# On a Conjecture Concerning a Characterization of the Exponential Distribution

Constance van Eeden
*Department of Statistics, University of British Columbia*
*2021 West Mall, Vancouver, B.C. V6T 1Z2*
and
*Département de mathématiques et d'informatique*
*Université du Québec à Montréal*
*C.P. 8888, Succ. A, Montréal, P.Q. H3C 3P8*

This paper contains the text of the author's *Gold Medalist Address* at the Annual Meeting of the Statistical Society of Canada, Toronto, 4-6 June, 1991. In it, the partial resolution by Leslie, van Eeden [14] of a conjecture of Dufour [7] concerning a characterization of the exponential distribution is presented. The importance of the characterization in a hypothesis testing problem is pointed out and several related characterizations are mentioned. The method of proof is outlined for the special case where the distribution function is absolutely continuous. Full details of the proofs are given in Leslie, van Eeden [14].

## 1 The conjecture

Dufour [7] states the following conjecture concerning a characterization of the exponential distribution:

Let $X_1, \ldots, X_n$ be independent, identically distributed non-negative random variables and let $Y_{1,n} \leq \ldots \leq Y_{n,n}$ be their order statistics. Let, for $i = 1, \ldots, n$,

$$D_{i,n} = (n - i + 1)(Y_{i,n} - Y_{i-1,n})$$

$$S_{i,n} = \sum_{j=1}^{i} D_{j,n},$$

where $Y_{0,n} = 0$. Further let, for some $r \in \{2, \ldots, n-1\}$,

$$W_{r,n} = \left( \frac{S_{1,n}}{S_{r,n}}, \ldots, \frac{S_{r-1,n}}{S_{r,n}} \right).$$

Then $X_1$ is exponentially distributed if $W_{r,n}$ is distributed as the order statistics of a sample of size $r-1$ from a $U(0,1)$ distribution.

That this conjecture characterizes the exponential distribution follows from the easily verifiable fact that $W_{r,n}$ has this uniform-order-statistics distribution if $X_1$ is exponentially distributed.

## 2 Use of the characterization in goodness-of-fit testing

Suppose, in the setting described in Section 1, one wants to test the hypothesis $H_0 : X_1$ has an exponential distribution with scale parameter $\theta > 0$, where $\theta$ is unknown and the available data are $Y_{1,n}, \ldots, Y_{r,n}$. Then, under $H_0, W_{r,n}$ has a distribution which is independent of $\theta$ and, as seen above, this distribution is the distribution of the order statistics of a sample of size $r-1$ from a $U(0,1)$ distribution.

So, if the conjecture is true, then the hypothesis $H_0$ is equivalent to the hypothesis

$$H_0^* : W_{r,n} \sim U_{(\cdot)}(r-1),$$

where $\sim$ stands for *is distributed as* and $U_{(\cdot)}(r-1)$ denotes the order statistics of a sample of size $r-1$ from a $U(0,1)$ distribution. The hypothesis $H_0^*$ can be tested by using the Kolmogorov statistic, or any other goodness-of-fit test for the uniform distribution (see e.g., d'Agostino, Stephens [6]).

However, if the conjecture is false then such a test has, as a test of $H_0$, power equal to its size for at least one alternative (i.e. non-exponential) distribution.

## 3 Some known characterizations of the exponential distribution

There are many characterizations of the exponential distribution and several books on characterizations of probability distributions contain sections on the special case of the exponential (see e.g. Kagan, Linnik and Rao [11], Galambos and Kotz [9]). The book by Azlarov and Volodin [4] specializes in characterizations of the exponential distribution, particularly those based on properties of the order statistics and those using the geometric distribution.

New characterizations of the exponential appear regularly in the literature. An example of a recent one is the following by Kopocinski [12] and Steutel and Thiemann [17]:

If $X$ is a non-negative random variable then $[cX]$ and $cX - [cX]$ are independent for all $c > 0$ if and only if $X$ is exponential.

