Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v16n3p694

On the Distribution of the Sum of Independent Exponential-Geometric Random Variables
By AL-Zaydi

15 December 2023

Electronic Journal of Applied Statistical Analysis
Vol. 16, Issue 3, 2023, 694-710
DOI: 10.1285/i20705948v16n3p694

On the Distribution of the Sum of Independent Exponential-Geometric Random Variables

Areej M. AL-Zaydi*[a]

[a] *Department of Mathematics and Statistics, Faculty of Science, Taif University , P.O.Box 11099, Taif 21944, Saudi Arabia.*

15 December 2023

In this article, we derive exact expressions for the probability density function and cumulative distribution function of the sum of independent and non-identical exponential-geometric random variables. Then we discuss the corresponding result for independent and identically distributed exponential-geometric random variables. A saddlepoint approximation is also utilized to approximate the derived distribution. Finally, numerical simulations are used to investigate the precision of the saddlepoint approximation.

Keywords: Exponential-geometric distribution, Divided differences, Independent and non-identically distributed random variables, Saddlepoint approximation.

1 Introduction

The exponential-geometric (EG) distribution is a two-parameter distribution with a decreasing failure rate, was proposed by Adamidis and Loukas (1998) to model life-time data. This distribution is used in the latent competing risk scenario (Louzada-Neto, 1999), where the lifetime associated with a specific risk is not observable and only the minimum lifetime value among all risks is observed. The probability density function (PDF) and the cumulative distribution function (CDF) of the EG distribution are given as

$$f(x) = \frac{\beta(1-p)e^{-\beta x}}{(1-p\ e^{-\beta x})^2},\ x>0,\ \beta>0,\ 0<p<1, \tag{1}$$

*Corresponding author: areej.f@tu.edu.sa.

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

$$F(x) = \frac{1 - e^{-\beta x}}{1 - p\ e^{-\beta x}},\ x > 0,\ \beta > 0,\ 0 < p < 1, \tag{2}$$

where β is the scale parameter and p is the shape parameter. When p approaches zero, this distribution tends to an exponential distribution with the parameter β.
Using the series expansion

$$(1 - z)^{-k} = \sum_{j=0}^{\infty} \frac{\Gamma(k + j)}{\Gamma(k)j!} z^j, \quad |z| < 1, \quad k > 0,$$

we can rewrite (1) as

$$f(x) = (1 - p) \sum_{j=0}^{\infty} p^j \beta(j + 1) e^{-\beta(j+1)x}. \tag{3}$$

According to (3), the EG distribution is an infinite mixture of the exponential distribution. The moment generating function (MGF) of the EG distribution can be derived by integration as

$$M_X(t) = (1 - p)B(1 - \frac{t}{\beta}, 1)\, {}_2F_1(2,\ 1 - \frac{t}{\beta};\ 2 - \frac{t}{\beta};\ p), \tag{4}$$

where B(. , .) is the beta function. We know that ${}_2F_1(a,\ b;\ c;\ x)$ is the hypergeometric function defined by

$${}_2F_1(a,\ b;\ c;\ x) = \sum_{j=0}^{\infty} \frac{(a)_j (b)_j}{(c)_j} \frac{x^j}{j!},$$

where $(v)_j = v(v + 1) \cdots (v + j - 1)$ denotes the ascending factorial (Gradshteyn and Ryzhik, 2014).
Therefore, we can write Equation (4) as

$$M_X(t) = (1 - p) \sum_{j=0}^{\infty} p^j \left(\frac{\beta(j + 1)}{\beta(j + 1) - t} \right). \tag{5}$$

For more details on the EG distribution, see Adamidis and Loukas (1998) and Balakrishnan et al. (2015).

Sums of random variables appear naturally in many applied fields, including communications and computer science, insurance and reliability, and performance evaluation, to name a few. Numerous researchers have derived and approximated the distribution of the sum of random variables for different distributions. For example, Mathai (1982) obtained the distribution of the sum of independent and non-identically distributed gamma random variables. Moschopoulos (1985) provided the expression of a single gamma series whose coefficients are computed using simple recursive relations. Additionally, Van Khuong and Kong (2006) obtained the PDF of the sum of independent exponential random variables using the characteristic function. Sadooghi-Alvandi et al. (2009) obtained the distribution of the sum of independent and non-identically distributed uniform random variables using a relatively simple approach. Recently, Kitani and Murakami (2020) obtained the exact distribution of the sum of independent and non-identically distributed extended exponential random variables. Levy (2021, 2022) used a divided difference perspective to find the density for the sums of independent Mittag-Leffler, exponential, Erlang, and gamma variates. Furthermore, Kitani et al. (2023) derived the distribution of the sum of independent and non-identically distributed generalized Lindley random variables. The PDFs for the sum of n independent random variables for Shanker, Akash, Ishita, Pranav, Rani, and Ram Awadh distributions were derived by Yaghoubi (2022) using the change-of-variables technique.

Saddlepoint approximation (SA) is an efficient method for approximating the distribution of a random variable if its cumulant generating function (CGF) is known. Several researchers have used the SA with great success. For example, Murakami (2014) and Nadarajah et al. (2015) considered the use of the SA for the sum of independent and non-identically distributed uniform and beta random variables, respectively. Murakami (2015) discussed the use of the SA for the sum of independent and non-identically distributed gamma random variables. More recently, Kitani and Murakami (2020) and Kitani et al. (2023) gave the approximation for the distribution of the sum of independent and non-identically distributed extended exponential and generalized Lindley random variables, respectively.

