-1}^{2k-2} (1-j\delta) = \delta_{k,m} k! c_k(\delta) \mathbb{E}q^k(X),\\ k,m \in \{0,1,\ldots,n\},\end{aligned} \tag{5.5}$$

where $\delta_{k,m}$ is Kronecker's delta and where an empty product should be treated as one.

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 $\delta \in [\frac{1}{2n-1}, \frac{1}{2n-2})$; in this case, however, $P_n \notin L^2(\mathbb{R},X)$ since 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 $\delta < \frac{1}{2n-1}$. Therefore, for each $k \in \{0,1,\ldots,n\}$ and for all $j \in \{k-1,\ldots,2k-2\}$, $1-j\delta > 0$ because $\{k-1,\ldots,2k-2\} \subseteq \{0,1,\ldots,2n-2\}$. Thus, $c_k(\delta) > 0$. Since $\mathbb{P}[q(X) > 0] = 1$, $\deg(q) \leqslant 2$ 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 the set $\{\phi_0,\phi_1,\ldots,\phi_n\} \subset L^2(\mathbb{R},X)$, where

$$\phi_k(x) := \frac{P_k(x)}{\left(k! c_k(\delta) \mathbb{E}q^k(X)\right)^{1/2}} = \frac{\frac{(-1)^k}{f(x)} \frac{\mathrm{d}^k}{\mathrm{d}x^k}[q^k(x)f(x)]}{\left(k! \mathbb{E}q^k(X) \prod_{j=k-1}^{2k-2}(1-j\delta)\right)^{1/2}}, \quad k = 0,1,\ldots,n, \tag{5.6}$$

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

$$\begin{aligned}\mathrm{lead}\,(\phi_k) := d_k(\mu;q) &= \left(\frac{\prod_{j=k-1}^{2k-2}(1-j\delta)}{k! \mathbb{E}q^k(X)}\right)^{1/2} \\ &= \left(\frac{c_k(\delta)}{k! \mathbb{E}q^k(X)}\right)^{1/2} > 0, \quad k = 0,1,\ldots,n.\end{aligned} \tag{5.7}$$

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 ± 1. It follows that the standardized Rodrigues polynomials ϕ_k of (5.6) are the unique orthonormal polynomials that can be defined for a density $f \sim \mathrm{IP}(\mu; \delta, \beta, \gamma)$, provided that $\mathrm{lead}(\phi_k) > 0$. Therefore, it is useful to express the L^2-norm of each P_k in terms of the parameters δ, β, γ and μ and, in view of (5.5) and (5.6), it remains to obtain an expression for $\mathbb{E}q^k(X)$. To this end, we first recall a definition from [20]; cf. [10].

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, $f \sim \mathrm{IP}(\mu; q) \equiv \mathrm{IP}(\mu; \delta, \beta, \gamma)$. 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

$$f^*(x) := \frac{q(x) f(x)}{\mathbb{E}q(X)}, \quad \alpha < x < \omega. \tag{5.8}$$

Since $P_1 = x - \mu$, setting $k = 1$ in the covariance identity (5.4) we get (see [7], [20])

$$\mathbb{E}[(X - \mu) g(X)] = \mathrm{Cov}\,[X, g(X)] = \mathbb{E}[q(X) g'(X)]. \tag{5.9}$$

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 (5.9) to the identity function $g(x) = x$ it is easily seen that $\mathbb{E}q(X) = \mathrm{Var} X = \sigma^2$, so that (cf. [10])

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

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 $J(X^*) = J(X) = (\alpha, \omega)$,

$$\mu^* = \frac{\mu + \beta}{1 - 2\delta}, \quad \text{and} \quad q^*(x) = \frac{q(x)}{1 - 2\delta}, \quad \alpha < x < \omega. \tag{5.10}$$

Proof. From Corollary 2.2 it follows that the assumption $\mathbb{E}|X|^3 < \infty$ is equivalent to $\delta < \frac{1}{2}$. Let $X^* \sim f^*(x) = q(x) f(x) / \mathbb{E}q(X) = q(x) f(x) / \sigma^2$, $\alpha < x < \omega$, where σ^2 is the variance of X. Then, it follows that

$$\mu^* = \mathbb{E}X^* = \frac{\mathbb{E}[X q(X)]}{\sigma^2}.$$

Define $P_1(x) = x - \mu$ and $P_2(x) = (x - \mu)^2 - (x - \mu) q'(x) - (1 - 2\delta) q(x)$. We have $\mathbb{E}P_1(X) = 0$ and $\mathbb{E}P_2(X) = \sigma^2 - \mathrm{Cov}\,[X, q'(X)] - (1 - 2\delta)\mathbb{E}q(X)$. Applying the covariance identity (5.9) to $g(x) = x$ and to $g(x) = q'(x)$ we see that $\mathbb{E}P_2(X) = 2\delta\sigma^2 - \mathbb{E}[q(X)q''(X)] = 2\delta\sigma^2 - 2\delta\mathbb{E}q(X) = 0$. Also,