However, given the nature of the characterization in the conjecture, I will restrict myself to a review of some of the known characterizations of the exponential distribution which are based on properties of order statistics and/or ratios of partial sums.

A) Fisz showed that, if $X_1$ and $X_2$ are i.i.d. and non-negative with $P(X_1 \leq x)$ strictly increasing on $(0, \infty)$, then $Y_{2,2} - Y_{1,2}$ and $Y_{1,2}$ are independent if and only if $X_1$ is exponential.

B) For the case where $X_1, \ldots, X_n$ are i.i.d. and non-negative, $X_1$ is exponential if and only if

   *i)* $D_{1,n}$ $(= nY_{1,n}) \sim X_1$. This can easily be seen by direct computation.

   *ii)* $D_{i,n} \sim X_1$ for some $i \in \{2, \ldots, n\}$. This result holds when $X_1$ has an absolutely continuous distribution function which is strictly increasing on $(0, \infty)$ and its failure rate is monotone. It is one of several characterizations by Ahsanullah [1,2,3].

*iii)* $Y_{j,n} - Y_{i,n} \sim Y_{j-i,n-i}$ for some $i$ and some pair $(j_1, j_2)$ of $j$,s with $1 \leq i < j_1 < j_2 \leq n$. This holds when $X_1$ has a continuous distribution function which is strictly increasing on $(0, \infty)$ (Gather [10]). Ahsanullah proved this characterization under the additional assumption that $X_1$ has a density and monotone failure rate.

*iv)* for $n \geq 3$,

$$\left( \frac{X_1}{\sum_{i=1}^{n} X_i}, \frac{X_1 + X_2}{\sum_{i=1}^{n} X_i}, \ldots, \frac{X_1 + \ldots + X_{n-1}}{\sum_{i=1}^{n} X_i} \right) \sim U_{(\cdot)}(n-1).$$

This characterization can be found in Seshadri, Csörgö and Stephens [16], in Csörgö, Seshadri and Yalovski [5] and in Menon, Seshadri [15]. For $n = 2$ the uniform distribution of $X_1/(X_1 + X_2)$ does not characterize the exponential distribution. This was shown by Kotlarski [13]. He notes that when $X_1$ and $X_2$ are i.i.d. with density $(1 + t^2)^{-3/2}$ on $(0, \infty)$, then $X_1/(X_1 + X_2)$ is $U(0, 1)$. Also, when $1/X_1$ is exponential, $X_1/(X_1 + X_2)$ is $U(0, 1)$.

C) If, in the conjecture, one takes $r = n$ it becomes

$$\left( \frac{S_{1,n}}{S_{n,n}}, \ldots, \frac{S_{n-1,n}}{S_{n,n}} \right) \sim U_{(\cdot)}(n-1) \Rightarrow X_1 \text{ is exponential.}$$

This characterization can be found in Seshadri, Csörgö and Stephens [16] and in Dufour, Maag and van Eeden [8]. It holds for $n \geq 3$. For $n = 2$ this characterization is not true: $(X_1)^{-1}$ exponential also gives a uniform distribution for $S_{1,2}/S_{2,2}$.

D) Finally we mention a characterization theorem of Kotlarski [13]. This result is needed in the proof of the partial resolution of the conjecture. Kotlarski shows that, if $T_1, T_2$ and $T_3$ are independent then $(R_1 = T_1/T_3, R_2 = T_2/T_3)$ has density $2(r_1 + r_2 + 1)^{-3}, r_1 \geq 0, r_2 \geq 0$, if and only if $T_1, T_2$ and $T_3$ are i.i.d. exponential. Note that the characterization B*iv)* above is a generalization of the one of Kotlarski.

None of the above characterizations can directly be used to resolve the conjecture. In fact, B *i)*-*iii)* are based on spacings, but not on ratios of partial sums of spacings, as is the conjecture. Further, B*iv)* is based on ratios of partial sums, but these are partial sums of i.i.d. random variables, whereas the spacings are not necessarily i.i.d. Also, the proof of C) above does, as far as I can see, not generalize to the case where $r \leq n$.

