Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/130137

An Approximate Model for Total Amount of Non-life Insurance Claims using Generalized Gamma Distribution and H-Function

E. Abderrahim[1], *C. Toufik*[2], *E. Youness*[3], *K. Yassine*[1,*] and *E. Mostafa*[1]

[1]Department of Statistics and Mathematics Applied to Economics and Management, Faculty of Economic and Social Juridical Sciences Ain Sebaâ, Hassan II University, Casablanca, Morocco
[2]Mathematics Department, Pluridisciplinary Faculty, Mohammed First University, Nador, Morocco
[3]IRF-SIC Laboratory, Faculty of Sciences, Ibn Zohr University, Agadir, Morocco

Received: 15 Mar. 2023, Revised: 28 Aug. 2023, Accepted: 6 Sep. 2023.
Published online: 1 Jan. 2024.

Abstract: This article proposes an analytical method to approximate the probability density function (PDF) and the cumulative distribution function (CDF) of the total amount of non-life claims to be paid by the insurer over a financial period considered. The individual claims amounts are independent positive random variables following the generalized gamma distribution (GGD) and distributed in a non-identical manner. The analytical approach suggested relies on the Fox H-function. The Fox H-function has various applications available in the literature. The method developed has demonstrated its performance both in respect of the result obtained (in comparison to the Monte-Carlo method) and in respect of simplicity (easily accessible for the most common claims amount distributions). The resulting PDF expression can be directly used to estimate the technical benefit, total cost, and ruin probability of the non-life insurance company.
Keywords: Generalized gamma distribution, Fox H-function, Amount of claims, Non-life insurance, Probability density function, Cumulative distribution function.

1 Introduction

The search for the total amount claims distribution is one of the main points in the mathematics of insurance. In fact, to quantify the amount of the premium and determine the solvency margins of a non-life insurance company, the actuary must study the distribution of the accumulation of benefits and know its expectation, its variance, its density function, or still its quantiles.

In this regard, we are faced a problem that incites distinct interest in the actuarial mathematics research. The principal purpose is to find the overall amount distribution related to the portfolio of insurance contracts that it has taken out. This issue evidently depends on the insurance company strategy that is specific to each insurer and depends on various criterion, namely, the individual amount claims distribution and the portfolio nature (independence, the homogeneity, etc.).

The overall amount of claims can be written as the sum, over the number of policies, of the total amount of claims generated by each policy or as the sum, over the number of claims, of the amounts of each claim. The first approach is called an individual model and the second is a collective model [1]. In fact, the individual model [1] makes it possible to approach the collective model [2].

The existing literature on this topic might be categorized into two segments based on the solving methods namely the exact approach and the numerical approach. The segment of the literature using exact approaches is initiated by Panjer [3] who proposed an algorithm which makes it possible to find, in recursive form, the PDF of the total amount of claims. Panjer's algorithm was then generalized by Dhaene and Collab [4] who proposed a more general recursive formula. Several other works have subsequently been proposed to approximate the total amount claims distribution via an exact approach, we find, among others, the approach of Heckman and Meyers [5] who used the FFT (Fast Fourier Transform) algorithm to express the law of the total amount of claims under certain conditions, and the approach of Partrat and Besson [6] who used Bowers' Gamma approximation to determine the total amount claims distribution function via the orthogonal Laguerre polynomials.

The second part of the literature offers a rather numerical vision to compete with the Panjer's algorithm and the FFT algorithm. The methods proposed in this context make it possible to significantly increase the quality of the precision without increasing the calculation times. Among these methods, we find the approximation of Mnatsakanov and Ruymgaart [7] which makes it possible to approximate the total amount claims distribution function via exponential moments from the Laplace transform, and the approximation of Gzyl and Tagliani [8] which makes it possible to get the total amount claims distribution function via fractional moments.

*Corresponding author e-mail: yassinekouach90@gmail.com

The recent generation like Goffard [9] proposed an approximation approach of the total amount claims distribution function using the polynomial approximation defined by a linear combination of moments of the distribution.

The common idea of the previous approaches in the literature aimed to find the total amount claims distribution. However, the problem with these distributions is that their probability density is only accessible in a certain number of cases. This is why many attempts have been made to improve these methods or to develop new methods (see for instance [10,11,12], etc.).

In the individual model context, this paper presents an analytical method for approximating the function of the probability density (PDF) and the function of the cumulative distribution (CDF) of the total amount of claims associated with independent random variables admitting a generalized gamma distribution (GGD) and distributed in a non-identical way. We are looking here for an approximate distribution of the total amount of claims based on GG distribution in agreement with the Fox H-function. The main objective is to significantly increase the quality of the precision and optimize calculation times. With a view to assess the validity of the proposed approximation method, the Monte Carlo simulations technique was used.

In order to carry out our study, we have divided our paper into four sections. the next chapter is a preliminary chapter which highlights the concepts and definitions used as mathematical tools on individual and collective models, usual distributions of the amount of claims, GGD, and Fox H-function. We then propose in the third section, an expression for the function of the probability density (PDF) and the function of the cumulative distribution (CDF) of the sum of independent, not identically random variables distributed using the H-function. An application of our study to non-life assurance is presented in the fourth section where we seek the total amount claims distribution, represent the results of the Monte Carlo simulations and offer insights into the performance of the approximation. In the last section, we will end this paper with a conclusion.

2. Preliminary

2.1 Modeling of the total amount of claims

Two models are studied to model the total amount of claims: individual and collective models. The individual model allows to quantify the total amount of claims payable by the insurance company over a given period, by summing policy by policy the amounts of claims suffered by each individual over this period.

We consider an insurance portfolio composed of n random variables $X_1, X_2, \dots, X_n$ with values in $\mathbb{R}^+$, independent, but not necessarily of the same law, represent the total amounts of claims generated by each policy. The total claim amount in the individual model S_{ind} is given by [13] :

$$S_{ind} = \sum_{i=1}^{n} X_i \tag{1}$$

The collective model no longer considers the policies individually but the portfolio as a whole: the overall claims charge is expressed as a function of the amount of each claim and no longer as a function of the total amount (possibly zero) of claims generated by each policy [14].

