Piotr Pawlas and Dominik Szynal

ON A CHARACTERIZATION OF A POWER DISTRIBUTION

Communicated by J. Wesołowski

Abstract. A measure of dependence called pseudo-covariance and related to covariance was proposed by Pawlas and Szynal [4]. It was used, among other things, in a characterization of a power distribution. Here we generalized that result.

1. Introduction

Krajka and Szynal [3] introduced the concept of Q-covariance for random variables $X$ and $Y$ to investigate their dependence when the classic and $F$-covariance fail (cf. Drouet-Mari and Kotz [2]). Next, a new measure of dependence called pseudo-covariance and related to covariance was proposed by Pawlas and Szynal [4]. It may be applied as a measure of dependence of uncorrelated random variables and used in characterizations of continuous distributions.

**Definition 1.** Let $(X, Y)$ be a pair of random variables with continuous distribution functions. By pseudo-covariance of $(X, Y)$ ($\text{Cov}(PD)(X, Y)$), we mean the quantity

$$\text{Cov}(PD)(X, Y) = \left( \int_0^1 (y(p))^2 dp \right)^{1/2} \left( \int_0^1 (E(X|Y = y(p)) - EX)^2 dp \right)^{1/2}$$

whenever the RHS is finite, where $y(p)$ is the quantile function of $Y$, $p \in (0, 1)$.

1.1. The characterization of a power distribution based on order statistics. Let $X_1, \ldots, X_n$ be i.i.d. random variables from power distribution on $(0, 1)$ and let $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ denote order statistics. We
show for simplicity of evaluation that for the power distribution on $(0, 1)$ with the cumulative distribution function

\[ F(x) = x^{1/m}, \]

for \( x \in (0, 1), \ m \in \mathbb{N}, \)

\begin{equation}
\text{Cov}(X_{i:n}, X_{j:n}) = \text{Cov}(X_{i:n}^o, X_{j:n}^o) = \text{Cov}^{(PD)}(X_{i:n}^o, X_{j:n}^o),
\end{equation}

where \( X_{i:n}^o = X_{i:n} - EX_{i:n} \).

It is known (cf. [1]) that for the power distribution with cumulative distribution function

\[ F(x) = x^{1/m}, \ \text{for} \ x \in (0, 1), \ m \in \mathbb{N}, \]

\[ \text{Cov}(X_{i:n}, X_{j:n}) = \frac{\Gamma(n + 1)\Gamma(m + i)\Gamma(2m + j)}{\Gamma(i)\Gamma(m + j)\Gamma(2m + n + 1)} - \frac{(\Gamma(n + 1))^2\Gamma(m + i)\Gamma(m + j)}{(\Gamma(m + n + 1))^2\Gamma(i)\Gamma(j)}. \]

Now, we note that by the definition of pseudo-covariance (1)

\begin{equation}
\text{Cov}^{(PD)}(X_{i:n}^o, X_{j:n}^o)
\end{equation}

\[ = \left( \int_0^1 (y^o(p))^2 dp \right)^{1/2} \left( \int_0^1 (E(X_{i:n}^o|X_{j:n}^o = y^o(p)) - EX_{i:n}^o)^2 dp \right)^{1/2}, \]

where \( y^o(p) = y(p) - EX_{j:n} \) and \( y(p) \) is the quantile of order \( p \) of \( X_{j:n} \).

Taking into account the formula

\[ \sigma^2 X = \int_0^1 \left[ F_X^{-1}(x) - EX \right]^2 dx \ (\text{cf. [5]}), \]

we get

\begin{equation}
\int_0^1 (y^o(p))^2 dp = \text{Var}(X_{j:n}) = \frac{\Gamma(n + 1)\Gamma(2m + j)}{\Gamma(j)\Gamma(2m + n + 1)} - \left( \frac{\Gamma(n + 1)\Gamma(m + j)}{\Gamma(j)\Gamma(m + n + 1)} \right)^2.
\end{equation}

The conditional density function of \( X_{i:n} \), given that \( X_{j:n} = y \), is given by

\[ f_{X_{i:n}}(x|X_{j:n} = y) = \frac{(j - 1)!}{(i - 1)!(j - i - 1)!} \left( \frac{F(x)}{F(y)} \right)^{i-1} \times \left( \frac{F(y) - F(x)}{F(y)} \right)^{j-i-1} \times \frac{f(x)}{F(y)}. \]