$$\begin{aligned}\mathbb{E}[P_1(X)P_2(X)] &= \mathbb{E}(X-\mu)^3 - \mathbb{E}[(X-\mu)^2 q'(X)] - (1-2\delta)\mathbb{E}[(X-\mu)q(X)] \\ &= \mathrm{Cov}\,[X,(X-\mu)^2] - \mathrm{Cov}\,[X,(X-\mu)q'(X)] - (1-2\delta)\,\mathrm{Cov}\,[X,q(X)]\end{aligned}$$

and, once again, (5.9) shows that $\mathbb{E}[P_1(X)P_2(X)] = 0$. Now observe that

$$x = \left(\frac{1}{2(1-\delta)(1-2\delta)}P_2(x) + \frac{\mu+\beta}{1-2\delta}P_1(x)\right)' = g'(x), \quad \text{say},$$

so that

$$\begin{aligned}\mathbb{E}Xq(X) &= \mathbb{E}q(X)g'(X) = \mathrm{Cov}\,[X,g(X)] = \mathbb{E}(X-\mu)g(X) = \mathbb{E}P_1(X)g(X) \\ &= \frac{1}{2(1-\delta)(1-2\delta)}\,\mathbb{E}P_1(X)P_2(X) + \frac{\mu+\beta}{1-2\delta}\,\mathbb{E}P_1^2(X) \\ &= 0 + \frac{\mu+\beta}{1-2\delta}\,\mathbb{E}(X-\mu)^2 = \frac{\mu+\beta}{1-2\delta}\,\sigma^2.\end{aligned}$$

It follows that $\mu^* = \mathbb{E}[Xq(X)]/\sigma^2 = (\mu+\beta)/(1-2\delta)$.

It remains to show that $q^*(x) = q(x)/(1-2\delta)$ is the quadratic polynomial of X^*, i.e. that

$$\int_{-\infty}^{x}(\mu^*-t)f^*(t)\,\mathrm{d}t = \frac{1}{1-2\delta}\,q(x)f^*(x), \quad x\in\mathbb{R}.$$

Equivalently, it suffices to verify the identity

$$\int_{-\infty}^{x}\{\mu+\beta-(1-2\delta)t\}q(t)f(t)\,\mathrm{d}t = q^2(x)f(x), \quad x\in\mathbb{R}. \tag{5.11}$$

Since $f(x) = 0$ for $x\notin(\alpha,\omega)$ it follows that the l.h.s. of (5.11) equals to zero for $x\leqslant\alpha$ (if $\alpha > -\infty$). Also, if $\omega < \infty$ and $x \geqslant \omega$ then the l.h.s. of (5.11) is equal to $(\mu+\beta)\mathbb{E}q(X) - (1-2\delta)\mathbb{E}Xq(X) = (\mu+\beta)\sigma^2 - (1-2\delta)\frac{\mu+\beta}{1-2\delta}\sigma^2 = 0$. Thus, (5.11) takes the form $0 = 0$ whenever $x\notin(\alpha,\omega)$. For $x\in(\alpha,\omega)$ it is easily seen that

$$\begin{aligned}&\left(q^2(x)f(x) - \int_{-\infty}^{x}\{\mu+\beta-(1-2\delta)t\}q(t)f(t)\,\mathrm{d}t\right)' \\ &\quad= (q(x)\cdot q(x)f(x))' - \{\mu+\beta-(1-2\delta)x\}q(x)f(x) \\ &\quad= q'(x)q(x)f(x) + q(x)(\mu-x)f(x) - q(x)f(x)\{\mu+\beta-(1-2\delta)x\} \\ &\quad= q(x)f(x)\,[q'(x)+(\mu-x)-(\mu+\beta)+(1-2\delta)x] = q(x)f(x)\,[q'(x)-2\delta x-\beta] = 0.\end{aligned}$$

Thus, there exists a constant $c\in\mathbb{R}$ such that

$$\int_{-\infty}^{x}\{\mu+\beta-(1-2\delta)t\}q(t)f(t)\,\mathrm{d}t = q^2(x)f(x)+c, \quad \alpha<x<\omega. \tag{5.12}$$

Now observe that $\lim_{x\nearrow\omega}\int_{-\infty}^{x}\{\mu+\beta-(1-2\delta)t\}q(t)f(t)\,\mathrm{d}t = \lim_{x\nearrow\omega}q^2(x)f(x) = 0$. Indeed, the first limit follows from dominated convergence and the fact that $\mathbb{E}q(X) =$