The number of claims suffered by policyholders over this period is modelled by a random variable N to value in N. Claims charges are modelled by random variables $(X_i)_{i\in[1,\dots,N]}$ positive, continuous and independent of N. In this model the total claim load is given by:

$$S_{coll} = \sum_{i=1}^{N} X_i \tag{2}$$

It is assumed by convention that $S_{coll} = 0$ if $N = 0$.

In fact, the collective model is an extension of the individual model; it allows to approximate it.

Note 2.1. In actuarial of non-life insurance, the usual probability laws most used to model the amount of claims X_i are those with positive support. The three most common positive laws are Gamma law, lognormal law, Weibull's law and Pareto law (case of serious claims).

2.2 Generalized Gamma Distribution

Generalized gamma (GGD) is a flexible distribution in the actuarial literature, it contains the Exponential law, Gamma and Weibull as subfamilies, and Log-normal as the limit distribution is introduced.

The GGD has a continuous probability density function with three parameters. It is a generalization of the gamma distribution with two parameters.

Definition 2.1. Let $(Z_i)_{i\in\{1,\dots,n\}} n$ independent positive random variables follow the Generalized gamma distribution (GGD) but distributed in a non-identical way.

For $z > 0$, the probability density function (PDF) of the generalized gamma (GGD) is given by:

$$f_{Z_i}(z) = \frac{p_i}{a_i\Gamma(d_i)}\left(\frac{z}{a_i}\right)^{d_i p_i - 1} e^{-(z/a_i)^{p_i}} \quad (3)$$

where $a_i > 0$ and $d_i > 0$ are the shape parameters, while $p_i > 0$ is the scale parameter and $\Gamma(.)$ denotes Euler's Gamma function.

The cumulative distribution function (CDF) is given by:

$$F_Z(z) = \frac{\gamma(d_i/p_i,(z/a_i)^{p_i})}{\Gamma(d_i/p_i)} \quad (4)$$

where $\gamma(.,.)$ denotes the lower incomplete gamma function.

Proposition 2.1. If $(Z_i)_{i\in\{1,\dots,n\}}$ has a generalized gamma distribution $Z_i \sim GG(a_i, d_i, p_i)$, then the power moments of GG distribution is given:

$$E({Z_i}^r) = {a_i}^r \frac{\Gamma\left(\frac{d_i+r}{p_i}\right)}{\Gamma\left(\frac{d_i}{p_i}\right)} \quad (5)$$

(Demonstration: see Johnson, Kotz and Balakrishnan [15]).

If $(Z)_{i\in\{1,\dots,n\}}$ has GG distribution, then

$$E(Z_i) = a_i \frac{\Gamma\left(\frac{d_i+1}{p_i}\right)}{\Gamma\left(\frac{d_i}{p_i}\right)} \quad \text{and} \quad Var(Z_i) = {a_i}^2 \frac{\Gamma\left(\frac{d_i+2}{p_i}\right)}{\Gamma\left(\frac{d_i}{p_i}\right)} - \left(a_i \frac{\Gamma\left(\frac{d_i+1}{p_i}\right)}{\Gamma\left(\frac{d_i}{p_i}\right)}\right)^2 \quad (6)$$

Note 2.2. If $d_i = p_i$ then the generalized gamma distribution becomes the Weibull distribution.

Else if $p_i = 1$ the generalized gamma becomes the gamma distribution.

Note 2.3. The GG distribution is sometimes parameterized by the substitution $r = \frac{d_i}{p_i}$ (Johnson, Kotz and Balakrishnan [15]). This gives a distribution called the Amoroso distribution, after the Italian mathematician and economist Luigi Amoroso.

2.3 Fox H-function

Despite its ancient roots, very versatile nature, the utilization of Fox's H-function in scientific research has only become prevalent in recent times as a possible distribution used to model physical and statistical phenomena [16].

The Fox H-function, introduced by Charles Fox [17], is an extension of the Meijer G-function and the Fox Wright-function. It is defined through a Mellin–Barnes integral.

$$H(z) \triangleq H_{p,q}^{m,n}\left[z \left| \begin{matrix} (a_1,A_1),(a_2,A_2),\dots,(a_p,A_p) \\ (b_1,B_1),(b_2,B_2),\dots,(b_q,B_q) \end{matrix} \right.\right] = \frac{1}{2\pi j}\int_L h(s)\, z^{-s} ds \quad (7)$$

Where:

$$h(s) = \frac{\prod_{k=1}^{m}\Gamma(b_k+B_k s)\prod_{k=1}^{n}\Gamma(1-a_k-A_k s)}{\prod_{k=m+1}^{q}\Gamma(1-b_k-B_k s)\prod_{k=n+1}^{p}\Gamma(a_k+A_k s)} \quad (8)$$

Where a_k, b_k, and $z \neq 0$are complex numbers. $A_k, B_k, m, n, p$and q are real numbers, with,$0 \leq n \leq p$ and $1 \leq m \leq q$. $j = \sqrt{-1}$an imaginary number.

The integration path, denoted as L, is a contour within the complex s-plane where all right-half-plane (RHP) poles of $\prod_{k=1}^{m}\Gamma(b_k + B_k s)$ are positioned to the right of L, and all LHP poles of $\prod_{k=n+1}^{p}\Gamma(a_k + A_k s)$ are located to the left of L.

3. An approximate distribution for the sum of independent and identically distributed generalized Gaussian (GG) random variables

Within this section, we present an equation representing the distribution of the RV, $(Z_i)_{i\in[1,\dots,n]}$, using Fox's H-function.

Consequently, the PDF of the cumulative sum of i.n.i.d. GG RVs can be estimated using a distribution derived from Fox's H-function. This approximation relies on calculating the moments of the H-function distribution.

