Three Characterizations of Exponential Distribution Involving Median of Sample of Size Three

Marko Obradović

Faculty of Mathematics, University of Belgrade, Studentski trg 16
Belgrade, Serbia
marcone@matf.bg.ac.rs

In this paper three new characterizing theorems of exponential distribution are presented. They are based on equidistribution of some functions of order statistics. All of them include the median of sample of size three.

Keywords: Characterization, exponential distribution, order statistics

2010 Mathematics Subject Classification: 62E10, 62G30.

1. Introduction

The characterization theorems have been very popular lately. Since exponential distribution has very wide applications, a major part of work on characterizations has been dedicated to this distribution. Many examples could be found in [1] and [4].

Recently Arnold and Villasenor [2] proposed a series of characterizations based on sample of size two and made some conjectures on their generalization. They also proposed a new method of proof which can be used when the density in question is analytic. Later Yanev and Chakraborty [5] proved by this method two characterization theorems concerning maximum of sample of size three. One of the benefits of these characterizations for small sample sizes is that they are suitable for creation of goodness of fit tests.

We propose three new characterization theorems whose common point is that in all of them appears the median of sample of size three. In some sense these characterizations could be considered as a generalization of the characterizations from [2] and [5].

2. Main Results

Let $\mathscr{F}$ be a class of absolutely continuous distribution functions F such that $F(0) = 0$ and whose density function f allows expansion in Maclaurin's series for all $x > 0$.

Marko Obradović

Firstly we present two lemmas which are important for the proofs of the theorems. They are very similar to lemmas proposed in [2] and [5] with an exception that the first derivative of the density function f at zero is also expressed via $f(0)$.

Lemma 2.1. *Let F be a distribution function from $\mathscr{F}$. If for all natural k holds:*

$$f^{(k)}(0) = (-1)^k f^{k+1}(0), \tag{2.1}$$

then $f(x) = \lambda e^{-\lambda x}$ for some $\lambda > 0$.

Proof. Expanding the function f in Maclaurin series for positive values of x we get:

$$f(x) = \sum_{k=0}^{\infty} f^{(k)}(0)\frac{x^k}{k!} = \sum_{k=0}^{\infty}(-1)^k f^{k+1}(0)\frac{x^k}{k!} = f(0)e^{-f(0)x}. \tag{2.2}$$

For $f(0) > 0$ this is the density of exponential distribution with $\lambda = f(0)$. □

Lemma 2.2. *Let F be a distribution from $\mathscr{F}$. Denote $G(x) = F(x)f(x)$ and $H(x) = F^2(x)f(x)$. Let the condition* (2.1) *be satisfied for $k \leq r-2$. Then the following equalities hold:*

$$G^{(k)}(0) = (-1)^{k-1}f^{k+1}(0)(2^k-1), \ \ 1 \leq k \leq r-1. \tag{2.3}$$
$$H^{(k)}(0) = (-1)^{k-2}f^{k+1}(0)(3^k - 2^{k+1}+1), \ \ 1 \leq k \leq r. \tag{2.4}$$

Proof. Applying Leibniz formula for the derivative of a product to $G(x)$ we get:

$$G^{(k)}(x) = \sum_{j=0}^{k}\binom{k}{j}F^{(j)}(x)f^{(k-j)}(x). \tag{2.5}$$

When we put $x = 0$, all summands that have $F^{(0)}(x)$ will be equal to zero, therefore we get

$$G^{(k)}(0) = \sum_{j=1}^{k}\binom{k}{j}f^{(j-1)}(0)f^{(k-j)}(0)$$

$$= \sum_{j=1}^{k}\binom{k}{j}(-1)^{j-1}f^j(0)(-1)^{k-j}f^{k-j+1}(0)$$

$$= (-1)^{k-1}f^{k+1}(0)\sum_{j=1}^{k}\binom{k}{j}$$

$$= (-1)^{k-1}f^{k+1}(0)(2^k-1).$$

Note that the expression on the right hand side of (2.4) is equal to zero for $k = 1$. Since $H(x) = F^2(x)f(x) = F(x)F(x)f(x)$, applying Leibniz formula for the derivative of product of three functions we obtain:

$$H^{(k)}(x) = \sum_{s=0}^{k}\sum_{j=0}^{s}\binom{k}{j, s-j, k-s}F^{(j)}(x)F^{(s-j)}(x)f^{(k-s)}(x). \tag{2.6}$$