## 4 Some restatements of the conjecture

In this section several equivalent forms of the conjecture will be given.

Dufour [7] shows that

$$W_{r,n} \sim U_{(\cdot)}(r-1) \Leftrightarrow (V_{1,n}, \ldots, V_{r-1,n}) \text{ has density}$$

$$\frac{n!(r-1)!}{(n-r)!}\left(\sum_{i=1}^{r-1} v_i + n - r + 1\right)^{-r}, 0 \le v_1 \le \ldots \le v_{r-1} \le 1,$$

where

$$V_{i,n} = \frac{Y_{i,n}}{Y_{r,n}}, \; i = 1, \ldots, r-1.$$

This result is obtained by the following straightforward, but tedious to implement, transformation

$$\left(\frac{S_{1,n}}{S_{r,n}}, \ldots, \frac{S_{r-1,n}}{S_{r,n}}\right) \Rightarrow \left(\frac{Y_{1,n}}{Y_{r,n}}, \ldots, \frac{Y_{r-1,n}}{Y_{r,n}}\right).$$

So now the conjecture can be written in the form

$(V_{1,n}, \ldots, V_{r-1,n})$ has the above density $\Rightarrow$ $X_1$ is exponential.

Now assume a density, $f$, for $X_1$. Then (see Dufour [7]), the density of $(V_{1,n}, \ldots, V_{r-1,n})$ becomes

$$\frac{n!}{(n-r)!}\int_0^\infty x^{r-1}\prod_{i=1}^{r-1}\{f(v_i x)\}f(x)(1-F(X))^{n-r}dx \;\; 0 \le v_1 \le \ldots \le v_{r-1} \le 1$$

and the conjecture becomes

$$\int_0^\infty \prod_{i=1}^{r-1}\{xf(v_i x)\}(n-r+1)f(x)(1-F(x))^{n-r}dx =$$

$$(n-r+1)(r-1)!\left(\sum_{i=1}^{r-1} v_i + n - r + 1\right)^{-r}, 0 \le v_i \le 1, i = 1, \ldots, r-1$$

implies that $X_1$ is exponential.

So now the conjecture can also be stated as $F(x) = 1 - e^{-\theta x}$, $x \ge 0$ and some $\theta > 0$ is the only solution to the above integral equation.

Another form of the conjecture can be obtained by noting that

$$(n-r+1)f(x)(1-F(x))^{n-r}, \; x \ge 0$$

is the density of $\min(X_1, \ldots, X_{n-r+1})$. So, the conjecture can also be stated as

if, on $[0,1]^{r-1}$, the density of $\left(\frac{X_1}{Z}, \ldots, \frac{X_{r-1}}{Z}\right)$ is given by

$$(n-r+1)(r-1)!\left(\sum_{i=1}^{r-1} v_i + n - r + 1\right)^{-r}$$

then $X_1$ is exponential, where $Z \sim \min(X_1, \ldots, X_{n-r+1})$ and is independent of $X_1, \ldots, X_{r-1}$. Note that the random variables $X_i/Z, i = 1, \ldots, r-1$, take non-negative values, but that knowing that $W_{r,n} \sim U_{(\cdot)}(r-1)$ gives us the joint density of the $X_i/Z$ only on the cube $[0,1]^{r-1}$.

In the next section it will be shown how the results of this section can be used to partially resolve the conjecture.

