2013

Characterizations of Exponential Distribution Based On Sample of Size Three

George Yanev
The University of Texas Rio Grande Valley

Santanu Chakraborty

Pliska Stud. Math. Bulgar. **23** (2013), 237–244

PLISKA
STUDIA MATHEMATICA
BULGARICA

CHARACTERIZATIONS OF EXPONENTIAL DISTRIBUTION BASED ON SAMPLE OF SIZE THREE

George P. Yanev, Santanu Chakraborty

Abstract. Two characterizations of the exponential distribution based on equalities among order statistics in a random sample of size three are proved. This proves two conjectures stated recently in Arnold and Villaseñor [4].

1. Introduction. The publications on characterizations of the exponential distribution are abundant. Comprehensive surveys can be found in Ahsanullah and Hamedani [1], Arnold and Huang [3], and Johnson, Kotz and Balakrishnan [5]. The Bulgarian probability school has its contribution with the works of Obretenov [6]-[8]. Recently, Arnold and Villaseñor [4] obtained a series of characterizations based on random sample of size two from a continuous distribution. They also identified a list of conjectures for possible extensions of their results to samples of size three and bigger. In this note we confirm that two of these conjectures are true.

Let X_1, X_2, X_3 be a random sample of size three from a parent random variable X. Denote $X_{2:2} := \max\{X_1, X_2\}$ and $X_{3:3} := \max\{X_1, X_2, X_3\}$. We write $X \sim \exp\{\lambda\}$ if the probability density function (pdf) of X equals $f_X(x) = \lambda e^{-\lambda x} I(x > 0)$, $\lambda > 0$. It is known (e.g., Arnold et al. (2008), p.77) that if $X \sim \exp\{\lambda\}$, then

$$X_1 + \frac{1}{2}X_2 + \frac{1}{3}X_3 \stackrel{d}{=} X_{3:3} \quad \text{and} \quad X_{2:2} + \frac{1}{3}X_3 \stackrel{d}{=} X_{3:3}, \tag{1}$$

2010 *Mathematics Subject Classification:* 62G30, 62E10.
Key words: characterization, exponential distribution, order statistics.

where $\stackrel{d}{=}$ denotes equality in distribution. Arnold and Villaseñor [4] conjectured that each one of the equalities in (1), under some regularity assumptions on the cumulative distribution function (cdf) F of X, is a sufficient condition for $X \sim \exp(\lambda)$ for some $\lambda > 0$. The theorem below proves these conjectures.

Theorem 1. *Let X_1, X_2, X_3 be a random sample from X, which has an absolutely continuous cdf F with $F(0) = 0$. Suppose the pdf f of X is analytic in a neighborhood of* 0.

(i) If

$$X_{2:2} + \frac{1}{3}X_3 \stackrel{d}{=} X_{3:3}, \tag{2}$$

then $X \sim \exp\{\lambda\}$ for some $\lambda > 0$.

(ii) If

$$X_1 + \frac{1}{2}X_2 + \frac{1}{3}X_3 \stackrel{d}{=} X_{3:3}, \tag{3}$$

then $X \sim \exp\{\lambda\}$ for some $\lambda > 0$.

2. Proofs. We begin with a useful lemma (see also Arnold and Villaseñor [4]).

Lemma 1. *If $F(0) = 0$, the pdf f is analytic in a neighborhood of 0, and*

$$f^{(k)}(0) = \left[\frac{f'(0)}{f(0)}\right]^{k-1} f'(0), \qquad k = 1, 2, \ldots, \tag{4}$$

then $X \sim \exp\{\lambda\}$ for some $\lambda > 0$.

P r o o f. For the Maclaurin series of $f(x)$, we have for $x > 0$

$$\begin{aligned} f(x) &= \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k \\ &= f(0) + \sum_{k=1}^{\infty} \left[\frac{f'(0)}{f(0)}\right]^{k-1} f'(0) \frac{x^k}{k!} \\ &= f(0) \exp\left\{\frac{f'(0)}{f(0)} x\right\}. \end{aligned} \tag{5}$$

Since $f(x)$ is a pdf, we have $f'(0)/f(0) < 0$. Denoting $\lambda = -f'(0)/f(0) > 0$ and setting the integral of (5) from 0 to ∞ to be 1, we obtain $\lambda = f(0)$. Therefore, $f(x) = \lambda e^{-\lambda x} I(x > 0)$, i.e., $X \sim \exp\{\lambda\}$. $\square$

We continue with the proof of the theorem.