In this paper, we present the distribution of the sum of independent EG random variables with distinct parameters as well as identical parameters. We use a similar procedure to that used in Levy (2022) to determine this distribution. The distribution of the sum of independent exponential random variables is also discussed as a special case. In addition, we discuss the approximate distribution using the SA. Furthermore, we compare the findings of a SA against those of a normal approximation (NA) to determine the best method to use for the distribution function.

The remainder of this paper is organized as follows: In Section 2, we provide some preliminaries that are required in the subsequent sections. In Section 3, we derive the exact PDF and CDF of the sum of independent and non-identically distributed EG random variables. Section 4 provides the exact PDF and CDF of the sum of independent, identically distributed EG random variables. The MGF for the sum of independent EG random variables is obtained in Section 5. In Section 6, we discuss the SA and present numerical results. Finally, a conclusion is given in Section 7.

2 Preliminaries

2.1 Newton's divided differences

For a function $f(.)$ defined at distinct points $a_1, ..., a_n$, the $(n-1)$th-order divided difference is denoted by $f[a_1, ..., a_n]$ and is defined by the recurrence relation:

$$f[a_1, ..., a_n] = \frac{f[a_2, ..., a_{n-1}, a_n] - f[a_1, ..., a_{n-2}, a_{n-1}]}{a_n - a_1}, \tag{6}$$

with $f[a] = f(a)$.
If $a_1 = ... = a_n = a$, $f[a_1, ..., a_n]$ is interpreted by its confluent form:

$$f[a, ..., a] = \frac{1}{(n-1)!} f^{(n-1)}(a),$$

assuming the derivative exists.
The divided difference $f[a_1, ..., a_n]$ can be expressed using Lagrange polynomials as follows:

$$f[a_1, ..., a_n] = \sum_{j=1}^{n} \frac{f(a_j)}{\prod_{\kappa=1,\kappa\neq j}^{n}(a_j - a_\kappa)}. \tag{7}$$

Examining (7) demonstrates that $f[a_1, ..., a_n]$ is a symmetric function of its arguments, and thus the calculations are invariant to permutations in the order of its arguments; for example, $f[a_1, a_2, a_3] = f[a_2, a_3, a_1]$. For a good introduction to the divided difference, see, for instance, Atkinson (1991).

2.2 Lemmas

We shall use the following lemmas to prove some of the paper's findings.

Lemma 1. *For any $n > 1$ distinct points $a_1, ..., a_n$ we have*

$$\sum_{j=1}^{n} \frac{1}{\prod_{\kappa=1,\kappa\neq j}^{n}(a_j - a_\kappa)} \equiv 0. \tag{8}$$

Lemma 2. *For distinct values $a_1, ..., a_n$ and $t \in \Re$, the decomposition of the rational function,*

$$U(t) = \frac{1}{\prod_{\kappa=1}^{n}(a_\kappa - t)},$$

as a sum of partial fractions, gives

$$\frac{1}{\prod_{\kappa=1}^{n}(a_\kappa - t)} = \sum_{j=1}^{n} \frac{1}{(a_j - t)\prod_{\kappa=1,\kappa\neq j}^{n}(a_\kappa - a_j)}.$$

(See Levy (2022), Lemma 3.1, and Lemma 3.2.)

2.3 The distribution of the sum of independent exponential random variables

The sum of n independent exponential random variables with pairwise distinct parameters, $\beta_i, i = 1, ..., n$, respectively, has the hypo-exponential density, $f_n(s)$, given by

$$f_n(s) = \left(\prod_{i=1}^{n} \beta_i\right) \sum_{j=1}^{n} \frac{e^{-\beta_j s}}{\prod_{\kappa=1,\kappa\neq j}^{n}(\beta_\kappa - \beta_j)}, \quad s \geq 0, \tag{9}$$

see, for example Ross (2014).
Levy (2022) demonstrates that Equation (9) can be determined and expressed more succinctly using a divided difference interpretation, as shown below:

$$f_n(s) = \left(\prod_{i=1}^{n} \beta_i\right) e[-\beta_1, ..., -\beta_n], \quad s \geq 0, \tag{10}$$

where $e[-\beta_1, ..., -\beta_n]$ is the $(n-1)$th-order divided difference for the function $e(a) = e^{at}$ at points $-\beta_1, ..., -\beta_n$.
When n independent exponential random variables have identical parameters, their sum has the Erlang distribution with parameters (n, β); see Akkouchi (2008).

3 The distribution of the sum of independent EG variables with different parameters

In this section, we derive the exact distribution of the sum of n independent and non-identically distributed EG random variables.
Assume that X_1 and X_2 are independent EG random variables with parameters (p_i, β_i), $i = 1, 2$ and where $\beta_i \neq \beta_j$, $p_i \neq p_j$ when $i \neq j$. Then the density function for the sum $S_2 = X_1 + X_2$, $f_2(s)$, can be found as follows:

$$\begin{aligned}
f_2(s) &= \int_0^s f_{X_1}(x_1) f_{X_2}(s - x_1) dx_1 \\
&= \left(\prod_{i=1}^{2} 1 - p_i\right) \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} \left(\prod_{i=1}^{2} p_i^{j_i} \beta_i (j_i + 1)\right) \\
&\times \int_0^s e^{-\beta_1(j_1+1)x_1} e^{-\beta_2(j_2+1)(s-x_1)} dx_1
\end{aligned}$$

