REVSTAT – Statistical Journal
Volume 0, Number 0, Month 0000, 000-000
https://doi.org/00.00000/revstat.v00i0.000

On a characterization of exponential and double exponential distributions

Authors: Reza Rastegar
– Occidental Petroleum Corporation,
Houston, TX 77046, USA
reza_rastegar2@oxy.com

Alexander Roitershtein ✉
– Department of Mathematics, Texas A&M University,
College Station, TX 77843, USA,
roiterst@tamu.edu

Received: April 2022 Revised: December 2022 Accepted: Month 0000

Abstract:

- Recently, G. Yanev [7] obtained a characterization of the exponential family of distributions in terms of a functional equation for certain mixture densities. The purpose of this note is twofold: we extend Yanev's theorem by relaxing a restriction on the sign of mixture coefficients and, in addition, obtain a similar characterization for the Laplace family of distributions.

Keywords:

- *Hypoexponential distribution; Laplace distribution; characterization of distributions; sums of independent random variables; characteristic functions, functional equations.*

AMS Subject Classification:

- 62E10, 60G50, 60E10.

✉ Corresponding author

1. Introduction

Our aim is to prove certain characterizations of the exponential and double exponential families of distributions. We will use the notation $X \sim \mathcal{E}(\lambda)$ to indicate that X is an exponential random variable with parameter $\lambda > 0$, that is $P(X > x) = e^{-\lambda x}$ for all $x > 0$. We will write $X \in \mathcal{E}$ if $X \sim \mathcal{E}(\lambda)$ for some $\lambda > 0$. Similarly, will write $X \in \mathcal{L}$ if X has a Laplace (double exponential) distribution [4], that is, for some $\lambda > 0$ and $Y \sim \mathcal{E}(\lambda)$,

$$P(X > x) = \frac{1}{2}\Big(P(Y > x) + P(-Y > x)\Big) = \frac{\lambda}{2}\int_x^\infty e^{-\lambda|y|}\,dy, \qquad \forall\, x \in \mathbb{R}.$$

For the exponential random variable we have:

Theorem 1.1. *Let X be a random variable and $\mu_1, \ldots, \mu_n$ be distinct non-zero real numbers. Let $\varphi(t) = E\big(e^{itX}\big)$, $t \in \mathbb{R}$, be the characteristic function of X, and suppose that φ is infinitely differentiable at zero and, furthermore,*

$$\prod_{k=1}^{n} \varphi(\mu_k t) = \sum_{k=1}^{n} \theta_k \varphi(\mu_k t), \qquad t \in \mathbb{R}, \tag{1.1}$$

where

$$\theta_k = \prod_{j=1, j\neq k}^{n} \frac{\mu_k}{\mu_k - \mu_j}, \qquad k = 1, \ldots, n. \tag{1.2}$$

If, in addition,

$$\sum_{(k_1,\ldots,k_n)\in W_{n,m}} \prod_{j=1}^{n} \mu_j^{k_j} \neq \sum_{k=1}^{n} \mu_k^m \quad \text{for any integer } m \geq 2, \tag{1.3}$$

where

$$W_{n,m} := \Big\{(k_1, \ldots, k_n) \in \mathbb{Z}^n : k_j \geq 0 \text{ and } \sum_{j=1}^{n} k_j = m\Big\}, \tag{1.4}$$

then, either $P(X = 0) = 1$ or $E(X) \neq 0$ and $X \cdot \operatorname{sign}\big(E(X)\big) \sim \mathcal{E}(\lambda)$ with $\lambda = 1/E(X)$.

The proof of the theorem is given in Section 2. Theorem 1.1 is an extension of a similar result of G. Yanev [7] obtained under the additional assumption that the coefficients μ_k are positive. In that case, the key technical condition (1.3) is trivial as the left-hand sides contains the μ_k^m terms and hence is always larger than the right-hand side.

To ensure the existence of the derivatives of φ at zero one can impose Cramér's condition, namely assume that there is $t_0 > 0$ such that $E\big(e^{tX}\big) < \infty$ for all $t \in (-t_0, t_0)$. Note also that the equality in (1.3) for any fixed $m \in \mathbb{N}$ describes a low-dimensional manifold in $\mathbb{R}^n$, and hence Theorem 1.1 is true for almost every vector $(\mu_1, \ldots, \mu_n)$ chosen at random from a continuous distribution on $\mathbb{R}^n$.