P r o o f o f P a r t (i). The pdf of the left-hand side of (2) is

$$\begin{aligned} f_{X_{2:2}+X_3/3}(x) &= \int_0^x f_{X_3/3}(y) f_{X_{2:2}}(x-y)\,dy \\ &= \int_0^x 3f(3y) \frac{d}{dx}[F^2(x-y)]\,dy \\ &= 6\int_0^x f(3y)F(x-y)f(x-y)\,dy. \end{aligned} \tag{6}$$

For the pdf of the right-hand side of (2), we have

$$\begin{aligned} f_{X_{3:3}}(x) &= \frac{d}{dx} F^3(x) \\ &= 3F^2(x)f(x) \\ &= 6f(x)\int_0^x F(y)f(y)\,dy. \end{aligned} \tag{7}$$

Define $G(x) := F(x)f(x)$. Referring to (6) and (7) we rewrite (2) as

$$\int_0^x f(3y)G(x-y)\,dy = f(x)\int_0^x G(y)\,dy. \tag{8}$$

For the nth derivative of the left-hand side of (8), we have

$$\frac{d^n}{dx^n}\int_0^x f(3y)G(x-y)\,dy = \sum_{i=0}^{n-1}[f(3x)]^{(n-1-i)}G^{(i)}(0) + \int_0^x f(3y)G^{(n)}(x-y)\,dy.$$

Applying the Leibnitz rule for the nth derivative of a product of two functions to the right-hand side of (8), we obtain

$$\frac{d^n}{dx^n}\left[f(x)\int_0^x G(y)\,dy\right] = \sum_{i=1}^{n}\binom{n}{i}[f(x)]^{(n-i)}G^{(i-1)}(x) + [f(x)]^{(n)}\int_0^x G(y)\,dy.$$

Now, differentiating both sides of (8) n times and evaluating the derivatives at $x = 0$, we obtain

$$\sum_{i=1}^{n} 3^{n-i} f^{(n-i)}(0) G^{(i-1)}(0) = \sum_{i=1}^{n} \binom{n}{i} f^{(n-i)}(0) G^{(i-1)}(0). \tag{9}$$

Since $G(0) = 0$ and $G'(0) = f^2(0)$, the above equation is equivalent to

$$\left[3^{n-2} - \binom{n}{2}\right] f^{(n-2)}(0) f^2(0) = \sum_{i=3}^{n} \left[\binom{n}{i} - 3^{n-i}\right] f^{(n-i)}(0) G^{(i-1)}(0), \tag{10}$$

where $n \geq 4$. We shall prove that (10) implies (4). Equation (4) is trivially true for $k = 1$. To proceed by induction, assume (4) for all $1 \leq k \leq n-3$, where $n \geq 4$. We need to prove it for $k = n-2$. Using the induction assumption, we have for $j = 1, 2, \ldots, n-2$

$$\begin{aligned} G^{(j)}(0) &= \sum_{i=0}^{j} \binom{j}{i} F^{(i)}(0) f^{(j-i)}(0) \\ &= \sum_{i=1}^{j} \binom{j}{i} f^{(i-1)}(0) f^{(j-i)}(0) \\ &= (j+1) f^{(j-1)}(0) f(0) + \sum_{i=2}^{j-1} \binom{j}{i} \left[\frac{f'(0)}{f(0)}\right]^{i-2} f'(0) \left[\frac{f'(0)}{f(0)}\right]^{j-i-1} f'(0) \\ &= \left[\frac{f'(0)}{f(0)}\right]^{j-1} f^2(0)(2^j - 1). \end{aligned}$$

Therefore, using the induction assumption again, we have for $i = 3, 4, \ldots, n-1$

$$\begin{aligned} f^{(n-i)}(0) G^{(i-1)}(0) &= \left[\frac{f'(0)}{f(0)}\right]^{n-i-1} f'(0) \left[\frac{f'(0)}{f(0)}\right]^{i-2} f^2(0)(2^{i-1} - 1) \\ &= \left[\frac{f'(0)}{f(0)}\right]^{n-3} f'(0) f^2(0)(2^{i-1} - 1). \end{aligned} \tag{11}$$