Proposition 3.1. Let $(Z_i)_{i\in[1,\dots,n]}$ be n i.n.i.d. $GG(a_i, d_i, p_i)$ RVs. The PDF of Z_i can be expressed as a H-function.

$$f_{Z_i}(z) = \frac{1}{a_i\Gamma(d_i)} H_{0,1}^{1,0}\left(\frac{z}{a_i}\middle|\begin{matrix}-\\ \left(d_i - \frac{1}{p_i}, \frac{1}{p_i}\right)\end{matrix}\right), \ z \geq 0 \tag{9}$$

In this context, d_i and p_i are the shape parameters, while a_i is the scale parameter.

Proof. By rewriting the exponential function as a Fox's H-function, according to equation (7) and equation (8) we have

$$e^{-(z/a_i)^{p_i}} = H_{0,1}^{1,0}\left(\left(\frac{z}{a_i}\right)^{p_i}\middle|\begin{matrix}-\\ (0,1)\end{matrix}\right) = \frac{1}{2\pi j}\int_L \Gamma(s)\left(\frac{z}{a_i}\right)^{-p_i s} ds \tag{10}$$

By substituting equation (10) into equation (3) and carrying out various calculations, we arrive at the following result:

$$f_{Z_i}(z) = \frac{p_i}{a_i\Gamma(d_i)}\frac{1}{2\pi j}\int_L \Gamma(s)\left(\frac{z}{a_i}\right)^{p_i(d_i-s)-1} ds \tag{11}$$

Lastly, through the utilization of the variable transformation, $t = p_i(s - d_i) + 1$, we derive equation (11), thus completing the proof of Proposition 3.1.

Property 3.1. The H-function $H_{0,1}^{1,0}$ can be written as an H-function $H_{1,1}^{1,0}$ as follows

$$H_{0,1}^{1,0}\left(\frac{z}{a_i}\middle|\begin{matrix}- & -\\ \left(d_i - \frac{2}{p_i}, \frac{2}{p_i}\right) & -\end{matrix}\right) = \frac{2\sqrt{\pi j}}{4^{d_i - \frac{2}{p_i}}} H_{1,1}^{1,0}\left(4^{\frac{2}{p_i}}\frac{z}{a_i}\middle|\begin{matrix}- & \left(d_i - \frac{2}{p_i} + \frac{1}{2}, \frac{2}{p_i}\right)\\ \left(2d_i - \frac{4}{p_i}, \frac{4}{p_i}\right) & -\end{matrix}\right), \ z \geq 0 \tag{12}$$

Proof. According to Bodenschatz [18], the H-function $H_{0,1}^{1,0}$ given in equation (10) can be written as an H-function $H_{1,1}^{1,0}$ as demonstrated in Property 3.1.

Proposition 3.2. introduces an approximation for the PDF of the cumulative sum of independent and identically distributed hyperbolic distributions (HDs). This approximation is based on the utilization of the moments-based approximation method, as outlined in references [18] and [19].

Proposition 3.2. The probability density function (PDF) of $Z = \sum_{i=1}^{n} Z_i$ can be estimated through approximation by a H-function as:

$$f_Z(z) \approx \alpha H_{1,1}^{1,0}\left(\frac{z}{\eta}\middle|\begin{matrix}- & (\beta,1)\\ (\delta,1) & -\end{matrix}\right), \ z \geq 0 \tag{13}$$

where η is a random arbitrary positive number, and

$$\alpha = \frac{\Gamma(\beta+1)}{\eta\Gamma(\delta+1)}, \beta = \frac{\eta\mu_1+\mu_2}{\mu_2-\mu_1^2} - 1, \text{ and } \delta = \left(\frac{\eta\mu_1-\mu_2}{\mu_2-\mu_1^2}\right)\frac{\mu_1}{\eta} - 1 \tag{14}$$

Moreover, the convergence of this PDF is contingent upon the validation of $\eta < \mu_1$ or $\eta > \frac{\mu_2}{\mu_1}$,

where μ_1 and μ_2 represent the first and second moments of Z, respectively.

Proof. The Bodenschatz theorem provides us with Equation (15):

$$\begin{aligned} f_Z(z) &\approx \alpha H_{1,1}^{1,0}\left(\frac{z}{\eta}\middle|\begin{matrix}- & (\beta,1)\\ (\delta,1) & -\end{matrix}\right), z \geq 0 \\ &= \frac{\alpha}{2\pi j}\int_L \frac{\Gamma(\delta+s)}{\Gamma(\beta+s)}\left(\frac{z}{\eta}\right)^{-s} ds \end{aligned} \tag{15}$$

Where, for any arbitrarily selected positive number, the solutions of the following nonlinear system of equations derived from ([20], eq. (6.3.3b)) are η, α, β, and δ

$$\mu_n = \alpha\eta^{n+1}\frac{\Gamma(\delta+n+1)}{\Gamma(\beta+n+1)}, n = 0,1,2 \quad (16)$$

Using the value of $\mu_0 = 1$, along with ([21], eq. (6.1.15)), we can readily derive equation (16). Additionally, the expectation of Z can be represented as follows:

$$\mu_1 = \sum_{i=1}^{n} a_i \frac{\Gamma(d_i+1/p_i)}{\Gamma(d_i)} \quad (17)$$

Where $(a_i, d_i, p_i)$are generalized gamma distribution parameters.

Meanwhile, the second moment can be obtained using the multinomial theorem, resulting in the expression:

$$\mu_2 = \sum_{i=1}^{n} a_i^2 \frac{\Gamma(d_i+1/p_i)}{\Gamma(d_i)} + 2\sum_{i=1}^{n-1}\sum_{l=i+1}^{n} a_i a_l \frac{\Gamma(d_i+1/p_i)}{\Gamma(d_i)}\frac{\Gamma(d_l+1/p_i)}{\Gamma(d_l)}. \quad (18)$$

It's important to note that a condition sufficient for the convergence of the H-function, as presented in equation (15), is: $\delta - \beta < 0$. Thus, as $\mu_2 - \mu_1^2 > 0$, we conclude that the sign of $\delta - \beta$ aligns with the sign of$\left(\mu_1 - \frac{\mu_2}{\eta}\right)(\mu_1 - \eta)$. This allows for the selection of any arbitrary value of η that is either greater than $\frac{\mu_2}{\mu_1}$ or less than μ_1, thereby concluding the proof of Proposition 3.2.