σ^2 and $\mathbb{E}[Xq(X)] = (\mu+\beta)\sigma^2/(1-2\delta)$, while the second one is obvious when $\omega < \infty$ because $q(\omega) = 0$ and $q(x)f(x) \to \mathbb{E}(\mu - X) = 0$ as $x \nearrow \omega$. Finally, if $\omega = \infty$ we have $q(x)f(x) = o(x^{-2})$ as $x \to \infty$ because $\mathbb{E}|X|^3 < \infty$ and for large enough x,

$$x^2 q(x)f(x) = x^2 \int_x^\infty (t-\mu)f(t)\mathrm{d}t \leqslant \int_x^\infty t^2(t-\mu)f(t)\mathrm{d}t \to 0, \ \text{ as } \ x \to \infty,$$

by dominated convergence. This shows that $\lim_{x\nearrow\omega} q^2(x)f(x) = 0$ in all cases. Therefore, taking limits as $x \nearrow \omega$ in (5.12) we conclude that $c = 0$ and (5.11) follows. □

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. $\delta < \frac{1}{2n}$) for some fixed $n \in \{0,1,\ldots\}$. Define the random variable X_k with density f_k given by

$$f_k(x) := \frac{q^k(x)f(x)}{\mathbb{E}q^k(X)}, \quad \alpha < x < \omega, \quad k = 0,1,\ldots,n. \tag{5.13}$$

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

$$\mu_k = \frac{\mu+k\beta}{1-2k\delta}, \ \text{ and } \ q_k(x) = \frac{q(x)}{1-2k\delta}, \ \alpha < x < \omega, \ k = 0,1,\ldots,n. \tag{5.14}$$

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, $n \geqslant 1$) the assertion 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 $k \leqslant n-1$. 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

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

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

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

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 $\delta < \frac{1}{2n+1}$) then for each $k \in \{0,1,\ldots,n\}$,

$$\sigma_k^2 := \mathrm{Var} X_k = \mathbb{E} q_k(X_k) = \frac{q(\frac{\mu+k\beta}{1-2k\delta})}{1-(2k+1)\delta}, \tag{5.15}$$

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

$$\sigma^2 := \mathrm{Var} X = \mathbb{E} q(X) = \frac{q(\mu)}{1-\delta}. \tag{5.16}$$

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 $\delta_k = \frac{\delta}{1-2k\delta} < 1$ because $\delta < \frac{1}{2n+1} \leqslant \frac{1}{2k+1}$. Note that it suffices to show only (5.16). Indeed, since $X_k \sim \mathrm{IP}(\mu_k;q_k)$ it follows from (5.9) (applied to the random variable X_k and to the function $g(x) = x$) that $\sigma_k^2 = \mathrm{Var} X_k = \mathbb{E} q_k(X_k)$. On the other hand, if we manage to show that $\mathrm{Var} X = \frac{q(\mu)}{1-\delta}$ for any $X \sim \mathrm{IP}(\mu;q)$ with $\delta < 1$ then, by (5.16) applied to X_k, we get

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

Since

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

(5.16) yields the identity (5.15) as follows:

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

It remains to verify that $\mathrm{Var} X = \sigma^2 = \frac{q(\mu)}{1-\delta}$ 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 (5.16). □

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

$$A_k = A_k(\mu;q) := \mathbb{E} q^k(X) = \frac{\prod_{j=0}^{k-1}(1-2j\delta)}{\prod_{j=0}^{k-1}(1-(2j+1)\delta)} \prod_{j=0}^{k-1} q\left(\frac{\mu+j\beta}{1-2j\delta}\right). \tag{5.17}$$

Proof. Observe that

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

where $A_0 := 1$, $q_0 = q$, $X_0 = X$. Multiplying these relations for $j = 0,1,\ldots,k-1$ and using (5.15) we get (5.17). □

REMARK 5.1. (a) It is important to note that the identity (5.4) enables a convenient calculation of the Fourier coefficients of any smooth enough function g with $\operatorname{Var} g(X) < \infty$ (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 $c_k = \mathbb{E}\phi_k(X)g(X)$ are given by $c_0 = \mathbb{E}g(X)$ and

$$c_k = \frac{\mathbb{E}q^k(X)g^{(k)}(X)}{(k!c_k(\delta)A_k(\mu;q))^{1/2}}, \quad k = 1,2,\ldots,n, \tag{5.18}$$

where $c_k(\delta)$ and $A_k(\mu;q)$ are given by (5.2) and (5.17), 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 $\delta \leqslant 0$ (i.e. if X is of Normal, Gamma or Beta-type) then $\mathbb{E}|X|^n < \infty$ for all n. Moreover, since there exist an $\varepsilon > 0$ such that $\mathbb{E}e^{tX} < \infty$ for $|t| < \varepsilon$ it follows that the corresponding polynomials $\{\phi_k\}_{k=0}^{\infty}$, given by (5.6), form a complete orthonormal system in $L^2(\mathbb{R};X)$; see, e.g., [24], [6], [3]. Therefore, for smooth enough g with $\operatorname{Var} g(X) < \infty$ and $\mathbb{E}q^k(X)|g^{(k)}(X)| < \infty$ for all $k \geqslant 1$, the Fourier coefficients are given by

