Ann. Inst. Statist. Math.
Vol. 47, No. 1, 171–176 (1995)

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 properties 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 polynomials, 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 $E(X) = \mu$, $\mathrm{Var}(X) = \sigma^2$ and g absolutely continuous with continuous derivative $g^{(2m+3)}$ of order $2m+3$ $(m \geq 0)$. Specifically,

$$S(2m+1) \leq \mathrm{Var}[g(X)] \leq S(2m), \tag{1.1}$$

where $S(n)$ is given by

$$S(n) = \sum_{\nu=0}^{n} \frac{(-1)^\nu}{\nu!(\nu+1)!} \int_{-\infty}^{+\infty} [g^{(\nu+1)}(t)]^2 a_\nu(t) dt \tag{1.2}$$

and

$$\begin{aligned} a_\nu(t) = (-1)^\nu E[X-t]^{\nu+1} \int_{-\infty}^{t} (x-t)^\nu f(x) dx \\ + (-1)^{\nu+1} E[X-t]^\nu \int_{-\infty}^{t} (x-t)^{\nu+1} f(x) dx. \end{aligned} \tag{1.3}$$

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,

$$\mathrm{Var}[g(X)] \leq \int_{-\infty}^{+\infty} [g'(t)]^2 a_0(t)dt \tag{1.4}$$

where

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

and we set

$$\sigma^2 w(t)f(t) = \int_{-\infty}^{t} (\mu - x)f(x)dx \tag{1.5}$$

if the support of f is an interval.

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

$$\mathrm{Cov}(X, g(X)) = \sigma^2 E[w(X)g'(X)], \tag{1.6}$$

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

$$\frac{d\ln f}{dx} = \frac{\alpha - x}{\beta_0 + \beta_1 x + \beta_2 x^2} \tag{1.7}$$

for some values of α, β_0, β_1, β_2.

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

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

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

$$a_\nu(t) = \nu! b\sigma^{2(\nu+1)}(w(t))^{\nu+1}f(t). \tag{1.8}$$

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

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

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*

$$\lim_{x\to r} x^j w(x) f(x) = \lim_{x\to s} x^j w(x) f(x) = 0, \qquad j = 0,1,2,\dots,n, \tag{2.1}$$

then

$$\operatorname{Cov}(X^{n+1}, g(X)) = \sum_{\nu=0}^{n} (-1)^\nu c_{\nu,n} E[g^{(\nu+1)}(X) X^{n-\nu} (w(X))^{\nu+1}] \tag{2.2}$$

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

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

PROOF. Let $n = 2m$. We apply the second inequality in (1.1) for the function $x^{n+1} + \theta g(x)$ where θ 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

$$\begin{aligned} &\theta^2 \operatorname{Var}[g(X)] + 2\theta \operatorname{Cov}[X^{n+1}, g(X)] \\ &\le \theta^2 \sum_{\nu=0}^{n} \frac{(-1)^\nu}{(\nu+1)!} b\sigma^{2(\nu+1)} E[(g^{(\nu+1)}(X))^2 (w(X))^{\nu+1}] \\ &+ 2\theta \sum_{\nu=0}^{n} \frac{(-1)^\nu}{(\nu+1)!} b\sigma^{2(\nu+1)} (n+1)_{\nu+1} \\ &\cdot E[(w(X))^{\nu+1} X^{n-\nu} g^{(\nu+1)}(X)]. \end{aligned} \tag{2.3}$$

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

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

$$x^{n+1} - \mu_{n+1} = -\sum_{\nu=0}^{n} c_{\nu,n} \frac{1}{f} \frac{d^{\nu+1}}{dx^{\nu+1}} [x^{n-\nu} (w(x))^{\nu+1} f(x)] \tag{2.4}$$

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

PROOF. In (2.2) take first $g(x) = \operatorname{Re}[e^{itx}]$, where $i = \sqrt{-1}$, and then $g(X) = \operatorname{Im}[e^{itx}]$. Integrating the RHS by parts $(\nu+1)$-times $\nu = 0,1,\dots,n$, 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. □

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

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

and

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

provided that

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

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

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

is a polynomial of degree j *and*

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

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

$$\begin{aligned}\frac{d^j}{dx^j}[Q_s(x) w^j(x) f(x)] = \frac{d^{j-1}}{dx^{j-1}} \Big[& (Q_s(x) w(x))' w^{j-2}(x)(w(x) f(x)) \\ & + (Q_s(x) w(x))(j-2) w'(x) w^{j-3}(x)(w(x) f(x)) \\ & + \frac{(Q_s(x) w(x)) w^{j-2}(x)(\mu - x) f(x)}{\sigma^2} \Big].\end{aligned}$$

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

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

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

$$\begin{aligned} E[h_i(X)h_j(X)] &= \int_r^s h_i(x)\frac{d^j}{dx^j}[w^j(x)f(x)]dx \\ &= \lim_{x\to s}\sum_{k=1}^{j}(-1)^{k-1}\frac{d^{k-1}}{dx^{k-1}}h_i(x)\frac{d^{j-k}}{dx^{j-k}}[w^j(x)f(x)] \\ &\quad - \lim_{x\to r}\sum_{k=1}^{j}(-1)^{k-1}\frac{d^{k-1}}{dx^{k-1}}h_i(x)\frac{d^{j-k}}{dx^{j-k}}[w^j(x)f(x)] \\ &\quad + (-1)^j\int_r^s \frac{d^j}{dx^j}[h_i(x)]w^j(x)f(x)dx. \end{aligned}$$

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

If all the moments μ_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*

$$\begin{aligned} &(-1)^{\nu+1}E[X-t]^{\nu+1}E[(X-t)^{\nu} \mid X > t] \\ &\qquad + (-1)^{\nu}E[X-t]^{\nu}E[(X-t)^{\nu+1} \mid X > t] = c_\nu(w(t))^{\nu+1}h(t) \end{aligned} \tag{3.1}$$

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

$$E(X \mid X > x) = E(X) + (a_0 + a_1 x + a_2 x^2)h(x)$$

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 $-\frac{f'}{f} = \frac{P_1}{w}$ ($P_1(x)$ 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 $a_0(t) = c_0 w(t) f(t)$ where

$$a_0(t) = E[X - t]F(t) - \int_r^t (x - t) f(x) dx = \int_r^t (\mu - x) f(x) dx.$$

Hence $w(t) = \frac{1}{c_0 f(t)} \int_r^t (\mu - x) f(x) dx$ 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). $\square$

REFERENCES

Cacoullos, T. and Papathanasiou, V. (1985). On upper bounds for the variance of functions of random variables, *Statist. Probab. Lett.*, **3**, 175–184.

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

Glänzel, W. (1991). Characterization through some conditional moments of Pearson-type distributions and discrete analogues, *Sankhyā Ser. B*, **53**, 17–24.

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

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

Papathanasiou, V. (1989). Variance bounds by a generalization of the Cauchy-Schwarz inequality, *Statist. Probab. Lett.*, **7**, 29–33.

Unnikrishnan, N. and Sankaran, P. (1991). Characterization of the Pearson family of distributions, *IEEE Transactions on Reliability*, **40**, 75–77.