$$\begin{aligned}
&= \left(\prod_{i=1}^{2} 1 - p_i\right) \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} \left(\prod_{i=1}^{2} p_i^{j_i} \beta_i (j_i + 1)\right) \\
&\times e^{-\beta_2(j_2+1)s} \int_0^s e^{-(\beta_1(j_1+1) - \beta_2(j_2+1))x_1} dx_1
\end{aligned}$$

$$\begin{aligned}
&= \left(\prod_{i=1}^{2} 1 - p_i\right) \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} \left(\prod_{i=1}^{2} p_i^{j_i} \beta_i (j_i + 1)\right) \\
&\times \left(\frac{e^{-\beta_2(j_2+1)s} - e^{-\beta_1(j_1+1)s}}{\beta_1(j_1 + 1) - \beta_2(j_2 + 1)}\right),
\end{aligned}$$

by using (7), we get

$$\begin{aligned} f_2(s) &= \left(\prod_{i=1}^{2} 1-p_i\right)\sum_{j_1=0}^{\infty}\sum_{j_2=0}^{\infty}\left(\prod_{i=1}^{2} p_i^{j_i}\beta_i(j_i+1)\right) \\ &\quad \times e[-\beta_1(j_1+1),-\beta_2(j_2+1)], \end{aligned} \tag{11}$$

where $e(a)=e^{as}$ at points $-\beta_1(j_1+1),-\beta_2(j_2+1)$.

Proposition 1. *Let $X_1,\cdots,X_n$ be independent EG random variables with parameters (p_i,β_i), $i=1,\cdots,n$, and where $\beta_i\neq\beta_j$, $p_i\neq p_j$ when $i\neq j$. Then, the density function of $S_n=\sum_{i=1}^{n} X_i$ is given by*

$$\begin{aligned} f_n(s) &= \left(\prod_{i=1}^{n} 1-p_i\right)\sum_{j_1=0}^{\infty}\sum_{j_2=0}^{\infty}\cdots\sum_{j_n=0}^{\infty}\left(\prod_{i=1}^{n} p_i^{j_i}\beta_i(j_i+1)\right) \\ &\quad \times e[-\beta_1(j_1+1),\cdots,-\beta_n(j_n+1)], \qquad s>0. \end{aligned} \tag{12}$$

Proof. We prove it using mathematical induction. For $n=1$, Equation (12) is trivially true. Assuming that (12) at $n-1$ is true, we examine the density for $S_n=S_{n-1}+X_n$,

$$f_n(s) = \int_0^s f_{n-1}(s-x_n)f_{X_n}(x_n)dx_n. \tag{13}$$

Using Equation (3) and Equation (12) in Equation (13), we get

$$\begin{aligned} f_n(s) &= \left(\prod_{i=1}^{n} 1-p_i\right)\sum_{j_1=0}^{\infty}\sum_{j_2=0}^{\infty}\cdots\sum_{j_n=0}^{\infty}\left(\prod_{i=1}^{n} p_i^{j_i}\beta_i(j_i+1)\right) \\ &\quad \times \int_0^s e[-\beta_1(j_1+1),\cdots,-\beta_{n-1}(j_{n-1}+1)]e^{-\beta_n(j_n+1)x_n}dx_n. \end{aligned} \tag{14}$$

Now using (7), we obtain

$$\begin{aligned} f_n(s) &= \left(\prod_{i=1}^{n} 1-p_i\right)\sum_{j_1=0}^{\infty}\sum_{j_2=0}^{\infty}\cdots\sum_{j_n=0}^{\infty}\left(\prod_{i=1}^{n} p_i^{j_i}\beta_i(j_i+1)\right) \\ &\quad \times \int_0^s \left(\sum_{\nu=1}^{n-1}\frac{e^{-\beta_\nu(j_\nu+1)(s-x_n)}}{\prod_{\kappa=1,\kappa\neq\nu}^{n-1}\beta_\kappa(j_\kappa+1)-\beta_\nu(j_\nu+1)}\right)e^{-\beta_n(j_n+1)x_n}dx_n \\ &= \left(\prod_{i=1}^{n} 1-p_i\right)\sum_{j_1=0}^{\infty}\sum_{j_2=0}^{\infty}\cdots\sum_{j_n=0}^{\infty}\left(\prod_{i=1}^{n} p_i^{j_i}\beta_i(j_i+1)\right) \\ &\quad \times \sum_{\nu=1}^{n-1} e^{-\beta_\nu(j_\nu+1)s}\int_0^s \frac{e^{-(\beta_n(j_n+1)-\beta_\nu(j_\nu+1))x_n}}{\prod_{\kappa=1,\kappa\neq\nu}^{n-1}\beta_\kappa(j_\kappa+1)-\beta_\nu(j_\nu+1)}dx_n. \end{aligned} \tag{15}$$

