A CHARACTERIZATION OF
THE PEARSON SYSTEM OF DISTRIBUTIONS AND
THE ASSOCIATED ORTHOGONAL POLYNOMIALS

V. PAPATHANASIOU

Department of Mathematics, University of Athens, Panepistemiopolis, 15784 Athens, Greece

(Received July 5, 1993; revised April 18, 1994)

Abstract. A new derivation of the classical orthogonal polynomials is given
by using the w-function which appears in the variance bounds and some prop-
erties of the Pearson system of distributions. Also a characterization of the
Pearson system of distributions through some conditional moments is obtained
by using a result obtained by Johnson (1993) concerning this family.

Key words and phrases: Variance bounds, Pearson system, orthogonal poly-
nomials, characterization.

1. Introduction and summary

Papathanasiou (1989) obtained upper and lower bounds for the variance of a
function g(X) where X is a continuous random variable with E(X) = μ, Var(X) =
σ^2 and g absolutely continuous with continuous derivative g^{(2m+3)} of order 2m+3
(m ≥ 0). Specifically,

\begin{equation}
S(2m+1) ≤ \text{Var}[g(X)] ≤ S(2m),
\end{equation}

where S(n) is given by

\begin{equation}
S(n) = \sum_{\nu=0}^{n} \frac{(-1)^\nu}{\nu!(\nu+1)!} \int_{-\infty}^{+\infty} [g^{(\nu+1)}(t)]^2 a_\nu(t) dt
\end{equation}

and

\begin{equation}
a_\nu(t) = (-1)^\nu E[X - t]^{\nu+1} \int_{-\infty}^{t} (x - t)^\nu f(x) dx
\end{equation}

+ (-1)^{\nu+1} E[X - t]^{\nu} \int_{-\infty}^{t} (x - t)^{\nu+1} f(x) dx.

171
The upper (lower) bound is attained if and only if $g$ is a polynomial of degree at most $2m+1 \ (2m+2)$. If $m = 0$ we have the upper bound for the variance obtained by Cacoullos and Papathanasiou (C-P) (1985), namely,

\begin{equation}
\text{Var}[g(X)] \leq \int_{-\infty}^{+\infty} [g'(t)]^2 a_0(t) dt
\end{equation}

where

\begin{equation}
a_0(t) = \int_{-\infty}^{t} (\mu - x) f(x) dx
\end{equation}

and we set

\begin{equation}
\sigma^2 w(t) f(t) = \int_{-\infty}^{t} (\mu - x) f(x) dx
\end{equation}

if the support of $f$ is an interval.

The same $w$-function features also in the lower bound $\text{Var}[g(X)] \geq \sigma^2 E[w(X)g'(X)]$ (cf. Cacoullos and Papathanasiou (1989)) obtained from the identity

\begin{equation}
\text{Cov}(X, g(X)) = \sigma^2 E[w(X)g'(X)],
\end{equation}

where the $w$-function is defined by (1.5). Furthermore, the $w$-function characterizes the corresponding distribution of $X$ (C-P (1989)).

Korwar (1991) showed that the $w$-function is a quadratic if and only if $X$ is a Pearson random variable, that is, its density $f$ satisfies the differential equation

\begin{equation}
\frac{d \ln f}{dx} = \frac{\alpha - x}{\beta_0 + \beta_1 x + \beta_2 x^2}
\end{equation}

for some values of $\alpha, \beta_0, \beta_1, \beta_2$.

Johnson (1993) proved that if $X$ is a Pearson random variable such that

\begin{equation}
\lim_{x \to r} x^j w(x)f(x) = \lim_{x \to s} x^j w(x)f(x) = 0, \quad j = 0, 1, 2, \ldots, n
\end{equation}

where $(r, s)$ is the support of $f$, then the function $a_\nu(t)$, given by (1.3), can be written as follows:

\begin{equation}
a_\nu(t) = \nu! b \sigma^{2(\nu+1)}(w(t))^{\nu+1} f(t).
\end{equation}

Here $b = 1 / \prod_{j=0}^{n} (1 - jk)$, $k = \beta_2/(1 - 2\beta_2)$ and $w$ is the $w$-function, as given by (1.5), which for the Pearson system is quadratic and

\begin{equation}
\sigma^2 w(x) = \frac{\beta_0 + \beta_1 x + \beta_2 x^2}{1 - 2\beta_2}.
\end{equation}

In the present paper using inequality (1.1) and Johnson’s result (1.8) a characterization of the Pearson system of distributions is obtained through some conditional moments. Furthermore using the fact that the $w$-function, which appears in the variance bounds, for the Pearson family is quadratic (see Korwar (1991) or Johnson (1993)) a new derivation of the classical orthogonal polynomial is obtained.
2. Some identities and orthogonal polynomials

For our purpose we obtain an identity, interesting in itself.