Substituting (11) in the right-hand side of (10) yields ($i = n$ corresponds to a 0 term)

$$\left[3^{n-2} - \binom{n}{2}\right] f^{(n-2)}(0) = \left[\frac{f'(0)}{f(0)}\right]^{n-3} f'(0) \sum_{i=3}^{n} \left[\binom{n}{i} - 3^{n-i}\right] (2^{i-1} - 1).$$

Thus, to prove (4) for $k = n-2$ it is sufficient to show that

$$3^{n-2} - \binom{n}{2} = \sum_{i=3}^{n} \left[\binom{n}{i} - 3^{n-i} \right] (2^{i-1} - 1)$$

or, equivalently,

$$\sum_{i=2}^{n} 3^{n-i}(2^{i-1} - 1) = \sum_{i=2}^{n} \binom{n}{i} (2^{i-1} - 1),$$

which is easily verified. This completes the proof of (4) by induction. The claim in Part (i) follows from (4) and the Lemma.

Proof of Part (ii). The pdf of the left-hand side of (3) is

$$\begin{aligned} (12)\quad f_{X_1+X_2/2+X_3/3}(x) &= \int_0^x f_{X_1}(y) f_{X_2/2+X_3/3}(x-y)\,dy \\ &= \int_0^x f_{X_1}(y) \int_0^{x-y} f_{X_2/2}(z) f_{X_3/3}(x-y-z)\,dz\,dy \\ &= 6\int_0^x f(y) \int_0^{x-y} f(2z) f(3(x-y-z))\,dz\,dy. \end{aligned}$$

Denoting

$$H(x-y) := \int_0^{x-y} f(2z) f(3(x-y-z))\,dz \tag{13}$$

and taking into account (7) and (12), we write (3) as

$$\int_0^x f(y) H(x-y)\,dy = f(x) \int_0^x G(y)\,dy. \tag{14}$$

Similarly to the proof of Part (i), differentiating n times with respect to x both sides of (14) and evaluating the derivatives at $x = 0$, we have

$$\sum_{i=1}^{n} f^{(n-i)}(0) H^{(i-1)}(0) = \sum_{i=1}^{n} \binom{n}{i} f^{(n-i)}(0) G^{(i-1)}(0).$$

Since $H(0) = G(0) = 0$ and $H'(0) = G'(0) = f^2(0)$, the last equation becomes

$$\left[1 - \binom{n}{2} \right] f^{(n-2)}(0) f^2(0) = \sum_{i=3}^{n} \left[\binom{n}{i} G^{(i-1)}(0) - H^{(i-1)}(0) \right] f^{(n-i)}(0). \tag{15}$$

Now we are in a position to prove (4) by induction. Equation (4) is trivially true for $k = 1$. Assuming (4) for all $1 \leq k \leq n-3$, where $n \geq 4$, we will prove it for $k = n-2$. Differentiating (13) n times with respect to x and evaluating the derivative at $x = y$, we have

$$H^{(n)}(0) = \sum_{i=1}^{n} 2^{n-i} f^{(n-i)}(0) 3^{i-1} f^{(i-1)}(0). \tag{16}$$

Under the induction assumption, (16) implies for $j = 1, 2, \ldots, n-2$

$$\begin{aligned} H^{(j)}(0) &= \sum_{i=1}^{j} 2^{j-i} \left[\frac{f'(0)}{f(0)}\right]^{j-i-1} f'(0) 3^{i-1} \left[\frac{f'(0)}{f(0)}\right]^{i-2} f'(0) \\ &= \left[\frac{f'(0)}{f(0)}\right]^{j-3} (f'(0))^2 \sum_{i=1}^{j} 2^{j-i} 3^{i-1} \\ &= \left[\frac{f'(0)}{f(0)}\right]^{j-1} f^2(0) \left(3^j - 2^j\right). \end{aligned}$$