Corollary 3.1. The H-function, as presented in equation (15), achieves convergence if:

$$\eta > \Delta, \forall z \geq 0 \quad (19)$$

with $$\Delta = \frac{\mu_2}{\mu_1}\left(1 + \sqrt{1 - \frac{\mu_1^2}{\mu_2}}\right) \quad (20)$$

Proof. If $\delta - \beta + 1 < 0$, We have the option to utilize the sum of either the residues from the LHP or the residues from the RHP at $z = \eta$. Additionally,

$$\delta - \beta + 1 = \frac{\mu_1\eta^2 - 2\mu_2\eta + \mu_1\mu_2}{\eta(\mu_2 - \mu_1^2)}$$

(21)

That is, as $\mu_2 - \mu_1^2 > 0$ any value of η greater than Δ can be chosen to have$\delta - \beta + 1 < 0$. because$\Delta > \frac{\mu_2}{\mu_1}$, the PDF provided in equation (15) exhibits convergence for any value of z if $\eta > \Delta$. This marks the completion of the proof for Corollary 3.1.

Considering equation (15), the existence of the PDF for the sum of random variables depends on satisfying the conditions $\eta > \frac{\mu_2}{\mu_1}$ and/or $\eta > \Delta, \forall z \geq 0$. Moreover, as increases$\eta$, the accuracy of this approximation improves. However, when dealing with large values of η, computational instability can arise, especially when using predefined routines in software like Mathematica or Matlab. Consequently, it is evident that for significantly large values of η, we can further simplify this expression. In light of this, to obtain the asymptotic expression, we will omit the parameter η from the PDF expression by evaluating the limit of equation (15) as η approaches infinity.

Theorem 3.1. If $(Z_i)_{i\in[1,\dots,n]}$ are i.n.i.d GG variates with PDFs $f_{Z_i}(z)$ described in equation (9), then the PDF of $Z = \sum_{i=1}^{n} Z_i$ can be approximated as follows:

$$f_Z(z) \approx \frac{\psi}{\Gamma(\phi)} H_{0,1}^{1,0}\left(\psi z \middle| \begin{matrix} - & - \\ (\phi - 1,1) & - \end{matrix}\right),\ z \geq 0 \quad (22)$$

Where $\phi = \frac{\mu_1^2}{\mu_2 - \mu_1^2}$ and $\psi = \frac{\mu_1}{\mu_2 - \mu_1^2}$.

Proof. By reproducing equation (15) ones gets.

$$f_Z(z) \approx \frac{\Gamma(\beta+1)}{\eta\Gamma(\delta+1)} H_{0,1}^{1,0}\left(\frac{z}{\eta} \middle| \begin{matrix} - & (\beta,1) \\ (\delta,1) & - \end{matrix}\right),\ z \geq 0$$

$$= \frac{\Gamma(\beta+1)}{\eta\Gamma(\delta+1)}\frac{1}{2\pi j}\int_L \frac{\Gamma(\delta+s)}{\Gamma(\beta+s)}\left(\frac{z}{\eta}\right)^{-s} ds$$

$$= \frac{1}{2\pi j\eta}\int_L \frac{\Gamma(\delta+s)}{\Gamma(\delta+1)}\frac{\Gamma(\beta+1)\eta^s}{\Gamma(\beta+s)} z^{-s} ds \quad (23)$$

Where:

$$\beta = \left(\eta - \frac{\mu_2}{\mu_1}\right)\psi - 1 \text{ and } \delta = \phi - \frac{\mu_2}{\eta}\psi - 1 \tag{24}$$

From equation (24), we have:

$$\frac{\Gamma(\delta+s)}{\Gamma(\delta+1)} \sim \frac{\Gamma(\phi-1+s)}{\Gamma(\phi)} \quad as\ \eta \to +\infty \tag{25}$$

Alternatively, by utilizing Stirling's relation ([21], eq. (6.1.39)), we can obtain:

$$\frac{\Gamma(\beta+1)}{\Gamma(\beta+s)}\eta^s \sim \eta\psi^{-s+1} \quad as\ \eta \to +\infty \tag{26}$$

Thus

$$\frac{1}{\eta}\frac{\Gamma(\delta+s)}{\Gamma(\delta+1)}\frac{\Gamma(\beta+1)\eta^s}{\Gamma(\beta+s)}z^{-s} \sim \frac{\Gamma(\phi-1+s)}{\Gamma(\phi)}\psi^{-s+1}z^{-s} \quad as\ \eta \to +\infty \tag{27}$$

It follows that the equation (23) becomes:

$$f_Z(z) \approx \frac{\psi}{\Gamma(\phi)}\frac{1}{2\pi j}\int_L \Gamma(\phi - 1 + s)(\psi z)^{-s} ds, \tag{28}$$

Thereby concluding the proof.

Corollary 3.2. The cumulative distribution of $Z = \sum_{i=1}^{n} Z_i$ can be accurately approximated as

$$F_Z(z) \approx \frac{1}{\Gamma(\phi)} H_{1,2}^{1,1}\left(\psi z \left| \begin{matrix} (1,1) & - \\ (\phi,1) & (0,1) \end{matrix} \right.\right), \ z \geq 0 \tag{29}$$

Proof. The CDF of Z can be straightforwardly derived from its PDF, as given in equation (15), as follows:

$$F_Z(z) = P(Z \leq z) = \int_0^z f_Z(x)dx$$

$$\approx \frac{\psi}{\Gamma(\phi)}\frac{1}{2\pi j}\int_L \Gamma(\phi - 1 + s)\psi^{-s}\left(\int_0^z x^{-s}dx\right)ds$$

$$\approx \frac{\psi}{\Gamma(\phi)}\frac{1}{2\pi j}\int_L \frac{\Gamma(\phi-1+s)}{-s+1}\psi^{-s}z^{-s+1}\,ds \tag{30}$$

Next, by leveraging ([21], eq. (6.1.15)), and performing some straightforward algebraic manipulations, along with implementing a linear change of variable $t = s - 1$, the CDF can be simplified to

$$F_Z(z) \approx \frac{1}{\Gamma(\phi)}\frac{1}{2\pi j}\int_{Cs} \frac{\Gamma(\phi+s)\Gamma(-s)}{\Gamma(1-s)}(\psi z)^{-s} ds \tag{31}$$

C_s is the domain of definition of s.