**THEOREM 2.1.** Let $X$ be a Pearson random variable with density $f$ and support an interval $(r, s)$. If

$$\lim_{x \to r} x^j w(x) f(x) = \lim_{x \to s} x^j w(x) f(x) = 0, \quad j = 0, 1, 2, \ldots, n,$$

then

$$\text{Cov}(X^{n+1}, g(X)) = \sum_{\nu=0}^{n} (-1)^\nu c_{\nu,n} E[(g^{(\nu+1)}(X)X^{n-\nu}(w(X))^{\nu+1}]
$$

provided these moments exist and $g$ has a continuous derivative of order $n+2$ and

$$c_{\nu,n} = \binom{n+1}{\nu+1} b^2(\nu+1).$$

**PROOF.** Let $n = 2m$. We apply the second inequality in (1.1) for the function $x^{n+1} + \theta g(x)$ where $\theta$ is an arbitrary real number. Since (2.1) holds we use Johnson’s result. For $x^{n+1}$ (1.1) becomes equality and we have

$$\theta^2 \text{Var}[g(X)] + 2\theta \text{Cov}[X^{n+1}, g(X)]
\leq \theta^2 \sum_{\nu=0}^{n} \frac{(-1)^\nu}{(\nu + 1)!} b^2(\nu+1) E[(g^{(\nu+1)}(X))^2(w(X))^{\nu+1}]
+ 2\theta \sum_{\nu=0}^{n} \frac{(-1)^\nu}{(\nu + 1)!} b^2(\nu+1)(n + 1)_{\nu+1}
\cdot E[(w(X))^{\nu+1}X^{n-\nu}g^{(\nu+1)}(X)].$$

Since (2.3) holds for arbitrary $\theta$ we conclude (2.2). The case $n = 2m + 1$ can be treated similarly. \qed

**THEOREM 2.2.** Under the assumptions of Theorem 2.1 we have

$$X^{n+1} - \mu_{n+1} = -\sum_{\nu=0}^{n} c_{\nu,n} \frac{1}{\int dx^{n+1}} [x^{n-\nu}(w(x))^{\nu+1}f(x)]$$

where $\mu_{n+1} = E[X^{n+1}]$.

**PROOF.** In (2.2) take first $g(x) = \text{Re}[e^{itx}]$, where $i = \sqrt{-1}$, and then $g(X) = \text{Im}[e^{itx}]$. Integrating the RHS by parts $(\nu + 1)$-times $\nu = 0, 1, \ldots, n$, since (2.1) holds we have (see the proof of Theorem 2.3)

$$\lim_{x \to r} \frac{d^j}{dx^j} [x^{n-\nu} w(x) \nu+1 f(x)] = \lim_{x \to s} [x^{n-\nu} w(x) \nu+1 f(x)] = 0$$
\[ j = 0, 1, \ldots, \nu, \nu = 0, 1, \ldots, n. \] By the uniqueness of the Fourier transform the proof is complete. \( \square \)

Using the same arguments, identities (2.2) and (2.4) can be easily extended in the general case as follows. We have

(2.5) \[ \text{Cov}(X^{n+1}, g(X)) = \sum_{\nu=0}^{n} \frac{(-1)^\nu}{\nu!} \binom{n+1}{\nu+1} E \left[ g^{(\nu+1)}(X) X^{n-\nu} a_\nu(X) \right] \]

and

(2.6) \[ X^{n+1} - \mu_{n+1} = -\sum_{\nu=0}^{n} \frac{1}{\nu!} \binom{n+1}{\nu+1} \frac{1}{f(x)} \frac{d^{\nu+1}}{dx^{\nu+1}} [x^{n-\nu} a_\nu(x)] \]

provided that

\[ \lim_{x \to \pm \infty} \frac{d^j}{dx^j} (x^{n-\nu} a_\nu(x)) = 0, \quad j = 0, \ldots, \nu, \quad \nu = 0, 1, \ldots, n. \]

**Theorem 2.3.** Under the assumptions of Theorem 2.1, the function \( h_j(x) \) defined by

(2.7) \[ h_j(x) = \frac{(-1)^j}{f} \frac{d^j}{dx^j} [(w(x))^j f(x)], \quad j = 0, 1, \ldots, n + 1 \]

is a polynomial of degree \( j \) and

\[ E[h_i(X)h_j(X)] = 0 \quad \text{for} \quad i \neq j, \quad i, j = 1, \ldots, n + 1. \]

**Proof.** We consider a polynomial \( Q_s(x) \) of degree \( s \). Since \( w(x) \) is quadratic, for \( j \leq n \), we observe that

\[ \frac{d^j}{dx^j} [Q_s(x)w^j(x)f(x)] = \frac{d^{j-1}}{dx^{j-1}} \left[ (Q_s(x)w(x))' w^{j-2} (x)(w(x)f(x)) + (Q_s(x)w(x))(j - 2)w'(x)w^{j-3}(x)(w(x)f(x)) + \frac{(Q_s(x)w(x))w^{j-2}(x)(\mu - x)f(x)}{\sigma^2} \right]. \]

Hence

\[ \frac{d^j}{dx^j} [Q_s(x)w^j(x)f(x)] = \frac{d^{j-1}}{dx^{j-1}} [P_{s+1}(x)w^{j-1}(x)f(x)] \]

where

\[ P_{s+1}(x) = (Q_s(x)w(x))' + Q_s(x)w'(x)(j - 2) + \frac{(\mu - x)Q_s(x)}{\sigma^2} \]

is a polynomial of degree \( s + 1 \).
If \( Q_s(x) = Q_{s-2}(x)w(x) \), then \( P_{s+1}(x) = P_{s-1}(x)w(x) \) (where \( Q_{s-2}(x) \) and \( P_{s-1}(x) \) are polynomials of degrees \( s-2 \) and \( s-1 \) respectively).