Putting $x = 0$ and using the same argument as in the previous case we get that $H'(0) = 0$ and for $k \geq 2$

$$H^{(k)}(0) = \sum_{s=2}^{k} \sum_{j=1}^{s} \binom{k}{j, s-j, k-s} f^{(j-1)}(0) f^{(s-j-1)}(0) f^{(k-s)}(0)$$

$$= \sum_{s=2}^{k} \sum_{j=1}^{s-1} \binom{k}{j, s-j, k-s} (-1)^{j-1} f^j(0)(-1)^{s-j-1} f^{s-j}(0)(-1)^{k-s} f^{k-s+1}(0)$$

$$= (-1)^{k-2} f^{k+1}(0) \sum_{s=2}^{k} \sum_{j=1}^{s-1} \binom{k}{j, s-j, k-s}$$

$$= (-1)^{k-2} f^{k+1}(0) \left(3^k - \sum_{s=0}^{1} \sum_{j=0}^{s} \binom{k}{j, s-j, k-s} - 2 \sum_{s=2}^{k} \binom{k}{s} \right)$$

$$= (-1)^{k-2} f^{k+1}(0)(3^k - 2^{k+1} + 1).$$

$\square$

Remark 2.1. Note that the expressions for $G'(0) = f^2(0)$ and $H''(0) = 2f^3(0)$ are always true regardless of the condition of the theorem. Besides, $G(0) = H(0) = H'(0) = 0$.

Let $X_{(k;n)}$ be the k th order statistics from the sample of size n. We now state and prove the characterization theorems.

Theorem 2.1. *Let X_1, X_2, X_3 be independent random variables from the distribution $F \in \mathscr{F}$. If*

$$\frac{1}{3} X_1 + \frac{1}{2} X_2 \overset{d}{=} X_{(2;3)} \tag{2.7}$$

then $X \sim \mathscr{E}(\lambda), \lambda > 0$.

Proof. Equalizing the densities on the left and the right hand side we get

$$\int_0^x 3f(3y) 2f(2(x-y)) dy = 6F(x)(1 - F(x)) f(x)$$

$$\int_0^x f(3y) f(2(x-y)) dy = f(x) \int_0^x f(y)(1 - 2F(y)) dy$$

$$\int_0^x f(3y) f(2(x-y)) dy = f(x) \int_0^x (f(y) - 2G(y)) dy. \tag{2.8}$$

We shall prove by induction that from (2.8) follows (2.1). Differentiating both sides twice with respect to x we get

$$f'(3x) 3f(0) + 2f(3x) f'(0) + 4 \int_0^x f(3y) f''(2(x-y)) dy$$

$$= 2f(x) f'(x) - 2G'(x) f(x) - 2G(x) f'(x) + (f(x) - 2G(x)) f'(x)$$

$$+ f''(x) \int_0^x (f(y) - 2G(y)) dy.$$

Letting $x = 0$ we get

$$5f'(0)f(0) = 3f(0)f'(0) - 2f(0)G'(0).$$

Using lemma 2.2 we get

$$f'(0) = -f^2(0).$$

Hence (2.1) is satisfied for $k = 1$. Let us now suppose that (2.1) is satisfied for $1 \leq k \leq r - 2$. Differentiating both sides of (2.8) r times we get

$$\sum_{j=0}^{r-1} 3^{r-1-j} f^{(r-1-j)}(3x) 2^j f^{(j)}(0) + \int_0^x f(3y) 2^r f^{(r)}(2(x-y)) dy$$

$$= \sum_{j=1}^{r} \binom{r}{j} f^{(r-j)}(x)(f^{(j-1)}(x) - 2G^{(j-1)}(x)) + f^{(r)}(x) \int_0^x (f(y) - 2G(y)) dy.$$