By integration and simplifying the resulting expression, we find that

$$\begin{aligned} f_n(s) &= \left(\prod_{i=1}^{n} 1-p_i\right)\sum_{j_1=0}^{\infty}\sum_{j_2=0}^{\infty}\cdots\sum_{j_n=0}^{\infty}\left(\prod_{i=1}^{n} p_i^{j_i}\beta_i(j_i+1)\right) \\ &\quad\times\left\{\sum_{\nu=1}^{n-1}\frac{e^{-\beta_\nu(j_\nu+1)s}}{\prod_{\kappa=1,\kappa\neq\nu}^{n}\beta_\kappa(j_\kappa+1)-\beta_\nu(j_\nu+1)}\right. \\ &\quad\left.-\sum_{\nu=1}^{n-1}\frac{e^{-\beta_n(j_n+1)s}}{\prod_{\kappa=1,\kappa\neq\nu}^{n}\beta_\kappa(j_\kappa+1)-\beta_\nu(j_\nu+1)}\right\}. \end{aligned} \tag{16}$$

From Lemma 1, we have

$$\begin{aligned} &\sum_{\nu=1}^{n}\frac{1}{\prod_{\kappa=1,\kappa\neq\nu}^{n}\beta_\kappa(j_\kappa+1)-\beta_\nu(j_\nu+1)} \\ &=\sum_{\nu=1}^{n-1}\frac{1}{\prod_{\kappa=1,\kappa\neq\nu}^{n}\beta_\kappa(j_\kappa+1)-\beta_\nu(j_\nu+1)}+\frac{1}{\prod_{\kappa=1}^{n-1}\beta_\kappa(j_\kappa+1)-\beta_n(j_n+1)}=0. \end{aligned} \tag{17}$$

Using (17) in (16), we obtain

$$\begin{aligned} f_n(s) &= \left(\prod_{i=1}^{n} 1-p_i\right)\sum_{j_1=0}^{\infty}\sum_{j_2=0}^{\infty}\cdots\sum_{j_n=0}^{\infty}\left(\prod_{i=1}^{n} p_i^{j_i}\beta_i(j_i+1)\right) \\ &\quad\times\left\{\sum_{\nu=1}^{n-1}\frac{e^{-\beta_\nu(j_\nu+1)s}}{\prod_{\kappa=1,\kappa\neq\nu}^{n}\beta_\kappa(j_\kappa+1)-\beta_\nu(j_\nu+1)}\right. \\ &\quad\left.+\frac{e^{-\beta_n(j_n+1)s}}{\prod_{\kappa=1}^{n-1}\beta_\kappa(j_\kappa+1)-\beta_n(j_n+1)}\right\}. \end{aligned}$$

$$\begin{aligned} f_n(s) &= \left(\prod_{i=1}^{n} 1-p_i\right)\sum_{j_1=0}^{\infty}\sum_{j_2=0}^{\infty}\cdots\sum_{j_n=0}^{\infty}\left(\prod_{i=1}^{n} p_i^{j_i}\beta_i(j_i+1)\right) \\ &\quad\times\left(\sum_{\nu=1}^{n}\frac{e^{-\beta_\nu(j_\nu+1)s}}{\prod_{\kappa=1,\kappa\neq\nu}^{n}\beta_\kappa(j_\kappa+1)-\beta_\nu(j_\nu+1)}\right), \end{aligned} \tag{18}$$

now, using (7) in (18), we get

$$\begin{aligned} f_n(s) &= \left(\prod_{i=1}^{n} 1-p_i\right)\sum_{j_1=0}^{\infty}\sum_{j_2=0}^{\infty}\cdots\sum_{j_n=0}^{\infty}\left(\prod_{i=1}^{n} p_i^{j_i}\beta_i(j_i+1)\right) \\ &\quad\times e[-\beta_1(j_1+1),\cdots,-\beta_n(j_n+1)], \qquad s>0. \end{aligned}$$

Thus, $f_n(s)$ is true whenever $f_{n-1}(s)$ is true. Hence, by the principle of mathematical induction, $f_n(s)$ is true for all $n \geq 2$. □

Remark 1. *When $p_i = 0$, $i = 1, \cdots, n$, Equation (12) is reduced to the density for the sum of independent exponential random variables with distinct parameters, β_i, $i = 1, \cdots, n$, which is equivalent to Levy (2022) as in Equation (10).*

Proposition 2. *The CDF of the sum of independent and non-identically distributed EG random variables is given by*

$$\begin{aligned} F_n(s) &= 1 - \left(\prod_{i=1}^{n} 1 - p_i \right) \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} \cdots \sum_{j_n=0}^{\infty} \left(\prod_{i=1}^{n} p_i^{j_i} \right) \\ &\quad \times \left(\sum_{\nu=1}^{n} \psi_{\nu,n} \quad e^{-\beta_\nu (j_\nu+1) s} \right), \end{aligned} \tag{19}$$

where $\psi_{\nu,n} = \prod_{\kappa=1,\kappa\neq\nu}^{n} \frac{\beta_\kappa (j_\kappa+1)}{\beta_\kappa (j_\kappa+1) - \beta_\nu (j_\nu+1)}$.