Moreover if \( P_n(x) \) is a polynomial of degree \( n \)

\[
\frac{1}{f} \frac{d}{dx} [P_n(x)w(x)f(x)] = P'_n(x)w(x) + \frac{P_n(x)(\mu - x)}{\sigma^2}
\]

is a polynomial of degree \( n + 1 \). Hence proceeding by induction we conclude that

\[
(2.8) \quad \frac{1}{f} \frac{d^j}{dx^j} [Q_s(x)w^j(x)f(x)] = P_{s+j}(x).
\]

For \( Q_s(x) = 1 \), (2.8) gives that \( h_j(x) \) is a polynomial of \( j \) degree. To show the orthogonality suppose without loss of generality \( i < j \). Then integrating by parts we have

\[
E[h_i(X)h_j(X)] = \int_r^s h_i(x) \frac{d^j}{dx^j} [w^j(x)f(x)] dx
\]

\[
= \lim_{x \to r} \sum_{k=1}^j (-1)^{k-1} \frac{d^{k-1}}{dx^{k-1}} h_i(x) \frac{d^{j-k}}{dx^{j-k}} [w^j(x)f(x)]
\]

\[
- \lim_{x \to r} \sum_{k=1}^j (-1)^{k-1} \frac{d^{k-1}}{dx^{k-1}} h_i(x) \frac{d^{j-k}}{dx^{j-k}} [w^j(x)f(x)]
\]

\[
+ (-1)^j \int_r^s \frac{d^j}{dx^j} [h_i(x)] w^j(x)f(x) dx.
\]

Hence \( E[h_i(X)h_j(X)] = 0 \) provided that relation (2.1) holds for \( j \leq 2n-1 \). □

If all the moments \( \mu_r, r = 1, 2, \ldots, \) of \( X \) exist then Theorem 2.3 yields the Rodrigues formula for the classical orthogonal polynomials by using only the properties of the \( w \)-function which appears in the variance bounds.

3. Characterization of the Pearson system

Relation (1.3) and Johnson’s result (1.8) gives the following

**Theorem 3.1.** Under the assumptions of Theorem 2.1 we have

\[
(3.1) \quad (-1)^{\nu+1} E[X - t]^\nu E[(X - t)^\nu \mid X > t]
\]

\[
+ (-1)^{\nu} E[X - t]^\nu E[(X - t)^{\nu+1} \mid X > t] = c_\nu (w(t))^{\nu+1} h(t)
\]

where \( c_\nu = \nu! b \sigma^{2(\nu+1)} \) and \( h(t) \) is the failure rate \( f(t)/(1 - F(t)) \).

Recently several characterizations for a Pearson random variable have been obtained through some conditional moments. Specifically Glänzel (1991) proved
that under certain conditions the distribution of the random variable $X$ belongs to the Pearson system iff

$$E(X^2 \mid X \geq x) = P(x)E(X \mid X \geq x) + Q(x)$$

where $P(x)$ and $Q(x)$ are polynomials of degree one at most with real coefficients.

Unnikrishnan and Sankaran (1991) proved that $X$ belongs to the Pearson family iff

$$E(X \mid X > x) = E(X) + (a_0 + a_1x + a_2x^2)h(x)$$

where $h(x)$ is the failure rate and $a_0, a_1, a_2$ real constants.

Now we prove the following

**Theorem 3.2.** Let $X$ be a continuous random variable with density $f$, distribution function $F$ and support an interval $(r, s)$. If there exist a function $w$ such that relation (3.1) holds for $\nu = 0$ (only) and $-\frac{f'(t)}{f(t)} = \frac{P_1(t)}{w(t)}$ ($P_1(x)$ is a linear function), then $X$ belongs to a Pearson system of distributions, provided that condition (2.1) holds for $j = 0$.

**Proof.** From the assumption of Theorem 3.2 we conclude that $a_0(t) = c_0w(t)f(t)$ where

$$a_0(t) = E[X-t]F(t) - \int_{t}^{s} (x-t)f(x)dx = \int_{t}^{s} (\mu-x)f(x)dx.$$  

Hence $w(t) = \frac{1}{c_0f(t)} \int_{t}^{s} (\mu-x)f(x)dx$ and $w'(t) = -\frac{f'(t)}{f(t)}w(t) + \frac{\mu-t}{c_0}$. Since $-\frac{f'(t)}{f(t)} = \frac{P_1(t)}{w(t)}$ we conclude that $w'(t) = P_1(t) + (\mu-t)/c_0$ which gives $w(t) = P_2(t)$ ($P_2$ is quadratic). Hence $-\frac{f'}{f} = \frac{P_1(t)}{P_2(t)}$ which completes the proof (or by Korwar’s result). □

**References**