$$c_k = \mathbb{E}\phi_k(X)g(X) = \frac{\mathbb{E}q^k(X)g^{(k)}(X)}{(k!c_k(\delta)A_k(\mu;q))^{1/2}}, \quad k = 0,1,2,\ldots, \tag{5.19}$$

and the variance of g can be calculated as (see [3], Theorem 5.1, pp. 522–523)

$$\operatorname{Var} g(X) = \sum_{k=1}^{\infty} \frac{\mathbb{E}^2 q^k(X)g^{(k)}(X)}{k!c_k(\delta)A_k(\mu;q)}. \tag{5.20}$$

Furthermore, the completeness of the Rodrigues polynomials (when $X \sim \mathrm{IP}(\mu;\delta,\beta,\gamma)$ and $\delta \leqslant 0$) enables one to write ([3], Theorem 5.2, p. 523)

$$\operatorname{Cov}[g_1(X),g_2(X)] = \sum_{k=1}^{\infty} \frac{\mathbb{E}[q^k(X)g_1^{(k)}(X)]\mathbb{E}[q^k(X)g_2^{(k)}(X)]}{k!c_k(\delta)A_k(\mu;q)}, \tag{5.21}$$

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 $k \geqslant 1$. The important thing in (5.20) and (5.21) is that we do not need explicit forms for the polynomials; in view of (5.2) and (5.17), 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, (5.20) yields the formulae

$$\operatorname{Var} g(X) = \sum_{k=1}^{\infty} \frac{\sigma^{2k}}{k!}\mathbb{E}^2 g^{(k)}(X), \qquad \text{if } X \sim N(\mu,\sigma^2), \tag{5.22}$$

$$\operatorname{Var} g(X) = \sum_{k=1}^{\infty} \frac{\Gamma(a)}{k!\Gamma(a+k)}\mathbb{E}^2 X^k g^{(k)}(X), \qquad \text{if } X \sim \Gamma(a,\lambda), \tag{5.23}$$

$$\operatorname{Var} g(X) = \sum_{k=1}^{\infty} \frac{(a+b+2k-1)\Gamma(a)\Gamma(b)\Gamma(a+b+k-1)}{k!\Gamma(a+b)\Gamma(a+k)\Gamma(b+k)}\mathbb{E}^2 X^k(1-X)^k g^{(k)}(X), \quad \text{if } X \sim B(a,b). \tag{5.24}$$

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 $n \geqslant 2$, i.e. with $\delta<\frac{1}{2n-1}$. 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

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

and has support (α,ω). Since $\delta<\frac{1}{2n-1}$ is equivalent to $\delta_1=\frac{\delta}{1-2\delta}<\frac{1}{2n-3}$ we conclude that $\mathbb{E}|X_1|^{2n-2}<\infty$ and, in particular, $\mathrm{Var}X_1<\infty$. 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 $f_1(x)=q(x)f(x)/\mathbb{E}q(X)$)

$$P_{k,1}(x) := \frac{(-1)^k}{f_1(x)}\frac{\mathrm{d}^k}{\mathrm{d}x^k}[q_1^k(x)f_1(x)] = \frac{(-1)^k}{(1-2\delta)^k q(x)f(x)}\frac{\mathrm{d}^k}{\mathrm{d}x^k}[q^{k+1}(x)f(x)], \quad \alpha<x<\omega,\ \ k=0,1,\ldots,n-1. \tag{5.25}$$

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 reobtain 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 $n\geqslant 1$ then the polynomials P_k of (5.1) and $P_{k,1}$ of (5.25) are related through

$$P'_{k+1}(x)=C_k(\delta)P_{k,1}(x),\ \ k=0,1,\ldots,n-1, \quad \text{where}\ \ C_k(\delta):=(k+1)(1-k\delta)(1-2\delta)^k. \tag{5.26}$$

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

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

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 $m+k+1\leqslant 2n-2$ 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,

$$\mathbb{E}P_{k,1}(X_1)P_{m,1}(X_1) = 0 \;\text{ for }\; k,m \in \{0,1,\ldots,n-1\} \;\text{ with }\; k \neq m.$$

Since $\deg(P'_{k+1}) = \deg(P_{k,1}) = k$, $k = 0,1,\ldots,n-1$, the uniqueness of the orthogonal polynomial system implies that there exist constants $C_k \neq 0$ such that $P'_{k+1}(x) = C_k P_{k,1}(x)$. Equating the leading coefficients we obtain $\text{lead}\,(P'_{k+1}) = C_k \text{lead}\,(P_{k,1})$, that is (see (5.2)),