The identity in (1.1) with θ_k introduced in (1.2) holds for any $X \in \mathcal{E}$, and Theorem 1.1 can be seen as a converse to this result. Let $X_1, \ldots, X_n$, $n \geq 2$, be independent copies of a random variable X. Equations (1.1) and (1.2) give an expression of the characteristic function of the sum

$$S = \mu_1 X_1 + \cdots + \mu_n X_n \tag{1.5}$$

as a linear combination of $\varphi(\mu_k t)$'s. If $X \sim \mathcal{E}(\lambda)$, then $\varphi(t) = \frac{\lambda}{\lambda - it}$, and thus (1.1) is the partial fraction decomposition of the complex-valued rational function $\psi(t) := E\big(e^{itS}\big)$. In the particular case when $X \in \mathcal{E}$ and $\mu_k = \frac{1}{L-k+1}$ for some integer $L > n$, the random variable S/λ is distributed as the n-th order statistic of a sample of L independent copies of X (this is the Rényi representation of order statistics; see, for instance, [3, p. 18]). For further background and earlier versions (particular cases) of Yanev's characterization theorem see [1, 6, 7].

It was pointed out in [7] that an extension of their result to a more general class of coefficients μ_k would be of interest from the viewpoint of both theory and applications[1]. When all the coefficients μ_k are positive and X is an exponential random variable, the random variable $S = \sum_{k=1}^{n} \mu_k X_k$ has a hypoexponential distribution. When some of the coefficients are negative, S is a difference of two hypoexponential random variables. Some applications of such differences are discussed, for instance, in [5]. An insightful theoretical exploration of the densities of hypoexponential distributions can be found in [2].

We remark that the theorem is not true if the particular form of the coefficients θ_k in (1.2) is not enforced. For instance, for the Laplace distribution we have:

Theorem 1.2. *Let X be a random variable and $\mu_1, \ldots, \mu_n$ be distinct positive numbers. Let $\varphi(t) = E\big(e^{itX}\big)$, $t \in \mathbb{R}$, be the characteristic function of X, and suppose that φ is infinitely differentiable at zero and, furthermore, (1.1) holds with*

$$\theta_k = \prod_{j=1, j\neq k}^{n} \frac{\mu_k^2}{\mu_k^2 - \mu_j^2}, \qquad k = 1, \ldots, n. \tag{1.6}$$

Then, either $P(X = 0) = 1$ or X has a Laplace distribution.

The result is closely related to the one stated in Theorem 1.1 because $X \in \mathcal{L}$ implies that for a suitable $Y \in \mathcal{E}$,

$$E\big(e^{itX}\big) = \frac{1}{2}\Big(E\big(e^{itY}\big) + E\big(e^{-itY}\big)\Big).$$

The proof of the theorem is similar to that of Theorem 1.1, and therefore is omitted. The key technical ingredient of the proof, namely an analogue of Lemma 2.1 for Laplace distributions, follows immediately from Lemma 2-(iii) in [7], and the rest of the proof of Theorem 1.1 can be carried over verbatim to the double exponential setup of Theorem 1.2.

We conclude the introduction with a brief discussion of condition (1.3). The equality with $n = 2$ and some $m \geq 2$ reads $\sum_{j=0}^{m} \mu_1^j \mu_2^{m-j} = \mu_1^m + \mu_2^m$, which is equivalent to $\frac{\mu_1^{m+1} - \mu_2^{m+1}}{\mu_1 - \mu_2} = \mu_1^m + \mu_2^m$. The latter implies that $\mu_2^{m-1} = \mu_1^{m-1}$, and hence m is odd and $\mu_2 = -\mu_1$. In that case, (1.1) becomes

$$\varphi(t)\varphi(-t) = \frac{1}{2}\big(\varphi(t) + \varphi(-t)\big), \qquad t \in \mathbb{R}. \tag{1.7}$$

The equation is satisfied when X is a Bernoulli random variable with $P(X = 0) = P(X = a) = \frac{1}{2}$ for some constant $a > 0$, in which case $\varphi(t) = \frac{1}{2}\big(1 + e^{iat}\big)$. More generally, (1.7) holds if and only if $\varphi(t) = \frac{1}{2}\big(1 + e^{i\rho(t)}\big)$, where $\rho : \mathbb{R} \to \mathbb{R}$ is an odd function. Unfortunately, we are not aware of any example where φ in this form would be a characteristic function of a random variable beyond the linear case $\rho(t) = at$ and linear fractional $\rho(t) = \frac{\lambda - ti}{\lambda + ti}$, $\rho(t) = \frac{\lambda + ti}{\lambda - ti}$ which correspond to, respectively, $X \in \mathcal{E}(\lambda)$ and $-X \in \mathcal{E}(\lambda)$.

