*Annales Univ. Sci. Budapest., Sect. Comp.* 50 *(2020) 57–61*

# AN EXPONENTIAL CHARACTERIZATION AND RELATED CONJECTURES

**Barry C. Arnold** (Riverside, USA)
**Jose A. Villasenor** (Montecillo, Mexico)

*To the memory of János Galambos*

Communicated by Imre Kátai

(Received January 24, 2020; accepted April 15, 2020)

**Abstract.** In a sequence of independent identically distributed exponential random varibles, the sum of the first two record values is distributed as a simple linear combination of exponential variables. The possiblity that this and certain related distributional properties might characterize the exponential distribution is investigated.

## 1. Introduction

Consider a sequence of independent identically distributed (i.i.d.) positive random variables $\{X_n\}_{n=1}^{\infty}$. Denote the corresponding sequence of upper records by $\{X^{(n)}\}_{n=1}^{\infty}$. Specifically, the first random variable in the sequence is identified as the first record, the second record is the first subsequent $X_n$ which exceeds $X_1$. It is well known that the record value sequence coresponding to a sequence of exponential variables has a particularly simple distributional structure. If we define the record spacings sequence $\{S_n\}_{n=1}^{\infty}$ by $S_1 = X_1 = X^{(1)}$ and for $n > 1$, $S_n = X^{(n)} - X^{(n-1)}$, then in the exponential case these spacings are independent random variables. Exponential characterizations based on

*Key words and phrases*: Upper reccords, hazard function, record spacings, monotone hazard function.
*2010 Mathematics Subject Classification*: 60E05, 62E10.

the independence of the record spacings are well known. In the present paper we will consider a simple relationship between the distribution of the first two records and the distribution of the first two $X_n$'s. To obtain a characterization, regularity conditions will be imposed. It is conjectured that these conditions can be relaxed. When we encountered this situation, we felt that the job of investigating whether the conditions could be relaxed might well be something that János would enjoy. Regrettably we will not be able to have input from him.

## 2. The conjectured characterizations

Consider a sequence of i.i.d. positive random variables $\{X_n\}_{n=1}^{\infty}$ with corresponding upper record sequence $\{X^{(n)}\}_{n=1}^{\infty}$. If the $X_i$'s have a common exponential distribution, then because the record spacings are themselves exponentially distributed with homogeneous intensity parameters, it follows that

$$X^{(1)} + X^{(2)} \stackrel{d}{=} X_1 + 2X_2. \tag{2.1}$$

After introspecting about this unusual relationship betwen the two sequences, $\{X_n\}_{n=1}^{\infty}$ and $\{X^{(n)}\}_{n=1}^{\infty}$, it became plausible that this was a characteristic property of the exponential distribution. Three conjectures were considered.

**Conjecture 2.1.** *Suppose that* $X^{(1)} + X^{(2)} \stackrel{d}{=} X_1 + 2X_2$, *then* $f_X(x) = \lambda e^{-\lambda x} I(x > 0)$ *for some* $\lambda > 0$.

**Conjecture 2.2.** *Suppose that, for some positive integer* $n$, $\sum_{i=1}^{n} X^{(i)} \stackrel{d}{=} \sum_{i=1}^{n} iX_i$, *then* $f_X(x) = \lambda e^{-\lambda x} I(x > 0)$ *for some* $\lambda > 0$.

**Conjecture 2.3.** *Suppose that, for every positive integer* $n$, $\sum_{i=1}^{n} X^{(i)} \stackrel{d}{=} \sum_{i=1}^{n} iX_i$, *then* $f_X(x) = \lambda e^{-\lambda x} I(x > 0)$ *for some* $\lambda > 0$.