Proof. The CDF of S_n is easily derived from its PDF, which is given in (12) as

$$\begin{aligned} F_n(s) &= P(S_n \leqslant s) \\ &= \int_0^s f_n(z) dz \\ &= \left(\prod_{i=1}^{n} 1 - p_i \right) \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} \cdots \sum_{j_n=0}^{\infty} \left(\prod_{i=1}^{n} p_i^{j_i} \beta_i (j_i + 1) \right) \\ &\quad \times \left(\sum_{\nu=1}^{n} \frac{\int_0^s e^{-\beta_\nu (j_\nu+1) z} dz}{\prod_{\kappa=1,\kappa\neq\nu}^{n} \beta_\kappa (j_\kappa + 1) - \beta_\nu (j_\nu + 1)} \right) \\ &= \left(\prod_{i=1}^{n} 1 - p_i \right) \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} \cdots \sum_{j_n=0}^{\infty} \left(\prod_{i=1}^{n} p_i^{j_i} \right) \\ &\quad \times \sum_{\nu=1}^{n} \psi_{\nu,n} \left(1 - e^{-\beta_\nu (j_\nu+1) s} \right). \end{aligned} \tag{20}$$

From Smaili et al. (2013), Corollary 2,

$$\sum_{\nu=1}^{n} \psi_{\nu,n} = 1. \tag{21}$$

Now, using (21) in (20) and simplifying, we get (19). □

4 The distribution of the sum of independent EG variables with identical parameters

By substituting $p_i = p$ and $\beta_i = \beta$, $i = 1, \cdots, n$, into propositions 1 and 2, we obtain the PDF and CDF of the sum of independent and identically distributed EG random variables as shown in corollaries 1 and 2.

Corollary 1. *The PDF of the sum of independent and identically distributed EG random variables is expressed as*

$$\begin{aligned} f_n(s) &= (1-p)^n \beta^n \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} \cdots \sum_{j_n=0}^{\infty} p^{\varphi} \left(\prod_{i=1}^{n} (j_i+1) \right) \\ & \quad \times e[-\beta(j_1+1), \cdots, -\beta(j_n+1)], \qquad s>0, \end{aligned} \tag{22}$$

where $\varphi = \sum_{i=1}^{n} j_i$.

Remark 2. *For* $p=0$*, we get the density for the sum of independent exponential random variables with identical parameters as*

$$\begin{aligned} f_n(s) &= \beta^n e[-\beta, \cdots, -\beta] \\ &= \beta^n \frac{e^{(n-1)}(-\beta)}{(n-1)!} = \frac{\beta^n s^{n-1} e^{-\beta s}}{(n-1)!}, \end{aligned} \tag{23}$$

as obtained by Levy (2022).

Corollary 2. *The CDF of the sum of independent and identically distributed EG random variables is expressed as*

$$F_n(s) = 1-(1-p)^n \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} \cdots \sum_{j_n=0}^{\infty} p^{\varphi} \left(\sum_{\nu=1}^{n} \psi_{\nu,n} \quad e^{-\beta_\nu (j_\nu+1)s} \right), \tag{24}$$

where $\psi_{\nu,n} = \prod_{\kappa=1,\kappa\neq\nu}^{n} \frac{(j_\kappa+1)}{(j_\kappa+1)-(j_\nu+1)}$.

5 The MGF for the sum of independent EG random variables

In this section, we obtain the MGF for the sum of independent EG random variables with distinct parameters.

Proposition 3. *Suppose* $X_i, i=1,\cdots,n$ *are independent EG random variables with parameters* (p_i, β_i)*. Then, the MGF of* $S_n = \sum_{i=1}^{n} X_i$ *is given by*

$$M_S(t) = \left(\prod_{i=1}^{n} 1-p_i \right) \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} \cdots \sum_{j_n=0}^{\infty} \left(\prod_{i=1}^{n} p_i^{j_i} \frac{\beta_i(j_i+1)}{\beta_i(j_i+1)-t} \right). \tag{25}$$

Proof. The MGF for the sum of independent and non-identically distributed EG random variables is defined as

$$M_S(t) = E[e^{ts}] = \int_0^{\infty} e^{ts} f_n(s) ds. \tag{26}$$

By substituting for $f_n(s)$ from (12) into (26), we get

$$
\begin{aligned}
M_S(t) &= \left(\prod_{i=1}^{n} 1-p_i\right) \sum_{j_1=0}^{\infty}\sum_{j_2=0}^{\infty}\cdots\sum_{j_n=0}^{\infty}\left(\prod_{i=1}^{n} p_i^{j_i}\beta_i(j_i+1)\right) \\
&\quad \times \int_0^{\infty} e^{ts} e[-\beta_1(j_1+1),\cdots,-\beta_n(j_n+1)]ds. \qquad (27)
\end{aligned}
$$

Using (7) in (27), we obtain

$$
\begin{aligned}
M_S(t) &= \left(\prod_{i=1}^{n} 1-p_i\right) \sum_{j_1=0}^{\infty}\sum_{j_2=0}^{\infty}\cdots\sum_{j_n=0}^{\infty}\left(\prod_{i=1}^{n} p_i^{j_i}\beta_i(j_i+1)\right) \\
&\quad \times \left(\sum_{\nu=1}^{n} \frac{\int_0^{\infty} e^{-(\beta_\nu(j_\nu+1)-t)s}ds}{\prod_{\kappa=1,\kappa\neq\nu}^{n}\beta_\kappa(j_\kappa+1)-\beta_\nu(j_\nu+1)}\right) \\
&= \left(\prod_{i=1}^{n} 1-p_i\right) \sum_{j_1=0}^{\infty}\sum_{j_2=0}^{\infty}\cdots\sum_{j_n=0}^{\infty}\left(\prod_{i=1}^{n} p_i^{j_i}\beta_i(j_i+1)\right) \\
&\quad \times \left(\sum_{\nu=1}^{n} \frac{1}{(\beta_\nu(j_\nu+1)-t)\prod_{\kappa=1,\kappa\neq\nu}^{n}\beta_\kappa(j_\kappa+1)-\beta_\nu(j_\nu+1)}\right). \qquad (28)
\end{aligned}
$$

