Applied Mathematics and Computation 226 (2014) 44–50

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

Characterization of general classes of distributions based on independent property of transformed record values

Matej Juhás *, Valéria Skřivánková

Institute of Mathematics, P. J. Šafárik University of Košice, Jesenná 5, 040 01 Košice, Slovakia

ARTICLE INFO

Keywords:
Record values
Exponential distribution
Fréchet distribution
Gumbel distribution
Log-normal distribution
Weibull distribution
Simulation
Hoeffding test

ABSTRACT

In this paper, general classes of probability distributions are characterized using the independence of suitable transformations of records in a sequence of independent, identically distributed random variables. Examples of special cases of general classes as Gumbel, Fréchet, Weibull, exponential and lognormal distributions are discussed. Further we use the theoretical results for application to simulated data.

1. Introduction

The model of record values was introduced by Chandler in 1952 [7], who studied the stochastic behavior of records and defined the basic terms as record value, record time, inter-record time and frequency of records. The problem of characterization of probability distributions by some properties of record values was opened by Ahsanullah in 1982 when he characterized the exponential distribution ([3]). Ahsanullah presented in 2005 in monograph [4] the recent developments of classical records (records of independent identically distributed random variables) distributed according to exponential law, generalized extreme value distribution, generalized Pareto distribution and power law. Korean mathematicians Lee, Lim and Chang dealt with the problem of characterization of Weibull and Pareto distribution by special transformations of records ([8,9,5]). Abu-Youssef [1], Malinowska and Szynal [10], Faizan and Khan [6] characterized wide classes of distributions using conditional expectation of function of record values.

In this paper, general classes of continuous probability distributions are characterized using independent property of transformed record values. Special cases of general classes as Gumbel, Fréchet, Weibull, exponential and lognormal distributions are also discussed. Some modifications of main theorems are formulated as corollaries. The theoretical results are applied for analyzing simulated data. After estimating unknown parameters of chosen distributions we compare the results by Kolmogorov–Smirnov (K–S), Anderson–Darling (A–D) goodness of fit test and Hoeffding test for transformed data.

Consider the sequence $\{X_n, n \geqslant 1\}$ of independent, identically distributed (iid) random variables with common absolutely continuous distribution function F and probability density function (p.d.f.) f. Variable X_n is an upper record if $X_n > max\{X_1, X_2, \ldots, X_{n-1}\}$ and lower record if $X_n < min\{X_1, X_2, \ldots, X_{n-1}\}$. By convention X_1 is an upper and lower record value. Let $\{T_n, n \geqslant 1\}$ be the lower record times at which record values occur. We consider discrete time and define $T_1 = 1$ and $T_n = min\{i; i > T_{n-1}, X_i < X_{T_{n-1}}\}$. Further let $\{T^n, n \geqslant 1\}$ be the sequence of upper record times where again $T^1 = 1$ and $T^n = min\{i; i > T^{n-1}, X_i > X_{T^{n-1}}\}$. So $\{L_n, n \geqslant 1\} = \{X_{T_n}, n \geqslant 1\}$ is a sequence of lower record values and

* Corresponding author.
E-mail addresses: matej.juhas@student.upjs.sk (M. Juhás), valeria.skrivankova@upjs.sk (V. Skřivánková).

0096-3003/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.amc.2013.10.037

$\{R_n, n \geqslant 1\} = \{X_{T^n}, n \geqslant 1\}$ is a sequence of upper record values. The distribution of record values is given in terms of hazard function and hazard rate (see [4, page 2]).

The function $H(x)$ defined as $H(x) = -\ln F(x)$ is called hazard function for lower records. The function $h(x) = -\frac{dH(x)}{dx} = \frac{f(x)}{F(x)}$ is called hazard rate for lower records.

The function $H(x)$ defined as $H(x) = -\ln(1 - F(x))$ is called hazard function for upper records and the function $h(x) = \frac{dH(x)}{dx} = \frac{f(x)}{1-F(x)}$ is the hazard rate for upper records.