All three conjectures are judged to be plausible. If Conjecture 2.1 could be proved then trivially Conjecture 2.3 would be also true. However it is conceivable that Conjecture 2.3 might be true even though Conjecture 2.1 is not true. Conjecture 2.2 would appear to be more difficult to resolve. In the next section we will provide a proof of conjecture 2.1 under certain regularity conditions. Under a different regularity condition, it may be shown that Conjecture 2.3 is trivially true. This observation is mentioned in the discusion section.

## 3. Proof of Conjecture 2.1, under regularity conditions

**Theorem 3.1.** *If $\{X_n\}_{n=1}^{\infty}$ are i.i.d. positive random variables with common density function $f(x)$, distribution function $F(x)$ and continuous monotone hazard function $h(x) = f(x)/[1-F(x)]$, and if $X^{(1)} + X^{(2)} \stackrel{d}{=} X_1 + 2X_2$, then $f(x) = \lambda e^{-\lambda x} I(x > 0)$ for some $\lambda > 0$.*

**Proof.** The joint density of the first two records is given by (see Arnold, Balakrishnan and Nagaraja (1998), for example)

$$f_{X^{(1)},X^{(2)}}(x_1, x_2) = h(x_1) f(x_2) I(0 < x_1 < x_2 < \infty). \tag{3.1}$$

If we define $Y_1 = X^{(1)}$ and $Y_2 = X^{(1)} + X^{(2)}$ , we obtain the joint density

$$f_{Y_1,Y_2}(y_1, y_2) = h(y_1) f(y_2 - y_1) I(0 < y_1 < \infty, 2y_1 < y_2 < \infty). \tag{3.2}$$

From this we obtain an expression for the marginal density of $Y_2$, i.e.,

$$f_{Y_2}(y_2) = \int_0^{y_2/2} h(y_1) f(y_2 - y_1) dy_1. \tag{3.3}$$

If we define $Y_3 = X_1 + 2X_2$, its density is given by the convolution formula

$$f_{Y_3}(y_3) = \int_0^{y_3/2} f(y_3 - 2y) f(y) dy, \tag{3.4}$$

But by hypothesis, $Y_2 \stackrel{d}{=} Y_3$, so, writing $z$ for $y_2$ and $y_3$ in the above equations, we have

$$\int_0^{z/2} h(y) f(z - y) dy = \int_0^{z/2} f(z - 2y) f(y) dy. \tag{3.5}$$

This equation is indeed satisfied if $f$ is an exponential density. We claim that the equation is only satisfied for such a choice of $f$.

Recalling the definition of a hazard function, and writing $\bar{F} = 1 - F$, we have

$$0 = \int_0^{z/2} [f(z - y)/\bar{F}(y) - f(z - 2y)] f(y) dy.$$

Then by the mean value theorem, there exists $y_0 \in (0, z/2)$ such that $\int_0^{z/2} [f(z-y)/\bar{F}(y) - f(z-2y)] f(y) dy = [f(z-y_0)/\bar{F}(y_0) - f(z-2y_0)] \int_0^{z/2} f(y) dy.$

Therefore, $0 = [f(z-y_0)/\bar{F}(y_0) - f(z-2y_0)]F(z/2)$. Hence, for $z > 0$,

$$0 = [f(z-y_0)/\bar{F}(y_0) - f(z-2y_0)].$$

If we let $u = z - y_0 > 0$ we have $0 = [f(u)/\bar{F}(y_0) - f(u-y_0)]$.

Hence $f(u) = f(u-y_0)\bar{F}(y_0)$.

Therefore, $\bar{F}(z) = \int_z^\infty f(u)du = \int_z^\infty f(u-y_0)\bar{F}(y_0)du$. Then considering the change of variable $w = u - y_0$ we have

$\bar{F}(z)/\bar{F}(y_0) = \int_{z-y_0}^\infty f(w)dw = \bar{F}(z-y_0)$.