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

REMARK 5.2. We note that the recurrence (5.26) 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 $k \geqslant 0$; thus, orthogonality is not, at all, needed for deriving it. Specifically, if $p_1 = a_0 + a_1 x$, $p_2 = b_0 + b_1 x + b_2 x^2$, and if h_k are the polynomials in (4.2) and $h_{k,1}$ are the polynomials given by

$$h_{k,1}(x) := \frac{1}{p_2(x)f(x)}\frac{\mathrm{d}^k}{\mathrm{d}x^k}[p_2^{k+1}(x)f(x)],$$

then, with Beale's notation, $h_{k+1}(x) = P_{k+1}(k+1,x)$ and $h_{k,1}(x) = P_k(k+1,x)$; see also [12], p. 401. Therefore, Beale's identity is equivalent to (cf. [4], eq. (2), p. 207)

$$h'_{k+1}(x) = (k+1)[a_1 + (k+2)b_2]h_{k,1}(x). \tag{5.27}$$

On the other hand, the current definition of P_k and $P_{k,1}$ can be translated to Beale's notation as follows: Since $X \sim \text{IP}(\mu;\delta,\beta,\gamma) \equiv \text{IP}(\mu;q)$ we have from Proposition 2.1 that $f'/f = p_1/p_2$ with $p_2 = q$ and $p_1 = \mu - x - q'$, that is, $a_0 = \mu - \beta$, $a_1 = -(1+2\delta)$, $b_0 = \gamma$, $b_1 = \beta$ and $b_2 = \delta$. Furthermore,

$$P_{k+1}(x) = \frac{(-1)^{k+1}}{f(x)}\frac{\mathrm{d}^{k+1}}{\mathrm{d}x^{k+1}}[q^{k+1}(x)f(x)] = \frac{(-1)^{k+1}}{f(x)}\frac{\mathrm{d}^{k+1}}{\mathrm{d}x^{k+1}}[p_2^{k+1}(x)f(x)] = (-1)^{k+1}h_{k+1}(x)$$

and

$$P_{k,1}(x) = \frac{(-1)^k}{f_1(x)}\frac{\mathrm{d}^k}{\mathrm{d}x^k}[q_1^k(x)f_1(x)] = \frac{(-1)^k}{q(x)f(x)}\frac{\mathrm{d}^k}{\mathrm{d}x^k}\left[\frac{q^k(x)}{(1-2\delta)^k}q(x)f(x)\right] = \frac{(-1)^k}{(1-2\delta)^k}h_{k,1}(x).$$

Thus, $h_{k+1} = (-1)^{k+1} P_{k+1}$, $h_{k,1} = (-1)^k (1-2\delta)^k P_{k,1}$ and (5.27) yields

$$(-1)^{k+1} P'_{k+1} = (k+1)[-(1+2\delta)+(k+2)\delta](-1)^k(1-2\delta)^k P_{k,1}.$$

That is, $P'_{k+1} = (k+1)[(1+2\delta)-(k+2)\delta](1-2\delta)^k P_{k,1} = (k+1)(1-k\delta)(1-2\delta)^k P_{k,1}$, which shows that (5.26) holds for all $k \in \{0,1,\ldots\}$.

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 $n \geqslant 1$ (i.e. $\delta < \frac{1}{2n-1}$) then

$$P^{(m)}_{k+m}(x) = C^{(m)}_k(\delta) P_{k,m}(x), \quad m = 1,2,\ldots,n, \quad k = 0,1,\ldots,n-m, \tag{5.28}$$

where

$$C^{(m)}_k(\delta) := \frac{(k+m)!}{k!}(1-2m\delta)^k \prod_{j=k+m-1}^{k+2m-2} (1-j\delta). \tag{5.29}$$

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 $X_m \sim \mathrm{IP}(\mu_m; q_m)$ of Theorem 5.2, i.e.,

$$P_{k,m}(x) := \frac{(-1)^k}{f_m(x)} \frac{\mathrm{d}^k}{\mathrm{d}x^k}[q^k_m(x) f_m(x)] = \frac{(-1)^k}{(1-2m\delta)^k q^m(x) f(x)} \frac{\mathrm{d}^k}{\mathrm{d}x^k}[q^{k+m}(x) f(x)], \quad \alpha < x < \omega, \quad k = 0,1,\ldots,n-m. \tag{5.30}$$

Proof. Apply first Lemma 5.2 to get

$$P'_{k+m} = P'_{(k+m-1)+1} = (k+m)(1-(k+m-1)\delta)(1-2\delta)^{k+m-1} P_{k+m-1,1}.$$