## 5 A PARTIAL RESOLUTION TO THE CONJECTURE

In Section 4 we saw that (assuming a density for $X_1$)

$$W_{r,n} \sim U_{(\cdot)}(r-1)$$

on $[0,1]^{r-1}$, determines the density of

$$\left(\frac{X_1}{Z}, \dots, \frac{X_{r-1}}{Z}\right)$$

where $Z \sim \min(X_1, \dots, X_{n-r+1})$ and is independent of $X_1, \dots, X_{r-1}$. Further, this density was obtained in explicit form. Now use (for a proof see Leslie, van Eeden [14]) the fact that

$$W_{r,n} \sim U_{(\cdot)}(r-1) \Rightarrow W_{r-k,n-k} \sim U_{(\cdot)}(r-k-1), 0 \leq k \leq r-2 \leq n-2.$$

Then

$$W_{r,n} \sim U_{(\cdot)}(r-1) \Rightarrow$$

the density of $\left(\frac{X_1}{Z}, \dots, \frac{X_{r-k-1}}{Z}\right)$ on $[0,1]^{r-k-1}$ is given by

$$(n-r+1)(r-k-1)! \left(\sum_{i=1}^{r-k-1} v_i + n - k + 1\right)^{-(r-k)}, 0 \leq k \leq r-2 \leq n-2,$$

where $Z \sim \min(X_1, \dots, X_{n-r+1})$ and independent of $X_1, \dots, X_{r-k-1}$. How this last result can be used to obtain the partial resolution of the conjecture is best explained for a small value of $r$. So, let us look at the case where $r = 3$. There we have:

$$W_{3,n} \sim U_{(\cdot)}(2)$$

determines the density of

$$\left(\frac{X_1}{Z}, \frac{X_2}{Z}\right) \text{ on } [0,1]^2$$

and the density of

$$\frac{X_1}{Z} \text{ on } [0,1].$$

From this, one can obtain

$$P\left(\frac{X_1}{Z} \leq v_1, \frac{X_2}{Z} > v_2\right) = P\left(\frac{X_1}{Z} \leq v_1\right) - P\left(\frac{X_1}{Z} \leq v_1, \frac{X_2}{Z} \leq v_2\right),$$

$$0 \leq v_i \leq 1, i = 1, 2$$

and, in the same way,

$$P\left(\frac{X_1}{Z} > v_1, \frac{X_2}{Z} > v_2\right), \quad 0 \leq v_i \leq 1, i = 1, 2.$$

Thus one can obtain

$$P(\min(\frac{X_1}{Z}, \frac{X_2}{Z}) > v), \quad 0 \le v \le 1.$$

In the general case the following result holds (see Leslie, van Eeden [14])

$$W_{r,n} \sim U_{(\cdot)}(r-1) \Rightarrow$$

$$P\left(\frac{\min(X_i, 1 \le i \le l)}{Z} > s_1, \frac{\min(X_i, l+1 \le i \le j-1)}{Z} > s_2\right) =$$

$$\left(1 + \tfrac{1}{n-r+1}s_1 + \tfrac{j-l-1}{n-r+1}s_2\right)^{-1}, \quad \begin{array}{l} 0 \le s_i \le 1, i = 1,2, \\ 2 \le j \le r \le n, 0 \le l \le j-2, \end{array}$$

where $Z \sim \min(X_1, \ldots, X_{n-r+1})$ and independent of $X_1, \ldots, X_{j-1}$. So, taking $l = j - l - 1 = n - r + 1$ (which is possible if and only if $r - 1 \ge 2n/3$), one obtains

$$W_{r,n} \sim U_{(\cdot)}(r-1) \Rightarrow$$

$$\text{on } [0,1]^2 \ (\frac{Z_1}{Z}, \frac{Z_2}{Z}) \text{ has density } 2(1 + s_1 + s_2)^{-3},$$

where $Z, Z_1, Z_2$ are i.i.d. and distributed as $\min(X_1, \ldots, X_{n-r+1})$. Now note that the above density of $(Z_1/Z, Z_2/Z)$ is the one in Kotlarski's theorem. Still, the theorem of Kotlarski cannot be applied because we know the density of $(Z_1/Z, Z_2/Z)$ only on $[0,1]^2$. However, using the transformation $V_1 = Z/Z_1, V_2 = Z_2/Z_1$ one gets

$$W_{r,n} \sim U_{(\cdot)}(r-1) \Rightarrow (V_1, V_2) \text{ has, on } \{v_1 > 0, 0 < v_2 < v_1\},$$

$$\text{density } 2(1 + v_1 + v_2)^{-3}.$$

From the fact that $(V_1, V_2)$ is exchangeable and has the same distribution as $(Z_1/Z, Z_2/Z)$ it then follows that

$$W_{r,n} \sim U_{(\cdot)}(r-1) \Rightarrow$$

$$\left(\frac{Z_1}{Z}, \frac{Z_2}{Z}\right) \text{ has density } 2(1 + s_1 + s_2)^{-3}, s_1, s_2 > 0.$$