Hence, for \( i < j \)

\[ E(X_{i:n}|X_{j:n} = y) = \left( \frac{i}{m + i} \right)^{(j-i)} y, \]
where
\[
\left( \frac{i}{m+i} \right)^{(j-i)} = \frac{i(i+1) \cdots (j-1)}{(m+i)(m+i+1) \cdots (m+j-1)}.
\]
Moreover,
\[
E(X_{i:n}) = \left( \frac{i}{m+i} \right)^{(j-i)} E(X_{j:n}).
\]
Therefore,
\[
\int_0^1 \left( E(X_{i:n}|X_{j:n} = y(p)) - E(X_{i:n}) \right)^2 dp
\]
\[
= \int_0^1 \left( \left( \frac{i}{m+1} \right)^{(j-i)} y(p) - \left( \frac{i}{m+1} \right)^{(j-i)} E(X_{j:n}) \right)^2 dp
\]
\[
= \left( \left( \frac{i}{m+1} \right)^{(j-i)} \right)^2 \text{Var}(X_{j:n})
\]
\[
= \left( \left( \frac{i}{m+i} \right)^{(j-i)} \right)^2 \left[ \frac{\Gamma(n+1)\Gamma(2m+j)}{\Gamma(j)\Gamma(2m+n+1)} - \left( \frac{\Gamma(n+1)\Gamma(m+j)}{\Gamma(j)\Gamma(m+n+1)} \right)^2 \right].
\]
Thus,
\[
\text{Cov}^{PD}(X_{i:n}^o, X_{j:n}^o) = \left( \frac{i}{m+i} \right)^{(j-i)} \text{Var}(X_{j:n})
\]
\[
= \left( \frac{i}{m+i} \right)^{(j-i)} \left[ \frac{\Gamma(n+1)\Gamma(2m+j)}{\Gamma(j)\Gamma(2m+n+1)} - \left( \frac{\Gamma(n+1)\Gamma(m+j)}{\Gamma(j)\Gamma(m+n+1)} \right)^2 \right]
\]
and
\[
\left( \frac{i}{m+i} \right)^{(j-i)} \left[ \frac{\Gamma(n+1)\Gamma(2m+j)}{\Gamma(j)\Gamma(2m+n+1)} - \left( \frac{\Gamma(n+1)\Gamma(m+j)}{\Gamma(j)\Gamma(m+n+1)} \right)^2 \right]
\]
\[
= \frac{\Gamma(j)\Gamma(m+i)}{\Gamma(i)\Gamma(m+j)} \left[ \frac{\Gamma(n+1)\Gamma(2m+j)}{\Gamma(j)\Gamma(2m+n+1)} - \left( \frac{\Gamma(n+1)\Gamma(m+j)}{\Gamma(j)\Gamma(m+n+1)} \right)^2 \right]
\]
\[
= \frac{\Gamma(n+1)\Gamma(m+i)\Gamma(2m+j)}{\Gamma(i)\Gamma(m+j)\Gamma(2m+n+1)} \left( \frac{(\Gamma(n+1))^2\Gamma(m+i)\Gamma(m+j)}{(\Gamma(m+n+1))^2\Gamma(i)\Gamma(j)} \right)
\]
\[
= \text{Cov}(X_{i:n}, X_{j:n}).
\]
The same statement holds true when we put in (2) \( a > 0 \) instead of \( 1/m \). Now, we are able to prove a stronger result.
**Theorem.** Let \((X_1, \ldots, X_n)\) be a sample from an absolutely continuous distribution function \(F\) on \((0, 1)\) and \(X_{1:n}, X_{2:n}, \ldots, X_{n:n}\) stand for order statistics. Then

\[
\text{Cov}(X_{i:n}, X_{j:n}) = \text{Cov}^{(Q)}(X_{i:n}, X_{j:n}) = \text{Cov}^{(PD)}(X_{i:n}^o, X_{j:n}^o)
\]

if and only if \(X\) has a power distribution.

**Proof.** Assume that \(X\) has a power distribution on \((0, 1)\). Then it was shown that (6) holds true. Now, suppose that the equality (6) is satisfied for \(X\) having an absolutely continuous distribution on \((0, 1)\) with a density function \(f\). Consider first the case when \(j = i + 1\). We know that

\[
|\text{Cov}^{(Q)}(X_{i:n}, X_{i+1:n})| = |\text{Cov}^{(Q)}(X_{i:n}^o, X_{i+1:n}^o)| \leq \text{Cov}^{(PD)}(X_{i:n}^o, X_{i+1:n}^o)
\]

\[
= \left( \int_0^1 (y^o(p))^2 dp \right)^{\frac{1}{2}} \left( \int_0^1 (E(X_{i:n}|X_{i+1:n} = y(p)) - EX_{i:n})^2 dp \right)^{\frac{1}{2}} (\text{cf. [4]}),
\]

where \(y^o(p) = y(p) - EX_{i+1:n}\).