Our proof technique differs significantly from the one used in [7]. However, interestingly enough, both rely on the validity of (1.3). Nevertheless, we believe that the following might be true:

Conjecture. For $n \geq 3$, (1.3) is an artifact of the proof and is not necessary.

2. Proof of Theorem 1.1

The following is a suitable version of Lemma 2-(iii) in [7].

[1] See also a recent preprint [8], where a special case of exponential convolutions with repeated coefficients is considered.

Lemma 2.1. *Assume (1.3). Then, for any integer $m \geq 2$,*

$$\sum_{k=1}^{n} \theta_k \mu_k^m \neq \sum_{k=1}^{n} \mu_k^m.$$

Proof of Lemma 2.1: Let $Y_1, \ldots, Y_n$ be i. i. d. random variables such that $Y_k \in \mathcal{E}(1)$ for each $k = 1, \ldots, n$. Similarly to (1.5), define

$$\widetilde{S} = \mu_1 Y_1 + \cdots + \mu_n Y_n.$$

Let $\widetilde{\varphi}(t) = E\big(e^{itY_1}\big) = \frac{1}{1-it}$. Thus, for $E\big(e^{i\widetilde{S}t}\big) = \prod_{k=1}^n \frac{1}{1-it\mu_k}$ we have a partial fraction decomposition similar to (1.1), namely

$$E\big(e^{i\widetilde{S}t}\big) = \sum_{k=1}^{n} \frac{\theta_k}{1-it\mu_k} = \sum_{k=1}^{n} \theta_k \widetilde{\varphi}(\mu_k t), \qquad t \in \mathbb{R},$$

where the coefficients θ_k are introduced in (1.2). Differentiating m times and taking in account that $E(Y_1^m) = m!$, we obtain

$$(-i)^m \frac{d^m}{dt^m} E\big(e^{i\widetilde{S}t}\big)\Big|_{t=0} = E\big(\widetilde{S}^m\big) = \sum_{k=1}^{n} \theta_k \mu_k^m E\big(Y_1^m\big) = m! \sum_{k=1}^{n} \theta_k \mu_k^m,$$

Recall $W_{n,m}$ from (1.4). Using the multinomial expansion

$$\widetilde{S}^m = \big(\mu_1 Y_1 + \cdots + \mu_n Y_n\big)^m = \sum_{(k_1,\ldots,k_n) \in W_{n,m}} \frac{m!}{k_1! \cdots k_n!} \prod_{j=1}^{n} (\mu_j Y_j)^{k_j},$$

and the fact that $E(Y_1^k) = k!$ for any $k \in \mathbb{N}$, we obtain:

$$\begin{aligned} \sum_{k=1}^{n} \theta_k \mu_k^m &= \frac{1}{m!} E\big(\widetilde{S}^m\big) = \sum_{(k_1,\ldots,k_n) \in W_{n,m}} \frac{1}{k_1! \cdots k_n!} \prod_{j=1}^{n} \mu_j^{k_j} E\big(Y_j^{k_j}\big) \\ &= \sum_{(k_1,\ldots,k_n) \in W_{n,m}} \prod_{j=1}^{n} \mu_j^{k_j}, \end{aligned} \tag{2.1}$$

which yields the result in view of (1.3). □

Differentiating both sides of (1.1) m times we obtain the identity

$$\frac{d^m}{dt^m} \prod_{k=1}^{n} \varphi(\mu_k t)\Big|_{t=0} = \sum_{k=1}^{n} \theta_k \mu_k^m \varphi^{(m)}(0), \qquad m \geq 2. \tag{2.2}$$