Hence, this completes the proof.

4. Application to non-life insurance

We seek to determine the PDF of the total amount of claims in the individual model, which corresponds to the cumulative amount of benefit to be paid by the insurer over a financial period considered, where the individual amounts of claims are independent positive random variables following the GG distribution but distributed in a non-identical manner. The parameters were chosen according to three subfamilies: the standard Gamma law, Log-normal and Weibull. These laws are the most common for modeling the amount of claims; they have quite different densities and have very different properties.

Let's a portfolio of non-life insurance company composed of n amount of claims $X_1, \ldots, X_n$ (i.n.i.d. GG RVs).

Consider a random variable X that represents the total amount of claims. Then, the distribution of the sum of $(X_i)_{i=[1,\ldots,n]}$ can be modeled using the PDF (equation (22)) presented in Theorem 3.1.

4.1 Numerical Results

In this part, the hypergeometric H-functions were assessed employing MATHEMATICA software. All the estimate expressions proposed are achieved using MATLAB software through Monte-Carlo simulations, involved by generating $10^7 n$ generalized gamma distributed random numbers. Additionally, the inverse transform sampling method [22] is applied, along with the use of an exponential decaying power delay profile with equispaced delays [23].

$$\mu_1^{(z_i)} = \mu_1^{(z_1)} e^{-\varphi(i-1)}, \ \forall i \in \{1, \dots, n\} \tag{32}$$

where φ is the average power decay factor. the meaning of the φ values is detailed in [23].

4.1.1 Approximated PDF

Weibull distribution

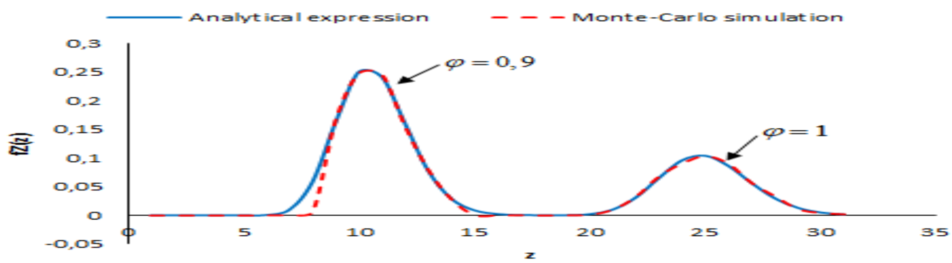


Fig. 1. *Sum PDF of i.n.i.d. GG RVs for d=1, p=3 (i.e., Weibull), n=32, and various values of φ.*

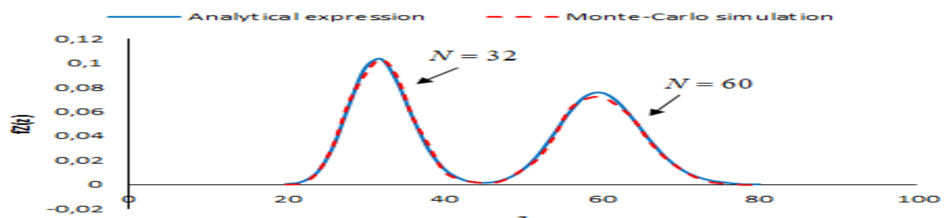


***Fig. 2**. Sum PDF of i.n.i.d. GG RVs for d=1, p=3 (i.e., Weibull), several values of n, and* $\varphi = 1$.

Gamma distribution

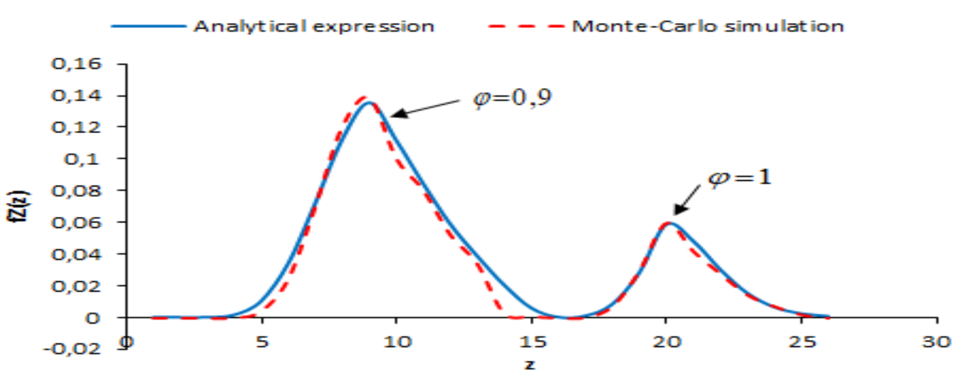


Fig. 3. *Sum PDF of i.n.i.d. GG RVs for d=3, p=1 (i.e., Gamma), n=32, and different values of φ.*