Letting $x = 0$ and eliminating zero terms we get

$$\sum_{j=1}^{r} 3^{r-j} f^{(r-j)}(0) 2^{j-1} f^{(j-1)}(0) = \sum_{j=1}^{r} \binom{r}{j} f^{(r-j)}(0)(f^{(j-1)}(0) - 2G^{(j-1)}(0)). \qquad (2.9)$$

Extracting the summands which contain $f^{(r-1)}(0)$ and grouping them to one side we get

$$f^{(r-1)}(0)f(0)(r+1-3^{r-1}-2^{r-1})$$
$$= \sum_{j=2}^{r-1} f^{(r-j)}(0) 3^{r-j} 2^{j-1} f^{(j-1)}(0) - \sum_{j=2}^{r-1} \binom{r}{j} f^{(r-j)}(0)(f^{(j-1)}(0) - 2G^{(j-1)}(0)).$$

Using the induction hypothesis and lemmas 2.1 and 2.2 we obtain

$$f^{(r-1)}(0)(r+1-3^{r-1}-2^{r-1}) = (-1)^{r-1} f^r(0) \left(\sum_{j=2}^{r-1} \left(3^{r-j} 2^{j-1} - \binom{r}{j} 2^{j-1} \right) - 2(2^{r-1} - 1) \right).$$

To prove that (2.1) is satisfied for $k = r - 1$ it suffices to show that

$$r+1-3^{r-1}-2^{r-1} = \sum_{j=2}^{r-1} \left(3^{r-j} 2^{j-1} - \binom{r}{j} 2^{j-1} \right) - 2(2^{r-1} - 1),$$

or equivalently

$$\sum_{j=1}^{r} 3^{r-j} 2^{j-1} = \sum_{j=1}^{r} \binom{r}{j} (2^j - 1).$$

It can be easily shown that both sums are equal to $3^r - 2^r$ which completes the proof. $\qquad \square$

Theorem 2.2. *Let X_0, X_1, X_2, X_3 be independent random variables from the distribution $F \in \mathscr{F}$. If*

$$X_0 + X_{(2;3)} \stackrel{d}{=} X_{(3;3)} \qquad (2.10)$$

then $X \sim \mathscr{E}(\lambda), \lambda > 0$.

Proof. Equalizing the respective densities we get:

$$\int_0^x f(y)6F(x-y)(1-F(x-y))f(x-y)dy = 3F^2(x)f(x)$$

$$6\int_0^x f(y)F(x-y)(1-F(x-y))f(x-y)dy = 6f(x)\int_0^x F(y)f(y)dy$$

$$\int_0^x f(y)G(x-y)dy - \int_0^x f(y)H(x-y)dy = f(x)\int_0^x G(y)dy. \tag{2.11}$$

As in the previous theorem we shall prove by induction that from (2.11) follows (2.1). Differentiating both sides three times with respect to x we get

$$f''(x)G(0) + f'(x)G'(0) + f(x)G''(0) + \int_0^x f(x)G^{(3)}(x-y)dy$$

$$- \left(f''(x)H(0) + f'(x)H'(0) + f(x)H''(0) + \int_0^x f(x)H^{(3)}(x-y)dy\right)$$

$$= 3f''(x)G(0) + 3f'(x)G'(0) + f(x)G''(0) + f^{(3)}\int_0^x G(y)dy.$$

Inserting $x = 0$ we get

$$f'(0)G'(0) - f(0)H''(0) = 3f'(0)G'(0),$$

and hence

$$f'(0) = -f^2(0).$$

Thus we got that (2.1) is satisfied for $k = 1$. Suppose now that (2.1) is satisfied for $1 \le k \le r-3$. Differentiating both sides of (2.11) r times we get

$$\sum_{j=0}^{r-1} f^{(r-1-j)}(x)G^{(j)}(0) + \int_0^x f(y)G^{(r)}(x-y)dy$$

$$- \sum_{j=0}^{r-1} f^{(r-1-j)}(x)H^{(j)}(0) - \int_0^x f(y)H^{(r)}(x-y)dy$$

$$= \sum_{j=1}^r \binom{r}{j} f^{(r-j)}(x)G^{(j-1)}(x) + f^{(r)}(x)\int_0^x G(y)dy.$$