Now, using Lemma 2, we get the result. □

Remark 3. *When $p_i = 0, i = 1,\cdots,n$, the MGF (25) is reduced to the MGF for the sum of independent exponential random variables with non-identical parameters.*

6 Numerical results

In Section 3, we obtained the exact CDF of the sum of independent and non-identically distributed EG random variables. However, as the number of random variables increases, it becomes more difficult to calculate the exact probability. Hence, we need to estimate the probability using an approximation method. In this section, we used the SA proposed by Daniels (1954, 1987) and developed by Lugannani and Rice (1980).
Let $X_i, i = 1,\cdots,n$ be independent EG random variables with parameters (p_i,β_i), and $S_n = \sum_{i=1}^{n} X_i$ denote the sum. The MGF and CGF of S_n are given by

$$
M_S(t) = \prod_{i=1}^{n} M_{X_i}(t), \qquad (29)
$$

and

$$
\begin{aligned}
K(t) &= log\, M_S(t) \\
&= \sum_{i=1}^{n} log(1-p_i) + \sum_{i=1}^{n} log\lambda(t,p_i,\beta_i), \qquad (30)
\end{aligned}
$$

respectively, where $\lambda(t,p_i,\beta_i)=\sum_{j=0}^{\infty}p_i^j(1-\frac{t}{\beta_i(j+1)})^{-1}$.
The first two derivatives of $K(t)$ with respect to t are

$$K^{'}(t) = \sum_{i=1}^{n}\frac{\lambda^{'}(t,p_i,\beta_i)}{\lambda(t,p_i,\beta_i)}, \tag{31}$$

$$K^{''}(t) = \sum_{i=1}^{n}\left\{\frac{\lambda^{''}(t,p_i,\beta_i)}{\lambda(t,p_i,\beta_i)}-\left(\frac{\lambda^{'}(t,p_i,\beta_i)}{\lambda(t,p_i,\beta_i)}\right)^2\right\}, \tag{32}$$

where $\lambda^{'}(t,p_i,\beta_i)=\sum_{j=0}^{\infty}\frac{p_i^j}{\beta_i(j+1)}(1-\frac{t}{\beta_i(j+1)})^{-2}$, and $\lambda^{''}(t,p_i,\beta_i)=\sum_{j=0}^{\infty}\frac{2p_i^j}{(\beta_i(j+1))^2}(1-\frac{t}{\beta_i(j+1)})^{-3}$.
According to Lugannani and Rice (1980), the SA of the CDF of S_n is as follows:

$$F_S(x) = \Phi(\hat{w})+\phi(\hat{w})(\hat{w}^{-1}-\hat{v}^{-1})+O(n^{-\frac{3}{2}}), \tag{33}$$

where $\phi(.)$ and $\Phi(.)$ are the PDF and CDF of a standard normal distribution, respectively,

$$\hat{w}=sgn(\hat{s})\{2[\hat{s}\,x-K(\hat{s})]\}^{\frac{1}{2}},$$
$$\hat{v}=\hat{s}\{K^{''}(\hat{s})\}^{\frac{1}{2}},$$

where $sgn(\hat{s})=+1$,-1 or 0, depending on whether $\hat{s}$ is positive, negative or zero; and $\hat{s}$ is the root of $K^{'}(s)=x$ which is solved numerically by the Newton-Raphson algorithm.
Here, we assumed that p_i and β_i for $n=2$, 3, and 5 as follows:
Case I: p_i simulated from the beta distribution $Beta(1.6,2.1)$ and β_i simulated from the uniform distribution $U(0,3)$

- n=2:
  $p_i=(0.83622,0.03969)$, $\beta_i=(2.29192,2.41286)$.
- n=3:
  $p_i=(0.35209,0.16908,0.45529)$, $\beta_i=(2.86754,2.10982,2.51217)$.
- n=5:
  $p_i=(0.20888,0.05032,0.78634,0.54273,0.68281)$,
  $\beta_i=(1.72461,2.4519,2.58802,2.38057,2.08476)$.

Case II: p_i simulated from the beta distribution $Beta(1.2,2.3)$ and β_i simulated from the gamma distribution $Gamma(2,1.5)$

- n=2:
  $p_i=(0.06627,0.22026)$, $\beta_i=(2.28903,2.98899)$.
- n=3:
  $p_i=(0.64500,0.05876,0.30688)$, $\beta_i=(2.54877,3.73808,2.10423)$.

- n=5:
  $p_i = (0.47353, 0.72432, 0.59492, 0.09691, 0.22232)$,
  $\beta_i = (1.19819, 1.86453, 3.06197, 1.82964, 2.48412)$.

Case III: p_i simulated from the beta distribution $Beta(1.6, 2.1)$ and β_i simulated from the Exponential distribution $Exponential(0.13)$

- n=2:
  $p_i = (0.29076, 0.12463)$, $\beta_i = (30.347, 3.26497)$.

- n=3:
  $p_i = (0.65733, 0.08973, 0.11112)$, $\beta_i = (14.8452, 5.98085, 4.92978)$.

- n=5:
  $p_i = (0.19850, 0.17897, 0.12182, 0.32765, 0.34383)$,
  $\beta_i = (5.71698, 6.13874, 12.2445, 8.26087, 10.1468)$.

