Far East Journal of Theoretical Statistics
Published Online: April 2016
http://dx.doi.org/10.17654/TS052020139
Volume 52, Number 2, 2016, Pages 139-148 **ISSN: 0972-0863**

CHARACTERIZATION OF DISTRIBUTIONS BASED ON *k*-TH LOWER RECORD VALUES

Piotr Pawlas

Institute of Mathematics
Maria Curie-Skłodowska University
pl. M. Curie-Skłodowskiej 1
PL-20-031 Lublin
Poland
e-mail: piotr.pawlas@poczta.umcs.lublin.pl

Abstract

We give new characterizations of negative exponential, negative Pareto and inverse power distribution in terms of *k*-th lower record values. The characterizations are based on properties of a measure of dependence called the pseudo-covariance.

1. Introduction

The concept of dependence for random variables X and Y defined by

$$Cov(X, Y) = EXY - EY\ EY,$$

when $X \in L^1$, $Y \in L^1$, $XY \in L^1$ (L^1 - the space of integrable variables) has been modified by many authors. For instance, Hoeffding [4] gave the

Received: December 11, 2015; Revised: February 3, 2016; Accepted: February 9, 2016
2010 Mathematics Subject Classification: 60E15, 62E10, 62G30.
Keywords and phrases: *k*-th upper record values, *k*-th lower record values, covariance, characterization, linearity of regression, exponential distribution, power distribution, Pareto distribution.
Communicated by K. K. Azad

formula

$$Cov^{(F)}(X, Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[F_{(X,Y)}(x, y) - F_X(x)F_Y(y)]dxdy.$$

Mardia and Thomson [6] generalized this formula introducing the quantity

$$Cov^{(F)}(X^r, Y^s) = rs\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x^{r-1}y^{s-1}[F_{(X,Y)}(x, y) - F_X(x)F_Y(y)]dxdy,$$

for $r, s \geq 1,$ provided that this integral is finite.

Krajka and Szynal [5] introduced the idea 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 [3]). Next a new measure of dependence called pseudo-covariance related to covariance was proposed by Pawlas and Szynal [9]. It may be applied as a measure of dependence of uncorrelated random variables (cf. [9]) and used in characterizations of continuous distributions (cf. [10]). Now we use this measure in a characterization of negative exponential, negative Pareto and inverse power distribution in terms of k-th lower record values. First we recall the concept of k-th lower and upper record values of $\{X_n, n \geq 1\}$.

Let $\{X_n, n \geq 1\}$ be a sequence of i.i.d random variables with common distribution function F and density function f of X. Let

$$X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$$

denote the order statistics of a sample $X_1, ..., X_n$. For a fixed integer $k \geq 1$ we define the sequences $\{U_k(n), n \geq 1\}$ and $\{L_k(n), n \geq 1\}$ of k-th upper and lower record times of $\{X_n, n \geq 1\}$ as follows:

$$U_k(1) = 1, \quad U_k(n+1) = \min\{j > U_k(n), X_{j:j+k-1} > X_{U_k(n):U_k(n)+k-1}\}$$

and

$$L_k(1) = 1, \quad L_k(n+1) = \min\{j > L_k(n), X_{k:L_k(n)+k-1} > X_{k:j+k-1}\}.$$

Then the sequences $\{Y_n^{(k)}, n \geq 1\}$ and $\{Z_n^{(k)}, n \geq 1\}$ with

$$Y_n^{(k)} = X_{U_k(n):U_k(n)+k-1} \text{ and } Z_n^{(k)} = X_{k:L_k(n)+k-1},$$

$n = 1, 2, \ldots,$ are called the *sequences* of k-th upper and lower record values of $\{X_n, n \geq 1\}$.

For more details on the k-th lower record values see [8] and [1]. Some characterizations of probability distributions via upper record values were given in Newzorov [7], Dembińska and Wesołowski [2], Pawlas and Szynal [9].

Now we present the notion of pseudo-covariance for random variables X and Y.

Let $(\Omega, \mathcal{A}, P)$ be a probability space and L^r denote the space of random variables X such that $E|X|^r < \infty$ for $r > 0$. L^0 stands for the space of all random variables having continuous distributions. For any $p \in (0, 1)$, $y(p)$ is the quantile function of the random variable Y i.e. $P[Y < y(p)] \leq p \leq P[Y \leq y(p)]$.