Now the theorem of Kotlarski can be applied and this gives

$$W_{r,n} \sim U_{(\cdot)}(r-1) \Rightarrow Z \text{ is exponential},$$

which implies that $X_1$ is exponential.

Thus, the conjecture is true for the case where $X_1$ has a density and $r \ge 2n/3 + 1$.

The first of these two conditions is not necessary. The proof of Leslie, van Eeden [14], which is in principle the one outlined above, works without the condition of a density for $X_1$. Whether the conjecture is true when $n \ge 3$ and $r < 2n/3 + 1$ we do not know.

REFERENCES


1. M. AHSANULLAH (1975). A characterization of exponential distribution. *Statistical Distributions in Scientific Work*, G.P. Patil, S. Kotz, J.K. Ord, Eds, Vol. 3, 131-135.
2. M. AHSANULLAH (1977). A characteristic property of the exponential distribution. *Ann. Statist. 5*, 580-582.
3. M. AHSANULLAH (1978). On a characterization of the exponential distribution by spacings. *Ann. Inst. Statist. Math. 30*, 163-166.
4. T.A. AZLAROV, N.A. VOLODIN (1986). *Characterization Problems Associated with the Exponential Distribution*, Springer-Verlag.
5. M. CSÖRGÖ, V. SESHADRI, M. YALOVSKY (1975). Applications of characterizations in the area of goodness of fit. *Statistical Distributions in Scientific Work*, G.P. Patil, S. Kotz, J.K. Ord, Eds, Vol. 2, 70-90.
6. R.B. D'AGOSTINO, M.A. STEPHENS (1986). *Goodness-of-fit Techniques*, Marcel Dekker.
7. R. DUFOUR (1982). *Tests d'ajustement pour des échantillons tronqués ou censurés*, PhD thesis, Université de Montréal.
8. R. DUFOUR, U.R. MAAG, C. VAN EEDEN (1984). Correcting a proof of a characterization of the exponential distribution, *J.R. Statist. Soc. B, 46*, 238-241.
9. J. GALAMBOS, S. KOTZ (1978). *Characterizations of Probability Distributions*, Lecture Notes in Mathematics 675, Springer-Verlag.
10. U. GATHER (1989). On a characterization of the exponential distribution by properties of the order statistics. *Statistics and Probability Letters 7*, 93-96.
11. A.M. KAGAN, YU.V. LINNIK, C.R. RAO (1973). *Characterization Problems in Mathematical Statistics*, John Wiley.
12. B. KOPOCINSKI (1988). Some characterizations of the exponential distribution function. *Prob. and Math. Statist. 9*, 105-111.
13. I. KOTLARSKI (1967). On characterizing the gamma and the normal distribution. *Pacif. J. Math. 20*, 69-76.
14. J. LESLIE, C. VAN EEDEN (1991). On a characterization of the exponential distribution based on a type 2 right censored sample (submitted).
15. M.V. MENON, V. SESHADRI (1975). A characterization theorem useful in hypothesis testing. *Contributions libres*, $40^{th}$ Session of the International Statistical Institute, 586-590.
16. V. SESHADRI, M. CSÖRGÖ, M.A. STEPHENS (1969). Tests for the exponential distribution using Kolmogorov-type statistics. *J.R. Statist. Soc. B, 31*, 499-509.
17. F.W. STEUTEL, J.G.F. THIEMANN (1989). On the independence of integer and fractional parts. *Statistica Neerlandica 43*, 53-59.