In Tables 1-3, $\hat{s}$ is a percentile derived from 10^6 random numbers generated by S_n, and P is the corresponding probability. In the tables, *r.e.* represents the relative error between the approximation and P, and F_M is the approximate CDF, which is truncated in the infinite series in Equation (19) after $M + 1$ terms. Moreover, F_S and F_N are the saddlepoint CDF from (33) and the NA, respectively. All computations here were conducted using Mathematica version 12. Tables 1, 2, and 3 demonstrate that the relative error for SA is always less than that for NA and that the relative errors decrease as P increases in general.

7 Conclusion

In this paper, we have obtained the exact distribution of the sum of independent and non-identically distributed EG random variables. We have also discussed the case of identically distributed EG random variables. Additionally, we have provided the approximate distribution by using the SA. Numerical results demonstrated that the SA outperformed the NA in terms of accuracy, and the SA is suitable for the CDF. The future challenge is to obtain the distribution of the sum of independent EG random variables when some of the parameters are identical and others are not.

Acknowledgement

The author is grateful to Prof. Bander Al-Zahrani, King Abdulaziz University, Jeddah, Saudi Arabia, for his helpful discussions throughout the preparation of this article.

Table 1: Numerical results for Case I

$\hat{s}$	P	F_M	F_N	F_S	$r.e.F_N$	$r.e.F_S$
			$n=2$			
0.5362	0.6000	0.6002	0.4796	0.6017	0.2006	0.0029
0.6750	0.7000	0.7006	0.5941	0.7012	0.1512	0.0018
0.8668	0.8000	0.8004	0.7383	0.8003	0.0772	0.0004
1.1904	0.9000	0.9002	0.9053	0.8997	0.0059	0.0004
1.5096	0.9500	0.9500	0.9760	0.9495	0.0274	0.0006
1.8260	0.9750	0.9749	0.9958	0.9745	0.0214	0.0005
2.2458	0.9900	0.9900	0.9998	0.9898	0.0099	0.0002
			$n=3$			
1.0149	0.6000	0.6000	0.5097	0.6002	0.1505	0.0003
1.2043	0.7000	0.7002	0.6249	0.7002	0.1073	0.0002
1.4547	0.8000	0.7999	0.7602	0.7998	0.0497	0.0002
1.8588	0.9000	0.9000	0.9090	0.9000	0.0100	0.0000
2.2448	0.9500	0.9501	0.9734	0.9502	0.0247	0.0002
2.6199	0.9750	0.9752	0.9941	0.9753	0.0196	0.0003
3.0981	0.9900	0.9901	0.9994	0.9901	0.0095	0.0001
			$n=5$			
1.6562	0.6000	0.6021	0.5227	0.6003	0.1288	0.0005
1.9088	0.7000	0.7011	0.6362	0.6998	0.0911	0.0003
2.2396	0.8000	0.8005	0.7673	0.7998	0.0409	0.0003
2.7592	0.9000	0.8997	0.9081	0.8994	0.0090	0.0007
3.2479	0.9500	0.9496	0.9708	0.9496	0.0219	0.0004
3.7175	0.9750	0.9747	0.9925	0.9748	0.0180	0.0002
4.3182	0.9900	0.9899	0.9991	0.9899	0.0092	0.0001

Table 2: Numerical results for Case II

$\hat{s}$	P	F_M	F_N	F_S	$r.e.F_N$	$r.e.F_S$
			$n = 2$			
0.7154	0.6000	0.6002	0.4990	0.6008	0.1684	0.0014
0.8707	0.7000	0.6999	0.6139	0.7004	0.1230	0.0006
1.0806	0.8000	0.7999	0.7531	0.8004	0.0587	0.0005
1.4232	0.9000	0.8999	0.9080	0.9003	0.0089	0.0003
1.7525	0.9500	0.9498	0.9743	0.9501	0.0256	0.0001
2.0759	0.9750	0.9749	0.9947	0.9751	0.0202	0.0001
2.5007	0.9900	0.9901	0.9996	0.9901	0.0097	0.0001
			$n = 3$			
0.8822	0.6000	0.6003	0.5038	0.6024	0.1603	0.0039
1.0513	0.7000	0.7003	0.6176	0.7019	0.1177	0.0026
1.2791	0.8000	0.8006	0.7547	0.8014	0.0566	0.0017
1.6504	0.9000	0.9003	0.9075	0.9005	0.0083	0.0006
2.0120	0.9500	0.9504	0.9741	0.9504	0.0254	0.0004
2.3608	0.9750	0.9750	0.9945	0.9750	0.0200	0.0000
2.8147	0.9900	0.9899	0.9995	0.9899	0.0096	0.0001
			$n = 5$			
1.9803	0.6000	0.6002	0.5180	0.6031	0.1366	0.0051
2.2854	0.7000	0.7002	0.6308	0.7032	0.0989	0.0046
2.6863	0.8000	0.8001	0.7622	0.8028	0.0472	0.0035
3.3326	0.9000	0.9004	0.9075	0.9021	0.0083	0.0023
3.9426	0.9500	0.9500	0.9715	0.9509	0.0226	0.0009
4.5500	0.9750	0.9753	0.9934	0.9757	0.0189	0.0007
5.3358	0.9900	0.9902	0.9994	0.9903	0.0095	0.0003