In view of Lemma 2.1 and the fact that $\varphi(0) = 1$, these identities can be used to determine all the derivatives of φ at zero in terms of $\varphi'(0)$, first $\varphi''(0)$ in terms of the parameter $\varphi'(0)$, then $\varphi'''(0)$ in terms of $\varphi'(0)$ and $\varphi''(0)$, and hence in terms of $\varphi'(0)$ only, and so on. For instance, (2.2) with $m = 2$ yields

$$\varphi''(0) \sum_{k=1}^{n} \mu_k^2 (\theta_k - 1) = \big(\varphi'(0)\big)^2 \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \mu_k \mu_j.$$

Note that $\sum_{k=1}^n \mu_k^2(\theta_k - 1) \neq 0$ by Lemma 2.1 with $m = 2$.

In general, for an arbitrary $m \in \mathbb{N}$, (2.2) can be written as

$$\varphi^{(m)}(0) \sum_{k=1}^{n} \mu_k^m (\theta_k - 1) = F_m\big(\mu_1, \ldots, \mu_n, \varphi'(0), \ldots, \varphi^{(m-1)}(0)\big),$$

where the multivariate functional $F_m(\cdot)$ in the right-hand side is independent of $\varphi^{(m)}(0)$. An explicit form of F_m is given by the general Leibnitz rule (an extension of the product differentiation rule to higher derivatives):

$$F_m = \sum_{(k_1,\dots,k_n)\in V_{n,m}} \frac{m!}{k_1!\cdots k_n!} \prod_{j=1}^{n} \mu_j^{k_j} \varphi^{(k_j)}(0),$$

where (cf. (1.4))

$$V_{n,m} := \big\{(k_1,\dots,k_n) \in \mathbb{Z}^n : 0 \le k_j < m \text{ and } \sum_{j=1}^{n} k_j = m\big\},$$

In view of Lemma 2.1, $\sum_{k=1}^n \mu_k^m(\theta_k - 1) \neq 0$, and thus we can define the derivative of φ at zero inductively, using the formula

$$\varphi^{(m)}(0) = \frac{F_m\big(\mu_1,\dots,\mu_n,\varphi'(0),\dots,\varphi^{(m-1)}(0)\big)}{\sum_{k=1}^n \mu_k^m(\theta_k - 1)}.$$

Let now $Z \in \mathcal{E}$ and $\psi(t) = E\big(e^{itZ}\big)$. The derivatives $\psi''(0), \psi'''(0), \dots$ as functions of the parameter $\psi'(0)$ can be in principle derived using the same inductive algorithm. Therefore, $\varphi'(0) = 0$ implies $P(X = 0) = 1$ while $\varphi'(0) = \psi'(0) = \lambda^{-1}$ for some $\lambda > 0$ implies that $\varphi^{(m)}(0) = \psi^{(m)}(0)$ for all $m \in \mathbb{N}$, and hence (since φ is analytic under the conditions of the theorem) $\varphi(t) = \psi(t) = \frac{\lambda}{\lambda - it}$ as desired. Finally, the case $\varphi(0) = -\lambda^{-1} < 0$ can be reduced to the previous one by switching from X to $-X$ in the above argument. □

REFERENCES

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

[2] Belton, A.; Guillot, D.; Khare, A. and Putinar, M. (2022). *Hirschman-Widder densities*, Appl. Comput. Harmon. Anal., **60**, 396–425.

[3] David, H. A. and Nagaraja, H. N. (2004). *Order Statistics*, Wiley.

[4] Kotz, S.; Kozubowski, T. J. and Podgórski, K. (2001). *The Laplace Distribution and Generalizations. A Revisit with Applications to Communications, Economics, Engineering, and Finance*, Birkhäuser.

[5] Li, K. H. and Li, C. T. (2019). *Linear combination of independent exponential random variables*, Methodol. Comput. Appl. Probab., **21**, 253–277.

[6] Miloševič, B. and Obradović, M. (2016). *Some characterizations of the exponential distribution based on order statistics*, Appl. Anal. Discrete Math., **10**, 394–407.

[7] Yanev, G. P. (2020). *Exponential and hypoexponential distributions: some characterizations*, Mathematics, **8**, 2207.

[8] Yanev, G. P. (2022). *On characterization of the exponential distribution via hypoexponential distributions*, arXiv: 2204.00867.