Now for $X \in L^1$ and $Y \in L^0$ with continuous distribution functions we write

$$L_{X,Y}(p) := E(X - EX)I[Y \geq y(p)],$$

and

$$\overline{L}_{X,Y}(p) := E(X - EX)I[Y < y(p)],$$

where $I[\cdot]$ is the indicator function. We see that

$$L_{X-EX,Y-EY}(p) = L_{X,Y}(p),$$

$$\overline{L}_{X-EX,Y-EY}(p) = \overline{L}_{X,Y}(p),$$

and

$$\overline{L}'_{X,Y}(p) = E(X \mid Y = y(p)) - EX,$$

where $\overline{L}'_{X,Y}(p) = \frac{d}{dp}\overline{L}_{X,Y}(p)$.

Definition 1 (cf. Krajka and Szynal [5])**.** For a random variable $X \in L^1$ and a random variable $Y \in L^0$, *Q-covariance* $Cov^{(Q)}(X, Y)$ is defined as follows:

$$Cov^{(Q)}(X, Y) = -\int_0^1 y(p)dL_{X,Y}(p) = \int_0^1 y(p)d\overline{L}_{X,Y}(p)$$

provided that one of the above integrals is finite.

The relation between $Cov(X, Y) = E(X - EX)(Y - EY)$ and $Cov^{(Q)}(X, Y)$ is given in

Theorem 1 (cf. Krajka and Szynal [5])**.** *Let X and Y be two random variables with continuous distribution functions such that* $X \in L^1$, $Y \in L^1$ *and* $XY \in L^1$. *Then*

$$Cov(X, Y) = Cov^{(Q)}(X, Y).$$

The new measure of dependence called *pseudo-covariance* (cf. Pawlas and Szynal [9]) was introduced using the following bound for $Cov^{(Q)}(X, Y)$.

Theorem 2 (cf. Krajka and Szynal [5])**.** *Let* (X, Y) *be a pair of random variables with continuous and strictly monotone marginal distribution functions. Suppose that* $X \in L^1$, $Y \in L^r$, $\overline{L}_{X,Y}(p)$ *is differentiable for* $p \in (0, 1)$ *and* $\overline{L}'_{X,Y}(p) \in L^s$, *where* $r, s > 0$ *and* $\frac{1}{r} + \frac{1}{s} = 1$. *Then*

$$|\, Cov^{(Q)}(X, Y)\,| \le \left(\int_0^1 |\, y(p)\,|^r\, dp\right)^{\frac{1}{r}} \left(\int_0^1 |\, \overline{L}'_{X,Y}(p)\,|^s\, dp\right)^{\frac{1}{s}}. \qquad (1)$$

Corollary 1. *Suppose that* $r = s = 2$ *in Theorem* 2. *Then*

$$| Cov^{(Q)}(X, Y) | \le \left(\int_0^1 | y(p) |^2 dp \right)^{\frac{1}{2}} \left(\int_0^1 (E(X | Y = y(p)) - EX)^2 dp \right)^{\frac{1}{2}}. \quad (2)$$

Note that the classical bound for $Cov(X, Y)$ is in this case

$$| Cov(X, Y) | \le \sigma X \sigma Y, \quad (3)$$

where $\sigma^2 X = \text{Var } X$, $\sigma^2 Y = \text{Var } Y$.

We see that the bound in (2),

$$\left(\int_0^1 | y(p) |^2 dp \right)^{\frac{1}{2}} \left(\int_0^1 (E(X | Y = y(p)) - EX)^2 dp \right)^{\frac{1}{2}} \quad (4)$$

is 0 for independent random variables X and Y while the bound $\sigma X \sigma Y$ in (3) is positive. This leads us to the new measure of dependence for (X, Y).

Definition 2 (cf. Pawlas and Szynal [9])**.** Let (X, Y) be a pair of random variables with continuous distribution functions. By *pseudo-covariance of* $(X, Y)(Cov^{(PD)}(X, Y))$ we mean the quantity

$$Cov^{(PD)}(X, Y) = \left(\int_0^1 | y(p) |^2 dp \right)^{\frac{1}{2}} \left(\int_0^1 (E(X | Y = y(p)) - EX)^2 dp \right)^{\frac{1}{2}} \quad (5)$$

whenever RHS is finite.

The following example shows some properties of the *PD*-covariance.

Example 1 (cf. Pawlas and Szynal [9])**.** Let the cumulative density function $f_{X,Y}(x, y)$ of a random vector (X, Y) has the form:

$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{4}[1 + xy(x^2 - y^2)]; & | x | < 1 \wedge | y | < 1, \\ 0; & | x | \ge 1 \vee | y | \ge 1. \end{cases}$$

Then the marginal distributions of X and Y are uniform on $(-1, 1)$. Moreover, we see that $f_{X,Y}(x, y) \neq f_X(x) \cdot f_Y(y)$, so X and Y are dependent. Since X and Y are uncorrelated we have $Cov(X, Y) = 0$. But $Cov^{(PD)}(X, Y) \neq 0$.