Table 3: Numerical results for Case III

$\hat{s}$	P	F_M	F_N	F_S	$r.e.F_N$	$r.e.F_S$
			$n = 2$			
0.2867	0.6000	0.6006	0.4635	0.6026	0.2275	0.0043
0.3708	0.7000	0.7006	0.5757	0.7011	0.1776	0.0015
0.4909	0.8000	0.8005	0.7239	0.7999	0.0951	0.0001
0.6989	0.9000	0.9003	0.9021	0.8994	0.0023	0.0006
0.9092	0.9500	0.9502	0.9773	0.9496	0.0287	0.0004
1.1182	0.9750	0.9749	0.9966	0.9746	0.0221	0.0005
1.4018	0.9900	0.9901	0.9999	0.9899	0.0100	0.0001
			$n = 3$			
0.3920	0.6000	0.6000	0.5059	0.6009	0.1569	0.0015
0.4680	0.7000	0.6997	0.6202	0.7004	0.1139	0.0005
0.5691	0.8000	0.7994	0.7563	0.7998	0.0546	0.0002
0.7333	0.9000	0.8994	0.9074	0.8996	0.0082	0.0004
0.8925	0.9500	0.9499	0.9736	0.9501	0.0248	0.0001
1.0454	0.9750	0.9748	0.9942	0.9749	0.0197	0.0001
1.2452	0.9900	0.9900	0.9995	0.9900	0.0096	0.0000
			$n = 5$			
0.5784	0.6000	0.6001	0.5293	0.6010	0.1179	0.0016
0.6591	0.7000	0.6998	0.6421	0.7006	0.0828	0.0009
0.7634	0.8000	0.7994	0.7701	0.8000	0.0373	0.0000
0.9275	0.9000	0.8995	0.9083	0.8999	0.0092	0.0001
1.0807	0.9500	0.9497	0.9701	0.9499	0.0211	0.0001
1.2299	0.9750	0.9751	0.9922	0.9752	0.0177	0.0002
1.4187	0.9900	0.9901	0.9990	0.9902	0.0091	0.0002

References

Adamidis, K. and Loukas, S. (1998). A lifetime distribution with decreasing failure rate. *Statistics & Probability Letters*, 39(1):35–42.

Akkouchi, M. (2008). On the convolution of exponential distributions. *Journal of the chungcheong mathematical society*, 21(4):501–501.

Atkinson, K. (1991). *An introduction to numerical analysis.* John wiley & sons.

Balakrishnan, N., Zhu, X., and Al-Zahrani, B. (2015). A recursive algorithm for the single and product moments of order statistics from the exponential-geometric distribution and some estimation methods. *Communications in Statistics-Theory and Methods*, 44(17):3576–3598.

Daniels, H. E. (1954). Saddlepoint approximations in statistics. *The Annals of Mathematical Statistics*, pages 631–650.

Daniels, H. E. (1987). Tail probability approximations. *International Statistical Review/Revue Internationale de Statistique*, pages 37–48.

Gradshteyn, I. S. and Ryzhik, I. M. (2014). *Table of integrals, series, and products.* Academic press.

Kitani, M. and Murakami, H. (2020). On the distribution of the sum of independent and non-identically extended exponential random variables. *Japanese Journal of Statistics and Data Science*, 3:23–37.

Kitani, M., Murakami, H., and Hashiguchi, H. (2023). The distribution of the sum of independent and non identically generalized lindley random variables. *Communications in Statistics-Theory and Methods*, 52(8):2597–2609.

Levy, E. (2021). On the density for sums of independent mittag-leffler variates with common order. *Statistics & Probability Letters*, 179:109211.

Levy, E. (2022). On the density for sums of independent exponential, erlang and gamma variates. *Statistical Papers*, 63(3):693–721.

Louzada-Neto, F. (1999). Polyhazard models for lifetime data. *Biometrics*, 55(4):1281–1285.

Lugannani, R. and Rice, S. (1980). Saddle point approximation for the distribution of the sum of independent random variables. *Advances in applied probability*, 12(2):475–490.

Mathai, A. M. (1982). Storage capacity of a dam with gamma type inputs. *Annals of the Institute of Statistical Mathematics*, 34:591–597.

Moschopoulos, P. G. (1985). The distribution of the sum of independent gamma random variables. *Annals of the Institute of Statistical Mathematics*, 37(1):541–544.

Murakami, H. (2014). A saddlepoint approximation to the distribution of the sum of independent non-identically uniform random variables. *Statistica Neerlandica*, 68(4):267–275.

Murakami, H. (2015). Approximations to the distribution of sum of independent non-identically gamma random variables. *Mathematical Sciences*, 9(4):205–213.

Nadarajah, S., Jiang, X., and Chu, J. (2015). A saddlepoint approximation to the distribution of the sum of independent non-identically beta random variables. *Statistica Neerlandica*, 69(2):102–114.

Ross, S. M. (2014). *Introduction to probability models.* Academic press.

Sadooghi-Alvandi, S., Nematollahi, A., Habibi, R., et al. (2009). On the distribution of the sum of independent uniform random variables. *Statistical papers*, 50(1):171–175.

Smaili, K., Kadri, T., and Kadry, S. (2013). Hypoexponential distribution with different parameters.

Van Khuong, H. and Kong, H.-Y. (2006). General expression for pdf of a sum of independent exponential random variables. *IEEE Communications Letters*, 10(3):159–161.

Yaghoubi, A. (2022). Sum of independent random variable for shanker, akash, ishita, pranav, rani and ram awadh distributions. *arXiv preprint arXiv:2208.08006*.