Now, since $P_{k+m-1,1}$ are the Rodrigues polynomials of f_1 we can apply again Lemma 5.2 to X_1 with $\delta_1 = \frac{\delta}{1-2\delta}$. It follows that

$$P'_{k+m-1,1} = P'_{(k+m-2)+1,1} = (k+m-1)(1-(k+m-2)\delta_1)(1-2\delta_1)^{k+m-2} P_{k+m-2,2}.$$

Combining the above equations we see that

$$\begin{aligned} P''_{k+m} &= (k+m)(1-(k+m-1)\delta)(1-2\delta)^{k+m-1} P'_{k+m-1,1} \\ &= (k+m)(k+m-1)(1-(k+m-1)\delta)(1-(k+m-2)\delta_1) \\ &\qquad \times (1-2\delta)^{k+m-1}(1-2\delta_1)^{k+m-2} P_{k+m-2,2}. \end{aligned}$$

By the same argument it follows that for any $j \in \{0,1,\ldots,m-1\}$,

$$P'_{k+m-j,j} = P'_{(k+m-j-1)+1,1} = (k+m-j)(1-(k+m-j-1)\delta_j)(1-2\delta_j)^{k+m-j-1} P_{k+m-j-1,j+1},$$

where $\delta_j = \delta/(1-2j\delta)$. Thus, we can easily show, using (finite) induction on s, that for any $s \in \{1,2,\ldots,m\}$,

$$P_{k+m}^{(s)} = \left\{ \left(\prod_{j=0}^{s-1} (k+m-j) \right) \left(\prod_{j=0}^{s-1} (1-(k+m-j-1)\delta_j) \right) \left(\prod_{j=0}^{s-1} (1-2\delta_j)^{k+m-j-1} \right) \right\} P_{k+m-s,s}.$$

Setting $s = m$ it follows that (5.28) is satisfied with

$$C_k^{(m)}(\delta) = \left(\prod_{j=0}^{m-1} (k+m-j) \right) \left(\prod_{j=0}^{m-1} (1-(k+m-j-1)\delta_j) \right) \left(\prod_{j=0}^{m-1} (1-2\delta_j)^{k+m-j-1} \right).$$

Now it suffices to observe that $\prod_{j=0}^{m-1}(k+m-j) = \frac{(k+m)!}{k!}$, that

$$\begin{aligned}\prod_{j=0}^{m-1} (1-(k+m-j-1)\delta_j) &= \prod_{j=0}^{m-1} \left(1-(k+m-j-1)\frac{\delta}{1-2j\delta} \right) \\ &= \frac{\prod_{j=0}^{m-1}(1-(k+m+j-1)\delta)}{\prod_{j=0}^{m-1}(1-2j\delta)} = \frac{\prod_{j=k+m-1}^{k+2m-2}(1-j\delta)}{\prod_{j=0}^{m-1}(1-2j\delta)},\end{aligned}$$

and that

$$\begin{aligned}\prod_{j=0}^{m-1}(1-2\delta_j)^{k+m-j-1} &= \prod_{j=0}^{m-1}\left(1-2\frac{\delta}{1-2j\delta}\right)^{k+m-j-1} = \prod_{j=0}^{m-1}\left(\frac{1-2(j+1)\delta}{1-2j\delta}\right)^{k+m-j-1} \\ &= \frac{\prod_{j=0}^{m-1}(1-2(j+1)\delta)^{k+m-j-1}}{\prod_{j=0}^{m-1}(1-2j\delta)^{k+m-j-1}} = \frac{\prod_{j=1}^{m}(1-2j\delta)^{k+m-j}}{\prod_{j=1}^{m-1}(1-2j\delta)^{k+m-j-1}} \\ &= (1-2m\delta)^k \frac{\prod_{j=1}^{m-1}(1-2j\delta)^{k+m-j}}{\prod_{j=1}^{m-1}(1-2j\delta)^{k+m-j-1}} \\ &= (1-2m\delta)^k \prod_{j=1}^{m-1}(1-2j\delta).\end{aligned}$$

□

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 constant C_k. Arguing as in the proof of Lemma 5.2 we see that C_k can be derived from the corresponding leading coefficients. Indeed, since $\text{lead}(P_{k+m}^{(m)}) = C_k \text{lead}(P_{k,m})$, we get, in view of (5.2), that

$$\begin{aligned}C_k &= \frac{\text{lead}(P_{k+m}^{(m)})}{\text{lead}(P_{k,m})} = \frac{\frac{(k+m)!}{k!}\text{lead}(P_{k+m})}{\text{lead}(P_{k,m})} = \frac{\frac{(k+m)!}{k!}c_{k+m}(\delta)}{c_k(\delta_m)} = \frac{\frac{(k+m)!}{k!}\prod_{j=k+m-1}^{2k+2m-2}(1-j\delta)}{\prod_{j=k-1}^{2k-2}(1-j\delta_m)} \\ &= \frac{\frac{(k+m)!}{k!}\prod_{j=k+m-1}^{2k+2m-2}(1-j\delta)}{\prod_{j=k-1}^{2k-2}(1-j\frac{\delta}{1-2m\delta})} = \frac{\frac{(k+m)!}{k!}(1-2m\delta)^k\prod_{j=k+m-1}^{2k+2m-2}(1-j\delta)}{\prod_{j=k+2m-1}^{2k+2m-2}(1-j\delta)} \\ &= \frac{(k+m)!}{k!}(1-2m\delta)^k \prod_{j=k+m-1}^{k+2m-2}(1-j\delta).\end{aligned}$$