We see that $Cov^{(PD)}(X, Y)$ provide a measure of dependence for uncorrelated random variables, i.e. when $Cov(X, Y) = 0$.

Moreover, it is known [9] that

$$Cov(X, Y) = Cov^{(Q)}(X, Y) = Cov^{(PD)}(X, Y) = 0$$

whenever X and Y are independent and satisfy some moment conditions.

Now we show that the equality

$$Cov^{Q}(X, Y) = Cov^{(PD)}(X, Y)$$

is true also for some dependent random variables. This equality can be used in characterizations of probability distributions.

2. The Characterization of Distributions based on *k*-th Lower Record Values

We need the following probability distribution functions:

- The negative exponential distribution with

$$F(x) = e^{\lambda(x-\nu)}, \quad x < \nu;\ \lambda > 0,\ \nu \in R. \tag{6}$$

- The inverse power distribution with

$$F(x) = \left(\frac{x-\alpha}{\beta-\alpha}\right)^{\theta}, \quad \alpha < x < \beta;\ \theta > 0,\ \alpha, \beta \in R,\ \alpha < \beta. \tag{7}$$

- The negative Pareto distribution with

$$F(x) = \left(\frac{\delta-\nu}{\delta-x}\right)^{\theta}, \quad x < \nu;\ \theta > 0,\ \nu, \delta \in R,\ \nu < \delta. \tag{8}$$

Theorem 3. *Let $Z_m^{(k)}$ and $Z_n^{(k)}$ be record values from $\{X_n, n \geq 1\}$, $m < n$. Write $Z_m^0 = Z_m^{(k)} - EZ_m^{(k)}$ and $Z_n^0 = Z_n^{(k)} - EZ_n^{(k)}$. Then*

$$Cov^{Q}(Z_n^0,\, Z_m^0) = Cov^{(PD)}(Z_n^0,\, Z_m^0) \tag{9}$$

holds true if and only if X has:

(1) *negative exponential distribution*,

(2) *negative Pareto distribution*,

(3) *inverse power distribution*.

Proof. Note that for the negative exponential, negative Pareto and inverse power distribution, the quality (9) holds true. Indeed we have:

I. Let F be the negative exponential distribution function (6) with parameters $\lambda = 1$ and $\nu = 0$.

Then we have

$$Cov^{(Q)}(Z_n^0,\, Z_m^0) = Cov(Z_n^0,\, Z_m^0) = \frac{m}{k^2},$$

$$\left(\int_0^1 |\, y^0(p)\,|^2\, dp\right)^{\frac{1}{2}} = \frac{\sqrt{m}}{k},$$

and

$$\left(\int_0^1 (E(Z_n^0 \,|\, Z_m^0 = y^0(p)) - EZ_n^0)^2\, dp\right)^{\frac{1}{2}}$$

$$= \left(\int_0^1 (E(Z_n^{(k)} \,|\, Z_m^{(k)} = y(p)) - EZ_n^{(k)})^2\, dp\right)^{\frac{1}{2}} = \frac{\sqrt{m}}{k},$$

which proves (9). The case when $\lambda > 0,\ \nu \in R$ can be proved similarly.

II. Let F be the inverse power distribution function (7).

Then we have

$$Cov^{(Q)}(Z_n^0,\, Z_m^0) = Cov(Z_n^0,\, Z_m^0) = \left(\frac{k\theta}{k\theta + 1}\right)^{n-m} Var(Z_m^0),$$

$$\left(\int_0^1 | y^0(p) |^2 dp\right)^{\frac{1}{2}} = \sqrt{Var(Z_m^0)},$$

and

$$\left(\int_0^1 (E(Z_n^0 | Z_m^0 = y^0(p)) - EZ_n^0)^2 dp\right)^{\frac{1}{2}} = \left(\frac{k\theta}{k\theta + 1}\right)^{n-m} \sqrt{Var(Z_m^0)}.$$

Hence we obtain the equality (9).