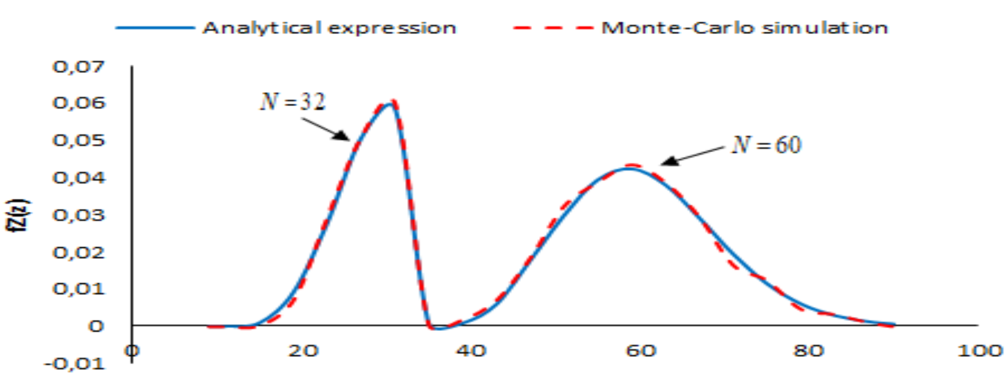
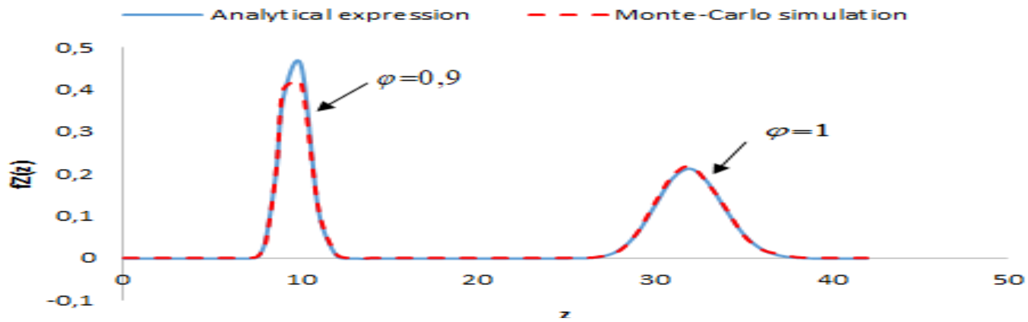


Fig. 4. *Sum PDF of i.n.i.d. GG RVs for d=3, p=1 (i.e., Gamma), various values of n, and* $\varphi = 1$.

Log-normal distribution



Fi g. 5*. Sum PDF of i.n.i.d. GG RVs for d=150, p=0.5 (i.e., Log-normal), n=32, and various values of φ.*

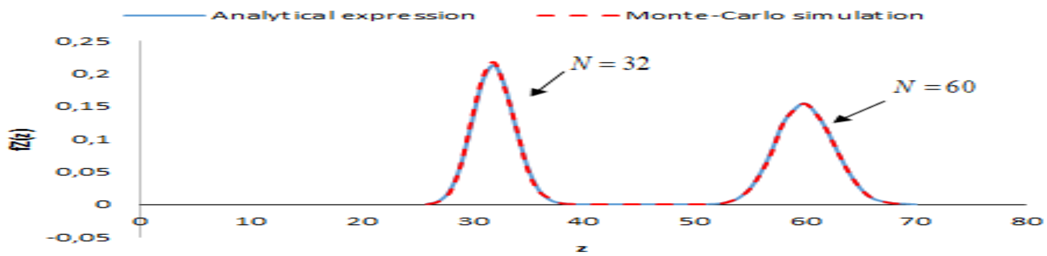


Fig. 6. *Sum PDF of i.n.i.d. GG Rvs for d=150, p=0.5 (i.e., Log-normal), various values of n, and* $\varphi = 1$.

In Figures 1, 2, 3, 4, 5, and 6, the tightness of the approximate PDF with the simulated ones is clearly noticed across the entire range of z for $W(1,3)$, $GA(3,1)$ and $LN(150,0.5)$, respectively, $e^{-\varphi} = \{1,0.9\}$, and different values of $N = \{32,60\}$. It can be seen that as N increases, it becomes evident that more of the curves shift to the right. This phenomenon occurs because with a larger N, the corresponding μ_1 also increases, and leading to a concentration of the probability distribution around this higher average value.

4.1.2 Approximated CDF

Weibull distribution

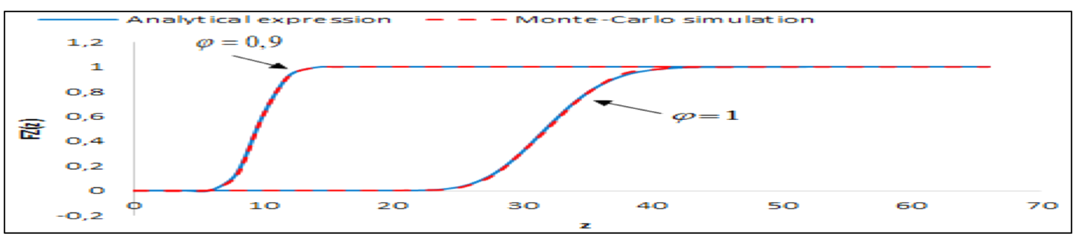


Fig. 7. *Sum CDF of i.n.i.d. GG RVs for d=1, p=3 (i.e., Weibull), n=32, and various values of φ.*

Gamma distribution

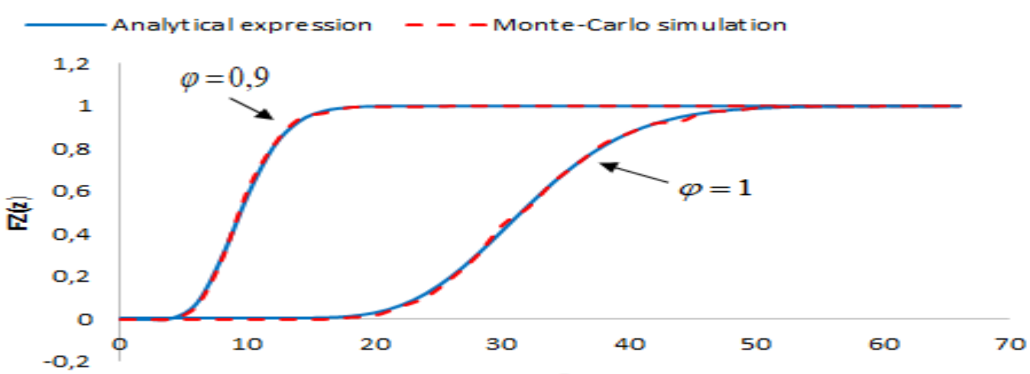


Fig. 8. *Sum CDF of i.n.i.d. GG RVs for d=3, p=1 (i.e., Gamma), n=32, and various values of* φ.