(b) We note that the recurrence (5.28) is contained in Beale (1937), eq. (4), p. 207, although it is stated in a quite different notation there. Specifically, if $p_1 = a_0 + a_1 x$, $p_2 = b_0 + b_1 x + b_2 x^2$, and if h_k are the polynomials in (4.2) and $h_{k,m}$ are the polynomials given by

$$h_{k,m}(x) := \frac{1}{p_2^m(x) f(x)} \frac{\mathrm{d}^k}{\mathrm{d}x^k}[p_2^{k+m}(x) f(x)],$$

then, with Beale's notation, $h_{k+m}(x) = P_{k+m}(k+m,x)$ and $h_{k,m}(x) = P_k(k+m,x)$. Therefore, putting $q \to m$, $k \to k+m$, $n \to k+m-1$, $N' \to -(1+2\delta)$ and $D'' \to 2\delta$ in eq. (4) of [4], we get

$$\begin{aligned} h_{k+m}^{(m)}(x) &= \left(\prod_{i=0}^{m-1} (k+m-i)((k+m+i-1)\delta - 1) \right) h_{k,m}(x) \\ &= (-1)^m \frac{(k+m)!}{k!} \left(\prod_{j=k+m-1}^{k+2m-2} (1 - j\delta) \right) h_{k,m}(x), \qquad k = 0,1,2,\ldots . \end{aligned} \tag{5.31}$$

On the other hand it is easy to see that $P_{k+m}(x) = (-1)^{k+m} h_{k+m}(x)$ and, with $p_2 = q$,

$$P_{k,m}(x) = \frac{(-1)^k}{f_m(x)} \frac{\mathrm{d}^k}{\mathrm{d}x^k}[q_m^k(x) f_m(x)] = \frac{(-1)^k}{p_2^m(x) f(x)} \frac{\mathrm{d}^k}{\mathrm{d}x^k} \left[\frac{p_2^k(x)}{(1-2m\delta)^k} p_2^m(x) f(x) \right] = \frac{(-1)^k h_{k,m}(x)}{(1-2m\delta)^k}.$$

Thus, $h_{k+m} = (-1)^{k+m} P_{k+m}$, $h_{k,m} = (-1)^k (1-2m\delta)^k P_{k,m}$, and (5.31) becomes

$$(-1)^{k+m} P_{k+m}^{(m)} = (-1)^m \frac{(k+m)!}{k!} \left(\prod_{j=k+m-1}^{k+2m-2} (1 - j\delta) \right) (-1)^k (1-2m\delta)^k P_{k,m}, \quad k = 0,1,\ldots;$$

equivalently, $P_{k+m}^{(m)} = \frac{(k+m)!}{k!} (1-2m\delta)^k \left(\prod_{j=k+m-1}^{k+2m-2} (1-j\delta) \right) P_{k,m}$, which shows that (5.28) holds for all $k \in \{0,1,\ldots\}$.

(c) Krall [16], [17] characterizes the Pearson system from 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 contains the main interest 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 (α,ω), and assume that $\mathbb{E}|X|^{2n} < \infty$ for some fixed $n \geqslant 1$ (equivalently, $\delta < \frac{1}{2n-1}$). Let $\{\phi_k\}_{k=0}^n$ be the orthonormal polynomials associated with X (with $\mathrm{lead}(\phi_k) > 0$ for all k; see (5.6), (5.7)), fix a number $m \in \{0,1,\ldots,n\}$, and consider the corresponding orthonormal polynomials $\{\phi_{k,m}\}_{k=0}^{n-m}$, with $\mathrm{lead}(\phi_{k,m}) > 0$, associated with

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

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

$$\phi_{k+m}^{(m)}(x) = \nu_k^{(m)} \phi_{k,m}(x), \quad \alpha < x < \omega, \quad k = 0,1,\ldots,n-m. \tag{5.32}$$