III. Let F be the negative Pareto distribution function (8).

Then we have

$$Cov^{(Q)}(Z_n^0, Z_m^0) = Cov(Z_n^0, Z_m^0) = \left(\frac{k\theta}{k\theta - 1}\right)^{n-m} Var(Z_m^0),$$

$$\left(\int_0^1 | y^0(p) |^2 dp\right)^{\frac{1}{2}} = \sqrt{Var(Z_m^0)},$$

and

$$\left(\int_0^1 (E(Z_n^0 | Z_m^0 = y^0(p)) - EZ_n^0)^2 dp\right)^{\frac{1}{2}} = \left(\frac{k\theta}{k\theta - 1}\right)^{n-m} \sqrt{Var(Z_m^0)},$$

which gives (9).

Now we show that the equality (9) implies that F belongs to one of the above mentioned classes of distributions (6)-(8). It is known from the Schwarz inequality that the equation (9) holds true iff there exists a constant c such that

$$c(y(p) - EZ_m^{(k)}) = E(Z_n^{(k)} | Z_m^{(k)} = y(p)) - EZ_n^{(k)}.$$

Hence we get

$$c = \frac{y(p) - [EZ_n^{(k)} + y(p) - E(Z_n^{(k)} | Z_m^{(k)} = y(p))]}{y(p) - EZ_m^{(k)}}. \tag{10}$$

We consider three cases:

(i) For $c = 1,$ we obtain

$$E(Z_n^{(k)} \mid Z_m^{(k)} = y(p)) = y(p) + EZ_n^{(k)} - EZ_m^{(k)}.$$

But we know from Bieniek and Szynal [1] that

$$E(Z_n^{(k)} \mid Z_m^{(k)} = x) = x + b$$

is satisfied only for negative exponential distribution (6).

In our case $b := b_{n,m,k} = -\dfrac{n-m}{k\lambda}.$

Now assume that $c \neq 1.$ Then from (10), we have

$$E(Z_n^{(k)} \mid Z_m^{(k)} = y(p)) = cy(p) + EZ_n^{(k)} - cEZ_m^{(k)}.$$

(ii) If $c < 1,$ then by results of [1] the equality

$$E(Z_n^{(k)} \mid Z_m^{(k)} = x) = cx + b$$

is true only for inverse power distribution (7). In our case $c := c_{n,m,k} = \left(\dfrac{k\theta}{k\theta+1}\right)^{n-m}$ and

$$b := b_{n,m,k} = \alpha\left\{\left(\frac{k\theta}{k\theta+1}\right)^{n-m} - 1\right\}.$$

(iii) If $c > 1,$ then from [1] the equation

$$E(Y_n^{(k)} \mid Y_m^{(k)} = x) = cx + b$$

is true only for negative Pareto distribution (8). In this case $c := c_{n,m,k} = \left(\dfrac{k\theta}{k\theta-1}\right)^{n-m}$ and $b := b_{n,m,k} = \delta\left\{1 - \left(\dfrac{k\theta}{k\theta-1}\right)^{n-m}\right\},$ which ends the proof and completes the proof of Theorem 3.

Acknowledgement

The author thanks the anonymous referees for their valuable suggestions which led to the improvement of the manuscript.

References

[1] M. Bieniek and D. Szynal, Characterizations based on *k*-th upper and lower record values, Demonstratio Math. 37(2) (2004), 263-273.

[2] A. Dembińska and J. Wesołowski, Linearity of regression for non-adjacent record values, J. Stat. Plan. Infer. 90 (2000), 195-205.

[3] D. Drouet-Mari and S. Kotz, Correlation and Dependence, Imperial College Press, London, 2001.

[4] W. Hoeffding, Mastabinvariante Korrelations-Theorie, Schriften Math. Inst. Univ. Berlin 5 (1940), 187-233.

[5] A. Krajka and D. Szynal, On *Q*-covariance and its applications, Proceedings of International Conference on Linear Statistical Inference LINSTAT'93, Kluwer Academic Publishers, Dordrecht, 1994, pp. 293-300.

[6] K. V. Mardia and J. W. Thomson, Unified treatment of moments-formulae, Sankhyā A 33 (1971), 121-132.

[7] V. B. Newzorov, Records: Mathematical Theory, Amer. Math. Soc., 2000.

[8] P. Pawlas and D. Szynal, Relations for single and product moments of *k*-th record values from exponential and Gumbel distributions, J. Appl. Statist. Sci. 7 (1998), 53-63.

[9] P. Pawlas and D. Szynal, On a new measure of dependence and its applications, Demonstratio Math. 45(2) (2012), 243-256.

[10] P. Pawlas and D. Szynal, On a characterization of a power distribution, Demonstratio Math. 48(1) (2015), 100-106.