Figures 7 and 8 depict both the approximated and simulated CDFs computed for $n = 32$, $W(1,3)$, $GA(3,1)$, and $e^{-\varphi} = \{1; 0,9\}$. It is evident that the curves closely align across the entire range.

4.2 Model validation

As shown in the figures above the analytical method proposed in this context can significantly increase the quality of precision with a more accepted execution rate. However, to validate the model, it is necessary to test the speed of the execution time.

The following figures show the execution time of the approximated PDF for the analytical model and the Monte-Carlo simulation (on a processor Intel Core i7-4800MQ CPU @ 2.70 GHz).

Weibull standard times

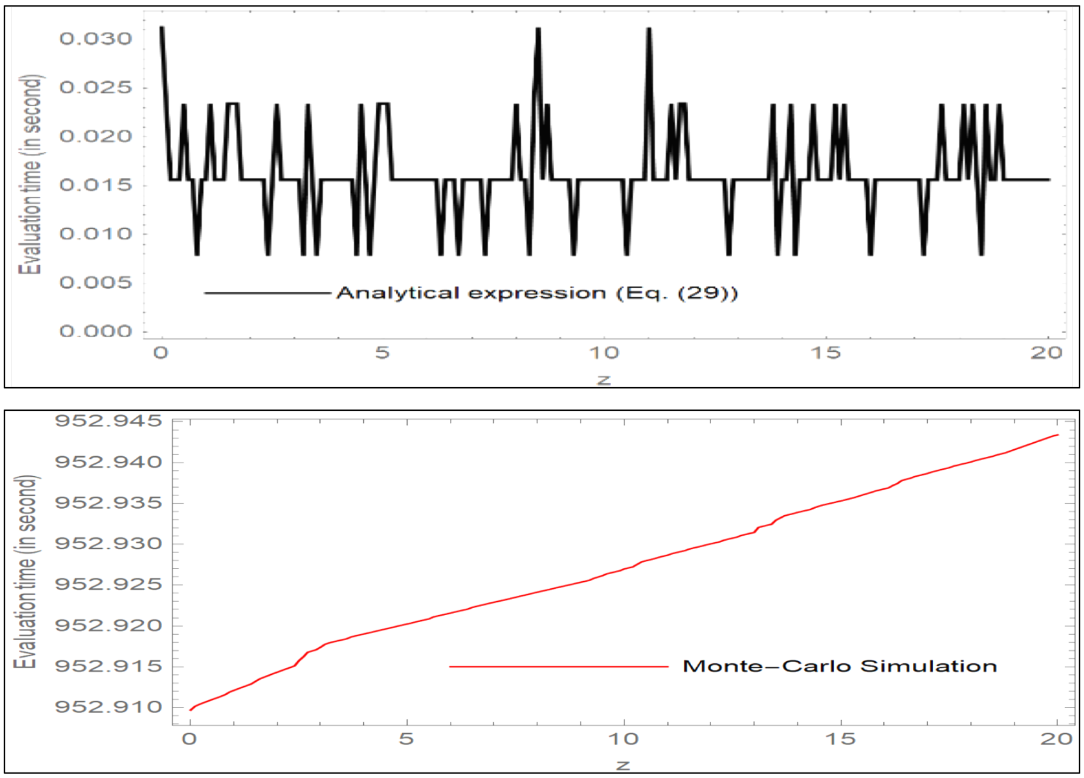
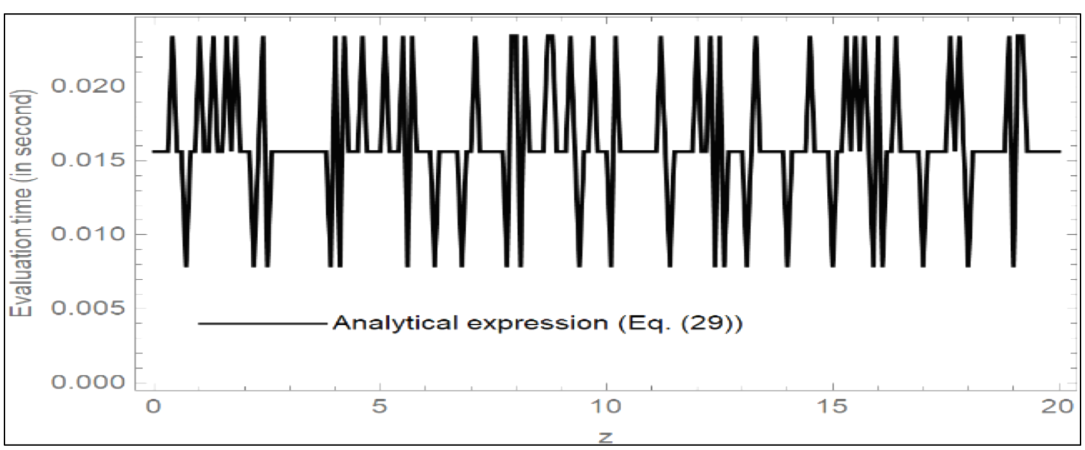


Fig. 9*. Evaluation time of the approximated and simulated PDF vs z, for W(1,3),* $e^{-\varphi} = 0.9$*, and n = 32.*

Gamma standard times



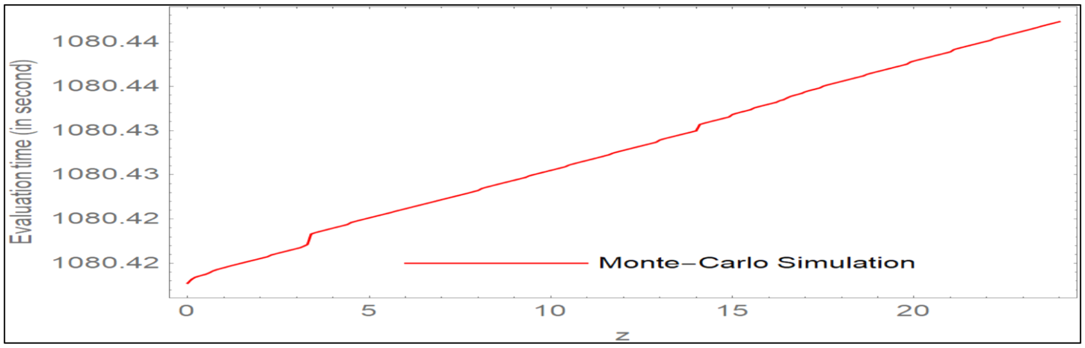


Fig. 10. *Evaluation time of the approximated and simulated PDF vs z, for GA(3,1),* $e^{-\varphi} = 0.9$*, and n = 32 .*