If $F_n(x)$ is the distribution function of random variable L_n (or R_n), $n \geqslant 1$ and $H(x)$ is the hazard function for lower records (or for upper records) according to the distribution function $F(x)$ of $X_n, n \geqslant 1$, then

$$F_n(x) = \int_{-\infty}^{x} \frac{H^{n-1}(u)}{(n-1)!} dF(u)$$

and the corresponding density function is

$$f_n(x) = \frac{H^{n-1}(x)}{(n-1)!} f(x). \tag{1}$$

The joint probability density function of $L_1, L_2, \ldots L_n$ is given by formula

$$f_{L_1,L_2,\ldots,L_n}(x_1, x_2, \ldots, x_n) = h(x_1)h(x_2)\ldots h(x_{n-1})f(x_n)$$

for $x_1 > x_2 > \cdots > x_n$ where h is hazard rate for lower records. The density of $R_1, R_2, \ldots R_n$ is

$$f_{R_1,R_2,\ldots,R_n}(x_1, x_2, \ldots, x_n) = h(x_1)h(x_2)\ldots h(x_{n-1})f(x_n)$$

for $x_1 < x_2 < \cdots < x_n$ and h is hazard rate for upper records.

The marginal density function of L_i, L_j (R_i, R_j) is given by

$$f_{i,j}(x_i, x_j) = \frac{(H(x_i))^{i-1}}{\Gamma(i)} h(x_i) \frac{(H(x_j) - H(x_i))^{j-i-1}}{\Gamma(j-i)} f(x_j) \tag{2}$$

for $-\infty < x_j < x_i < \infty$ $(-\infty < x_i < x_j < \infty)$.

2. Main results

Theorem 1. *Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of iid random variables with absolutely continuous distribution function $F(x), x \in (a, b), (a, b) \subseteq \mathbb{R}$ with $F(a) = 0$ and $F(b) = 1$. Let function $g : (a, b) \to (0, \infty)$ with properties: g is differentiable function, $g'(x) < 0$ for all $x \in (a, b), \lim_{x\to a^+} g(x) = \infty, \lim_{x\to b^-} g(x) = 0$. Then the distribution function of $X_1, X_2, \ldots$ is of the form $F(x) = e^{-cg(x)}, c > 0, x \in (a, b)$ if and only if random variables $g(L_n)$ and $g(L_{n+1}) - g(L_n), n \geqslant 1$ are independent.*

Remark 1. Before we give the proof of Theorem 1 we mention some facts. If $g'(x) < 0$ for all $x \in (a, b)$ then g is decreasing function and thus injective function and it holds $g'(g^{-1}(x))(g^{-1})'(x) = 1$. Interval (a, b) can be of the form $(-\infty, b), (a, \infty)$ or $(-\infty, \infty)$ so $a = -\infty$ or $b = \infty$. From assumptions of Theorem 1 holds $\lim_{x\to\infty} g^{-1}(x) = a$ and $\lim_{x\to 0^+} g^{-1}(x) = b$.

Proof. Let $F(x) = e^{-cg(x)}, x \in (a, b), c > 0$. Then $f(x) = -cg'(x)e^{-cg(x)}, H(x) = -\ln(F(x)) = cg(x)$ and $h(x) = -cg'(x)$. If we use relation (2) then the density of L_n, L_{n+1} is of the form

$$f_{n,n+1}(x, y) = \frac{(cg(x))^{n-1}}{\Gamma(n)} cg'(x) cg'(y) e^{-cg(y)}, \quad x, y \in (a, b).$$

Consider the transformation

$$t : \begin{pmatrix} L_n \\ L_{n+1} \end{pmatrix} \to \begin{pmatrix} g(L_n) \\ g(L_{n+1}) - g(L_n) \end{pmatrix} = \begin{pmatrix} U \\ V \end{pmatrix}, \quad \tau : \begin{pmatrix} U \\ V \end{pmatrix} \to \begin{pmatrix} g^{-1}(U) \\ g^{-1}(U + V) \end{pmatrix}. \tag{3}$$

The determinant of this transformation is

$$D_\tau = \begin{vmatrix} \frac{\partial g^{-1}(u)}{\partial u} & 0 \\ \frac{\partial g^{-1}(u+v)}{\partial u} & \frac{\partial g^{-1}(u+v)}{\partial v} \end{vmatrix} = \begin{vmatrix} (g^{-1})'(x)_{|x=u} \cdot 1 & 0 \\ (g^{-1})'(x)_{|x=u+v} \cdot 1 & (g^{-1})'(x)_{|x=u+v} \cdot 1 \end{vmatrix} = \begin{vmatrix} (g^{-1})'(u) & 0 \\ (g^{-1})'(u+v) & (g^{-1})'(u+v) \end{vmatrix} = (g^{-1})'(u)(g^{-1})'(u+v). \tag{4}$$

Then the density of U, V is given by formula