That is,

$$\bar{F}(z) = \bar{F}(z-y_0)\bar{F}(y_0). \tag{3.6}$$

From (3.6), for $0 < y(h) < z + h$, we can write

$$\begin{aligned}\frac{\bar{F}(z+h)-\bar{F}(z)}{h} &= \frac{\bar{F}(z+h-y(h))\bar{F}(y(h)) - \bar{F}(z-y_0)\bar{F}(y_0)}{h} = \\ &= \frac{\bar{F}(y_0)\left[\bar{F}(z+h-y(h))\frac{\bar{F}(y(h)}{\bar{F}(y_0)} - \bar{F}(z-y_0)\right]}{h},\end{aligned} \tag{3.7}$$

in which $y(h) \to y_0$ as $h \to 0$ so that $\bar{F}(y(h))/\bar{F}(y_0) \to 1$ as $h \to 0$. Consequently, taking the limit as $h \to 0$ in (3.7) we obtain

$$-f(z) = \bar{F}(y_0)[-f(z-y_0)].$$

Since the hazard function $h(x) = f(z)/\bar{F}(z)$ it follows from this equation , together with (3.6) that $h(z) = h(z-y_0)$.

Thus far, we have that for every $z > 0$ there exists $y$ with $0 < y < z$ such that $h(y) = h(z)$. Since the hazard function is monotone, this implies that the hazard function $h(x)$ is constant on the interval $[y, z]$.

Fix $z = 1$. Consider the set

$A = \{a : h(x)$ is constant and equals $\lambda = h(1)$ on the interval $[a, 1]\}$.

This set is non-empty. Suppose that $\inf\{a \in A\} = a_0 > 0$. By the continuity of $h(x)$ we have $h(a_0) = \lambda$. However we can then find $a_1 < a_0$ with $h(x) = \lambda$ for every $x \in [a_1, a_0]$, so $a_1 \in A$ which contradicts the claim that $\inf\{a \in A\} = a_0$. Thus it must be true that $h(x) = \lambda$ for every $x \in (0, 1]$.

Next, let $n$ be an arbitrary positive integer, by an analogous argument we must have $h(x)$ constant on the inteval $(0, n]$, but then it must be equal to $\lambda$ throughout this interval. But $n$ is arbitrary, so we must have $h(x) = \lambda$ $\forall x \in (0, \infty)$. The constant $\lambda$ must be positive in order that $h(x)$ will be a valid hazard function. It then follows that $X_i \sim \exp(\lambda)$, as claimed. ■

## 4. Discussion

The proof of Theorem 3.1 was more convoluted than was expected. It did use crucially the assumptions that $h(x)$ was continuous and monotone. Relaxation of these conditions, as mentioned in the introduction, might well have appealed to János. Conjecture 2.2 does not appear to be readily resolved even under quite strong regularity conditions. In contrast, Conjecture 2.3 is amenable to resolution under a simple moment condition. Thus, if we assume that the $X_i$'s have a finite $(1+\delta)$'th moment, which implies that first moments of the $X_i$'s and of the records exist, then $\sum_{i=1}^n X^{(i)} \stackrel{d}{=} \sum_{i=1}^n iX_i$ implies that $E(\sum_{i=1}^n X^{(i)}) = E(\sum_{i=1}^n iX_i)$. If this holds for every $n$, then it follows that $E(X^{(n)}) = nE(X_1)$ for every $n$, and a common exponential distribution is ensured since the sequence of expected records determines the parent distribution (Kirmani and Beg, 1984). Whether the moment condition used here can be relaxed remains open.

## References

[1] **Arnold, B.C., N. Balakrishnan and H.N. Nagaraja,** *Records*, Wiley, New York, 1998.

[2] **Kirmani, S.N.U.A. and M.I. Beg,** On characterization of distributions by expected records. *Sankhya*, **A.46** (1984), 463–465.

**B. C. Arnold**
University of California
Riverside
USA
barry.arnold@ucr.edu

**J. A. Villasenor**
Colegio de Postgraduados
Montecillo
Mexico
jvillasr@colpos.mx