From Figures 9 and 10, we can easily notice that the execution time of the Monte-Carlo algorithm is proportional to the value of z, such that the greater the value of z, the greater the running time is greater. In addition, the Monte-Carlo algorithm takes several minutes to perform such a simulation. On the other hand, in the analytical model, the execution time does not respond to z. Indeed, the PDF estimate can be done in a second.

Therefore, it can be deduced that the execution of the PDF estimation for the analytical model is less restrictive compared to the Monte-Carlo algorithm. This provides an additional advantage when adopting the analytical process over the Monte-Carlo approach.

5. Conclusion

This work represents an analytical expression for PDF and CDF of the sum of independent GG RVs following a distribution described by Fox's H-function. The approach has proven to be effective, both in terms of simplicity (it is easy to understand and implement using statistical tools readily available in most data mining software.) and in terms of the results obtained compared to Monte Carlo simulations. It not only reduces execution time but also significantly improves precision.

The application can be extended to the whole field of activity of non-life insurance. For example, calculating the technical profit, the total cost and the probability of ruin of an insurance company.

Nonetheless, our research raises several questions that warrant exploration in future studies. For instance, it would be intriguing to investigate the distribution of the total claims amount in a collective model, where the claim amounts follow independent and identically distributed generalized gamma random variables and the number of claims is modeled as a Poisson random variable. Thus, the open nature of this research.

Conflicts of Interest Statement

The authors confirm that there are no known conflicts of interest or personal relationships that could have influenced the work presented in this paper.

Ethics Statement

This work did not require ethical approval.

References

[1] A. El Attar, M. El Hachloufi, and Y. Kouach, Automobile Takaful Pricing by Machine Learning Algorithms, *Recherches et Applications en Finance Islamique (RAFI)*, 6, 201–218 (2022).

[2] P.-E. Therond, Assurance non-vie, (cours), Institut de Science Financière et d'Assurances, (2005).

[3] H. H. Panjer, Recursive Evaluation of a Family of Compound Distributions, *ASTIN Bulletin: The Journal of the IAA*, 12, 22–26 (1981).

[4] J. Dhaene, G. Willmot, and B. Sundt, Recursions for distribution functions and stop-loss premiums, *Scandinavian Actuarial Journal*, 1999, 52–65 (1999).

[5] P. E. Heckman and G. G. Meyers, The Calculation of Aggregate Loss Distributions from Claim Severity and Claim Count Distributions, *Proceedings of the Casualty Actuarial Society*, 70, 133–134 (1983).

[6] C. Partrat and J.-L. Besson, *Assurance non-vie*. Economica (2004).

[7] R. Mnatsakanov and F. Ruymgaart, Some properties of moment-empirical cdf's with application to some inverse estimation problems, *Journal of Nonparametric Statistics*, 12, 478–495 (2003).

[8] H. Gzyl and A. Tagliani, Determination of the distribution of total loss from the fractional moments of its exponential, *Applied Mathematics and Computation*, 219, 2124–2133, (2012).

[9] Goffard, Approximations polynomiales de densités de probabilité et applications en assurance, Thèse de doctorat, Institut des Sciences Financières et d'Assurance,Université d'Aix-Marseille, France, (2015).

[10] T. Jin, S. B. Provost, and J. Ren, Moment-based density approximations for aggregate losses, *Scandinavian Actuarial Journal*, 2016, 216–245 (2016).

[11] S. Asmussen, P.-O. Goffard, and P. Laub, Orthonormal polynomial expansions and lognormal sum densities, in *In Risk and Stochastics,* 127–150, (2019).

[12] P.-O. Goffard and P. J. Laub, Orthogonal polynomial expansions to evaluate stop-loss premiums, *Journal of Computational and Applied Mathematics*, (2020).

[13] M. Denuit, A. Charpentier, and C. Bébéar, *Mathématiques de l'assurance non-vie : Tome 1, Principes fondamentaux de théorie du risque*. Economica, Paris, (2004).

[14] M. Denuit, Modèle collectif de théorie du risque, ISFA, (2003).

[15] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, Wiley-Interscience, New York, (1994).

[16] A. K. Rathie, A new generalization of generalized hypergeometric functions, *Le Matematiche*, 297–310 (1997).

[17] C. Fox, The G and H Functions as Symmetrical Fourier Kernels, *Transactions of the American Mathematical Society*, 98, 395–429 (1961).

[18] C. D. Bodenschatz, Finding an H-function distribution for the sum of independent H-function variates, Ph. D. thesis, University of Texas, Austin, (1992).

[19] H. Jacobs, J. W. Barnes, and I. D. Cook, Applications of the H-Function Distribution in Classifying & Fitting Classical Probability Distributions, *American Journal of Mathematical and Management Sciences*, 7, 131–147 (1987).

[20] M. D. Springer, *The algebra of random variables*, Wiley, New York, (1979).

[21] M. Abramowitz, I. A. Stegun, and D. Miller, Handbook of Mathematical Functions With Formulas, Graphs and Mathematical Tables, *Journal of Applied Mechanics*, 32, 239 (1965).

[22] S. W. Taqdees, Gengamrand (n, Alpha, Beta, a). MATLAB Central File Exchange, (2020).

[23] N. Kong and L. B. Milstein, SNR of generalized diversity selection combining with nonidentical Rayleigh fading statistics, *IEEE Transactions on Communications*, 48, 1266–1271 (2000).