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

$$\nu_k^{(m)} = \nu_k^{(m)}(\mu;q) := \left\{ \frac{\frac{(k+m)!}{k!} \prod_{j=k+m-1}^{k+2m-2}(1-j\delta)}{A_m(\mu;q)} \right\}^{1/2}, \tag{5.33}$$

where $A_m(\mu;q) = \mathbb{E}q^m(X)$ is given by (5.17). In particular, setting $\sigma^2 = \operatorname{Var} X$ we have

$$\phi'_{k+1}(x) = \frac{\sqrt{(k+1)(1-k\delta)}}{\sigma} \phi_{k,1}(x) = \sqrt{\frac{(k+1)(1-\delta)(1-k\delta)}{q(\mu)}} \phi_{k,1}(x), \ k = 0,1,\ldots,n-1. \tag{5.34}$$

Proof. Observe that

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

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

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

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,

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

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

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

Combining the preceding relations we obtain

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

and the proof is complete. □

ACKNOWLEDGEMENTS. We would like to thank M.C. Jones and H. Papageorgiou for a number of suggestions and helpful comments that improved the presentation.

References

[1] AFENDRAS, G. and PAPADATOS, N. (2011). On matrix variance inequalities. *J. Statist. Plann. Inference* **141** 3628–3631.

[2] AFENDRAS, G., PAPADATOS, N. and PAPATHANASIOU, V. (2007). The discrete Mohr and Noll inequality with applications to variance bounds. *Sankhyā* **69** 162–189.

[3] AFENDRAS, G., PAPADATOS, N. and PAPATHANASIOU, V. (2011). An extended Stein-type covariance identity for the Pearson family, with applications to lower variance bounds. *Bernoulli* **17**(2) 507–529.

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

[5] BEALE, F.S. (1941). On a certain class of orthogonal polynomials. *Ann. Math. Statist.* **12** 97–103. MR0003852

[6] BERG, C. and CHRISTENSEN, J.P.R. (1981). Density questions in the classical theory of moments. *Ann. Inst. Fourier* (Grenoble) **31** 99–114. MR0638619

[7] CACOULLOS, T. and PAPATHANASIOU, V. (1989). Characterizations of distributions by variance bounds. *Statist. Probab. Lett.* **7** 351-356. MR1001133

[8] CRAIG, C.C. (1936). A new exposition and chart for the Pearson system of frequency curves. *Ann. Math. Statist.* **7**(1) 16–28.

[9] DIACONIS, P. and ZABELL, S. (1991). Closed form summation for classical distributions: variations on a theme of De Moivre. *Statist. Science* **6** 284–302. MR1144242

[10] GOLDSTEIN, L. and REINERT, G. (1997). Stein's method and the zero-bias transformation with application to simple random sampling. *Ann. Appl. Probab.* **7** 935–952. MR1484792

[11] GOLDSTEIN, L. and REINERT, G. (2005). Distributional transformations, orthogonal polynomials and Stein characterizations. *J. Theoret. Probab.* **18** 237–260. MR2132278

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

[13] JOHNSON, N.L., KOTZ, S. and BALAKRISHNAN, N. (1994). *Continuous Univariate Distributions*, vol. 1, 2nd ed., Wiley, N.Y.

[14] JOHNSON, R.W. (1993). A note on variance bounds for a function of a Pearson variate. *Statist. Decisions* **11** 273–278. MR1257861

[15] KORWAR, R.M. (1991). On characterizations of distributions by mean absolute deviation and variance bounds. *Ann. Inst. Statist. Math.* **43** 287–295. MR1128869

[16] KRALL, H.L. (1936). On derivatives of orthogonal polynomials. *Bull. Amer. Math. Soc.* **42** 423–428. MR1563314

[17] KRALL, H.L. (1941). On derivatives of orthogonal polynomials, II. *Bull. Amer. Math. Soc.* **47** 261–264. MR0003854

[18] ORD, J.K. (1967). On a system of discrete distributions. *Biometrika* **54** 649–656.

[19] ORD, J.K. (1972). *Families of Frequency Distributions.* Griffin, London.

[20] PAPADATOS, N. and PAPATHANASIOU, V. (2001). Unified variance bounds and a Stein-type identity. In: *Probability and Statistical Models with Applications* (Ch.A. Charalambides, M.V. Koutras and N. Balakrishnan, Eds.), Chapman & Hall/CRC, New York, pp. 87-100.

[21] PAPATHANASIOU, V. (1995). A characterization of the Pearson system of distributions and the associated orthogonal polynomials. *Ann. Inst. Statist. Math.* **47** 171–176. MR1341214

[22] PEARSON, K. (1895). Contributions to the mathematical theory of evolution. *Philosophical Transactions of the Royal Society of London, Ser. A* **186** 343–414.

[23] PEARSON, K. (1924). On the mean error of frequency distributions. *Biometrika* **16** 198–200.

[24] RIESZ, M. (1923). Sur le problème des moments et le théorème de Parseval (in French). *Acta Litt. Ac. Sci.* (Szeged) **1** 209–225.