Putting $x = 0$ and eliminating zero terms we get

$$\sum_{j=2}^{r} f^{(r-j)}(0)G^{(j-1)}(0) - \sum_{j=2}^{r-1} f^{(r-1-j)}(0)H^{(j)}(0) = \sum_{j=2}^{r} \binom{r}{j} f^{(r-j)}(0)G^{(j-1)}(0)$$

The terms for $j = r$ are equal on both sides so they cancel out. Extracting the summands containing $f^{(r-2)}(0)$ and grouping them on one side we get

$$f^{(r-2)}(0)G'(0)\left(1 - \binom{r}{2}\right)$$
$$= \sum_{j=3}^{r-1} \left(\binom{r}{j} - 1\right) f^{(r-j)}(0)G^{(j-1)}(0) + \sum_{j=2}^{r-1} f^{(r-1-j)}(0)H^{(j)}(0)$$

Using the induction hypothesis we obtain

$$f^{(r-2)}(0)\left(1 - \binom{r}{2}\right)$$
$$= (-1)^{r-2} f^{r-1}(0)\left(\sum_{j=3}^{r-1}(2^{j-1} - 1)\left(\binom{r}{j} - 1\right) - \sum_{j=2}^{r-1}(3^j - 2^{j+1} + 1)\right),$$

To prove that the condition (2.1) is satisfied for $k = r - 2$ it remains to show that

$$1 - \binom{r}{2} = \left(-2 + \sum_{j=3}^{r-1}(2^{j-1} - 1)\left(\binom{r}{j} - 1\right) - (3^j - 2^{j+1} + 1)\right),$$

or, equivalently,

$$\sum_{j=2}^{r-1}(2^{j-1} - 1)\left(\binom{r}{j} - 1\right) = \sum_{j=2}^{r-1}(3^j - 2^{j+1} + 1). \tag{2.12}$$

It can easily be calculated that both sums are equal to $\frac{3^r}{2} - 2^{r+1} + r + \frac{3}{2}$, which completes the proof. $\qquad\square$

In the following theorem appear two order statistics from the samples of different sizes.

Theorem 2.3. *Let X_1, X_2, X_3, X_4, and X_1', X_2', X_3', X_4' be two sets of independent random variables from the distribution $F \in \mathscr{F}$. If*

$$X_{(2;3)} + \frac{1}{4}X_4 \stackrel{d}{=} X_{(3;4)}, \tag{2.13}$$

where $X_{(2;3)}$ is taken from the first three variables of the first sample and $X_{(3;4)}$ from the second sample, then $X \sim \mathscr{E}(\lambda), \lambda > 0$.

Proof. Equalizing the respective densities we get:

$$\int_0^x 6F(x-y)(1-F(x-y))f(x-y)4f(4y)dy = 12F^2(x)(1-F(x))f(x)$$

$$2\int_0^x f(4y)F(x-y)(1-F(x-y))f(x-y)dy = f(x)\int_0^x (F^2(y)-F^3(y))f(y)dy$$

$$2\int_0^x f(4y)(G(x-y)-H(x-y))dy = f(x)\int_0^x (2G(y)-3H(y))dy. \qquad (2.14)$$

As in the previous cases, we shall prove that from (2.14) follows (2.1).