But equality in the above Schwarz inequality holds iff

\[
cy^o(p) = E(X_{i:n}|X_{i+1:n} = y(p)) - EX_{i:n}
\]

or

\[
c(y(p) - EX_{i+1:n}) = E(X_{i:n}|X_{i+1:n} = y(p)) - EX_{i:n}
\]

for some constant \(c \in \mathbb{R}\). Taking into account that

\[
f_{X_{i:n}}(x|X_{i+1:n} = y) = \frac{(i + 1 - 1)!}{(i - 1)!(i + 1 - i - 1)!} \left( \frac{F(x)}{F(y)} \right)^{i-1} \frac{f(x)}{F(y)}
\]

\[
= i \left( \frac{F(x)}{F(y)} \right)^{i-1} \frac{f(x)}{F(y)},
\]

we get

\[
E\left( X_{i:n}|X_{i+1:n} = y(p) \right) = \frac{i}{F^i(y(p))} \int_0^{y(p)} x F^{i-1}(x) f(x) dx
\]

\[
= \frac{1}{F^i(y(p))} \int_0^{y(p)} x (F^i(x))' dx
\]

\[
= \frac{1}{F^i(y(p))} \left[ y(p) F^i(y(p)) - \int_0^{y(p)} F^i(x) dx \right]
\]

\[
= y(p) - \frac{1}{F^i(y(p))} \int_0^{y(p)} F^i(x) dx.
\]

Therefore, \(7\) takes the form

\[
c(y(p) - EX_{i+1:n}) = y(p) - \frac{1}{F^i(y(p))} \int_0^{y(p)} F^i(x) dx - EX_{i:n}
\]
or
\[ y(p)(1 - c) + cEX_{i+1:n} - EX_{i:n} = \frac{1}{F^i(y(p))} \int_0^{y(p)} F^i(x)dx. \]

Differentiating the above equality with respect to \( p \), we obtain
\[ y'(p)(1 - c) = \frac{F^i(y(p))y'(p)F^i(y(p)) - iF^{i-1}(y(p))f(y(p))y'(p)\int_0^{y(p)} F^i(x)dx}{F^{2i}(y(p))} \]
or
\[ y'(p)(1 - c) = y'(p) - \frac{if(y(p))y'(p)\int_0^{y(p)} F^i(x)dx}{F^{i+1}(y(p))}. \]

Hence, we get
\[ \frac{c}{\int} - \frac{f(y(p))\int_0^{y(p)} F^i(x)dx}{F^{i+1}(y(p))} = 0, \]
which leads us to
\[ c \int \frac{F^i(y(p))y'(p)}{\int_0^{y(p)} F^i(x)dx}dp = i \int \frac{f(y(p))y'(p)}{F(y(p))}dp. \]

Hence, we have
\[ \ln \int_0^{y(p)} F^i(x)dx = \frac{i}{c} \ln F(y(p)) + \ln A \]
or
\[ \int_0^{y(p)} F^i(x)dx = A(F(y(p)))^{\frac{i}{c}}. \]

Differentiating the above equality with respect to \( p \), we obtain
\[ F^i(y(p))y'(p) = \frac{Ai}{c} (F(y(p)))^{\frac{i}{c} - 1} f(y(p))y'(p). \]

Thus
\[ F^{i-\frac{i}{c} + 1}(y(p)) = \frac{Ai}{c} f(y(p)), \]
for all \( p \in (0, 1) \).

Letting in (9) \( y(p) := y, p \in (0, 1) \), we have
\[ F^{i-\frac{i}{c} + 1}(y) = \frac{Ai}{c} f(y) = \frac{Ai}{c} dF(y)/dy. \]

Solving this differential equation, we get
\[ y = \frac{A}{(1 - c)} F^{\frac{i}{c} - i}(y) + B. \]
But by the assumptions $F(0) = 0$ and $F(1) = 1$, we see that $B = 0$ and $A = 1 - c$. Hence

$$F(y) = y^{\frac{c}{1-\epsilon}}, \quad \text{for } y \in (0, 1) \text{ and } c \in (0, 1).$$

The above arguing for $j = i + 1$, we can repeat for $j > i + 1$. Then we conclude that the equality in

$$|\text{Cov}^{(Q)}(X_{i:n}, X_{j:n})| \leq |\text{Cov}^{(PD)}(X_{i:n}, X_{j:n})|$$

holds true iff

$$c(y(p) - E X_{j:n}) = E(X_{i:n} \mid X_{j:n} = y(p)) - E X_{i:n},$$

for some constant $c \in \mathbb{R}$ and the equality (8) takes in this case the form

$$\sum_{k=0}^{j-i-1} \binom{j-i-1}{k} (-1)^k \left( \frac{c}{i+k} - \frac{f(y(p))}{F^i+k(y(p))} \frac{F^{i+k}(x)dx}{F^{i+k+1}(y(p))} \right) = 0.$$ (11)

One can see that the above equality holds true for $F(x) = x^a$, for $x \in (0, 1)$, $a > 0$, with

$$c = \prod_{h=0}^{j-1} \frac{i + h}{\frac{1}{a} + i + h},$$

which completes the proof. The characterization of a power distribution in Theorem generalizes the similar result in [4] (where $i=1$).

**Acknowledgments.** We thank a referee for suggestions improving our article.

**References**


Received February 5, 2014; revised version June 18, 2014.