Therefore, using the induction assumption again, we have for $i = 3, 4, \ldots, n-1$

$$\begin{aligned} f^{(n-i)}(0) H^{(i-1)}(0) &= \left[\frac{f'(0)}{f(0)}\right]^{n-i-1} f'(0) \left[\frac{f'(0)}{f(0)}\right]^{i-2} f^2(0) \left(3^{i-1} - 2^{i-1}\right) \\ &= \left[\frac{f'(0)}{f(0)}\right]^{n-3} f'(0) f^2(0) \left(3^{i-1} - 2^{i-1}\right). \end{aligned}$$

Recalling (11) from the proof of Part (i) we rewrite (15) as (note that $i = n$ corresponds to a 0 term)

$$\left[1 - \binom{n}{2}\right] f^{(n-2)}(0) = \left[\frac{f'(0)}{f(0)}\right]^{n-3} f'(0) \sum_{i=3}^{n} \left[\binom{n}{i}(2^{i-1} - 1) - (3^{i-1} - 2^{i-1})\right]$$

Thus, to prove (4) for $k = n-2$ it is sufficient to show that

$$1 - \binom{n}{2} = \sum_{i=3}^{n} \left[\binom{n}{i}(2^{i-1} - 1) - (3^{i-1} - 2^{i-1})\right],$$

or equivalently

$$\sum_{i=2}^{n} \left[\binom{n}{i}(2^{i-1} - 1) - 3^{i-1} + 2^{i-1}\right] = 0$$

which is easily verified. This proves (4). Now, referring to the Lemma we complete the proof of the Theorem. □

3. Concluding remarks. The more general cases of samples of size $n \geq 4$ and relations

$$X_{n-1:n-1} + \frac{1}{n}X_n \stackrel{d}{=} X_{n:n} \quad \text{and} \quad X_{n-2:n-2} + \frac{1}{n-1}X_{n-1} + \frac{1}{n}X_n \stackrel{d}{=} X_{n:n},$$

where $X_{j:j} := \max\{X_1, X_2, ..., X_j\}$ for $j = n-1$ and $j = n$ are still under investigation. Finally, it is worth noticing that if we assume for i.i.d. random variables $X_1, X_2, \ldots$ with $E|X_1| < \infty$, that for every $n = 1, 2, \ldots$

$$X_1 + \frac{1}{2}X_2 + \frac{1}{3}X_3 + \ldots + \frac{1}{n}X_n \stackrel{d}{=} X_{n:n},$$

then the X_i's have a common exponential distribution (see e.g. Arnold and Villaseñor [4]).

REFERENCES

[1] M. Ahsanullah, G. G. Hamedani. Exponential Distribution: Theory and Methods. NOVA Science, New York, 2010.

[2] B. C. Arnold, N. Balakrishnan, H. N. Nagaraja. A First Course in Order Statistics. SIAM Classics in Applied Probab. vol. **54**, Philadelphia, 2008.

[3] B. C. Arnold, J. S. Huang. Chapter 12: Characterizations. In: , The Exponential Distribution: Theory, Methods and Applications (Eds N. Balakrishnan, A. P. Basu). Gordon and Breach, Amsterdam, 1995, 185–203.

[4] B. C. Arnold, J. A. Villaseñor. Exponential characterizations motivated by the structure of order statistics in sample of size two. *Statistics and Probability Letters* **83** (2013), 596–601.

[5] N. L. Johnson, S. Kotz, N. Balakrishnan. Continuous Univariate Distributions, Vol. 1, 2nd Edn. Wiley, New York, 1994.

[6] A. Obretenov. A property of the exponential distribution. *Physics-Mathematics Journal, Bulgarian Acad. Sci.* **13** (1970), No 46, 51–53 (in Bulgarian).

[7] A. Obretenov. Characterizations of exponential and geometric distributions based on order statistics. *Pliska Stud. Math. Bulgar.* **7** (1984), 97–101.

[8] A. Obretenov. A characterization of an exponential distribution based on renewals. In: Exploring stochastic laws. Festschrift in honour of the 70th birthday of Academician Vladimir Semenovich Korolyuk (Eds A. V. Skorokhod et al.) VSP, Utrecht, 1995, 351–355.

George P. Yanev, Santanu Chakraborty
Department of Mathematics
University of Texas – Pan American
1201 W. University Drive
Edinburg, Texas 78539
e-mail: `yanevgp@utpa.edu`
`schakraborty@utpa.edu`