After evaluating the third derivative of both sides at zero and eliminating zero terms we get

$$8f'(0)G'(0) - 2f(0)H''(0) = 6f'(0)G'(0) - 3f(x)H''(0)$$

and hence again

$$f'(0) = -f^2(0).$$

The condition (2.1) is thus valid for $k=1$. Suppose now that it is satisfied for $1 \le k \le r-3$. Differentiating both sides of (2.14) r times we get

$$2\sum_{j=0}^{r-1} 4^{r-1-j}f^{(r-1-j)}(4x)(G^{(j)}(0)-H^{(j)}(0)) + 2\int_0^x 4^r f^{(r)}(4y)(G(x-y)-H(x-y))dy$$

$$= \sum_{j=1}^{r}\binom{r}{j}f^{(r-j)}(x)(2G^{(j-1)}(x)-3H^{(j-1)}(x)) + f^r(x)\int_0^x (2G(y)-3H(y))dy.$$

Letting $x=0$ and eliminating zero terms we get

$$2\sum_{j=2}^{r-1} 4^{r-j}f^{(r-j)}(0)G^{(j-1)}(0) - 2\sum_{j=2}^{r-1} 4^{r-1-j}f^{(r-1-j)}(0)H^{(j)}(0)$$

$$= 2\sum_{j=2}^{r}\binom{r}{j}f^{(r-j)}(0)G^{(j-1)}(0) - 3\sum_{j=2}^{r-1}\binom{r}{j+1}f^{(r-1-j)}(0)H^{(j)}(0).$$

The terms for $j=r$ in the first and the third sum coincide so they cancel out. Extracting terms containing $f^{(r-2)}(0)$ on one side we get

$$2f^{(r-2)}(0)f^2(0)\left(4^{r-2}-\binom{r}{2}\right)$$

$$= 2\sum_{j=3}^{r-1} f^{(r-j)}(0)G^{(j-1)}(0)\left(\binom{r}{j}-4^{r-j}\right)$$

$$- \sum_{j=2}^{r-1} f^{(r-1-j)}(0)H^{(j)}(0)\left(3\binom{r}{j+1}-2\cdot 4^{r-1-j}\right).$$

Applying the induction hypothesis to the right hand side we get

$$2f^{(r-2)}(0)\left(4^{r-2} - \binom{r}{2}\right)$$

$$= (-1)^{r-2}f^{r-1}(0)\left(2\sum_{j=3}^{r-1}(2^{j-1}-1)\left(\binom{r}{j} - 4^{r-j}\right)\right.$$

$$\left. + \sum_{j=2}^{r-1}(3^j - 2^{j+1} + 1)\left(3\binom{r}{j+1} - 2 \cdot 4^{r-1-j}\right)\right).$$

It remains to prove

$$\sum_{j=2}^{r-1}(2^j - 2)\left(4^{r-j} - \binom{r}{j}\right) = \sum_{j=2}^{r-1}(3^j - 2^{j+1} + 1)\left(3\binom{r}{j+1} - 2 \cdot 4^{r-j-1}\right). \qquad (2.15)$$

It can easily be calculated that both sums are equal to $\frac{4^r}{3} - 3^r + 2^r - \frac{1}{3}$, which completes the proof. □

3. Discussion

In this paper we stated and proved three characterizations theorems for exponential distribution. Although they have similar formulation they could be considered as different types. We suppose that these three types of characterizations could be valid for any order statistics. For example, the theorem 2.1 could be generalized that $\sum_{j=1}^{k} \frac{X_j}{n-j+1}$ and $X_{(k;n)}$ are identically distributed. Other two characterizations could also be generalized in a similar manner. So far, the generalization of the theorem 2.3 for the special case of $k = n$ (consecutive maxima) has been considered in [3]. We hope to extend the results in the future.

References

[1] M. Ahsanullah and G.G. Hamedani, *Exponential distribution: Theory and Methods* (NOVA Science, New York, 2010).

[2] B.C. Arnold and J.A. Villasenor, Exponential characterizations motivated by the structure of order statistics in samples of size two, *Stat Probab Lett* **83**(2) (2013), 596 – 601.

[3] S. Chakraborty and G.P. Yanev, Characterization of exponential distribution through equidistribution conditions for consecutive maxima, *J Stat Appl Pro* **2**(3) (2013), 237 – 242.

[4] J. Galambos and S. Kotz, *Characterizations of Probability Distributions*, (Springer-Verlag, Berlin-Heidelberg-New York, 1978).

[5] G.P. Yanev and S. Chakraborty, Characterizations of exponential distribution based on sample of size three, *Pliska Stud Math Bulgar* **23** (2013), 237 – 244.