Journal of Statistics & Management Systems
ISSN 0972-0510 (Print), ISSN 2169-0014 (Online)
Vol. 26 (2023), No. 4, pp. 885–896
DOI : 10.47974/JSMS-941

Tail probabilities of aggregate claim distribution when claims are weighted against Lindley distribution

K. Thiagarajah
Department of Mathematics
Illinois State University
Normal
IL 61790-4520
U.S.A.

Abstract

This paper proposes an approximation method for computing the tail probabilities of the aggregate claims distribution when claims follow a two-parameter weighted Lindley distribution. We have compared this approximation with some existing methods. Numerical comparisons are made with the exact results based on relative errors. The proposed approximation gives satisfactory accuracy in all applications considered in this paper.

Subject Classification: *(2010) 91B30, 62P05, 62E17.*

Keywords: *Frequency distribution, Lindley distribution, Aggregate claims distribution, Approximation methods.*

1. Introduction

The right tail probability of an aggregate claims distribution is of interest to insurers in determining the company's profit or loss. In the context of insurance theory, aggregate claims can be viewed as the sum of individual claim amounts for a random number of claim counts over a fixed time frame. In other words, it can be represented as the sum (S) of individual claim amounts $X_1, X_2, \ldots, X_N$, where N is the random number of claims over a fixed time. Conditional on $N = n$, the random variables $X_1, X_2, \ldots, X_n$ are assumed to be positive, mutually independent, and identically distributed. It is also assumed that the common distribution of

E-mail: **kthiaga@ilstu.edu**

$X_i s$ is independent of N. The (right) tail probability of aggregate claims random variable (S) can be written as:

$$SF(x) = P(S > x) = \sum_{n=0}^{\infty} P(N = n) P(X_1 + X_2 + \cdots + X_N > x \mid N = n). \tag{1.1}$$

The tail probability computation in (1.1) requires tedious numerical integrations for continuous severity distributions. [9] provided some (time-intensive) numerical procedures discussed by [13], which require a discretization of the continuous severity distribution. Several other approximation methods have been developed and studied by many authors in the literature ([16], [17], [20]). Here, we have considered only four: normal power, improved normal power, gamma, and IG approximations. The normal power approximation provided by [14] and the gamma approximation introduced by [2] are well-known to actuarial researchers. [15] Discussed an improvement to the Normal Power approximation. [3] Proposed the IG approximation and a mixture of Gamma and IG approximations. [11] Introduced an approximation method for computing the tail probability of a distribution, which requires only the first three moments. Several authors have applied this method in various situations ([1], [18], [19]). We propose this approximation, referred to as the LR approximation, for aggregate claim distribution when claim sizes follow two-parameter weighted Lindley. The Gamma approximation requires the numerical evaluation of incomplete gamma functions. The improved Normal Power, Gamma, and IG approximations are based on the lowest three moments equated with the distribution of interest. The Gamma-IG mixture approximation requires kurtosis of the distributions and the evaluation of the incomplete gamma function; all other approximations are based on the lowest three moments. Therefore, to make a meaningful comparison, we have omitted the Gamma-IG mixture approximation in the computation of relative errors.

The paper is organized as follows: Section 2 presents a brief description of the approximations mentioned above for computing the tail probability of the distribution of S. Section 3 compares the accuracy of these approximations numerically in terms of relative errors for some compound distributions when claim sizes follow a two-parameter weighted Lindley distribution with parameters α and θ. The relative error is computed as (approximate probability - exact probability) / exact probability. The exact probabilities are obtained through simulations. The final section contains the conclusion.

2. Approximations

In this section, we discuss five approximations for calculating the tail probability of the distribution of the aggregate claims. For the approximation methods, the required quantities such as mean, variance, and skewness of the distribution can easily be obtained from the cumulant generating function of *S*. Let us denote the cumulant generating function (*c.g.f*) as $C_S(t)$. Then, we can express the mean $= \mu_s = C_S'(0)$, the second central moment = variance = $\sigma_s^2 = \mu_2(s) = C_S''(0)$, and the third central moment = $\mu_3(s) = C_S'''(0)$.

2.1 *NP and Improved NP Approximations*

One can approximate the distribution of random sum, *S*, by a normal distribution, using the well-known central limit theorem. This may be appropriate for a large volume of risks. [14] provided the following expression, which is called Normal-Power (*NP*) approximation, as an adjustment to the normal approximation:

$$SF(x) \approx 1 - \Phi\left[\sqrt{1 + \frac{9}{\gamma_s^2} + \frac{6Z}{\gamma_s}} - \frac{3}{\gamma_s}\right], \tag{2.1}$$

where $Z = \frac{(x-\mu_s)}{\sigma_s}$ and γ_s is the skewness of S. [15] stated that the accuracy of the NP approximation can be improved by matching the first three moments. This would result in the following expression for the tail probability of *S*:

$$SF(x) \approx 1 - \Phi\left[\sqrt{1 + \frac{Z}{\delta_0} + \frac{c_0^2}{4\delta_0^2}} - \frac{c_0}{2\delta_0}\right], \tag{2.2}$$

where δ_0 is the unique root of the equation $\gamma_s = 6\delta - 4\delta^3$. The coefficient c_0 is defined as $c_0 = \sqrt{1 - 2\delta_0^2}$ for $|\delta| \leq 0.707$ and $|\gamma_s| \leq 2.828$. One may notice that if $c_0 = 1$ and $\delta_0 = \frac{\gamma_s}{6}$ we obtain the traditional NP approximation given in (2.1).

2.2 *Gamma Approximation*

The gamma approximation provided by [2] was based on the method of matching the lowest three moments of a gamma distribution to the distribution of interest. The approximate tail probability can be obtained as:

$$SF(x) \approx 1 - \frac{1}{\Gamma(\alpha)} \int_0^{\alpha + Z\sqrt{\alpha}} e^{-y} y^{\alpha-1} dy, \tag{2.3}$$

where $Z = \frac{(x-\mu_s)}{\sigma_s}$ and $\alpha = \frac{4}{\gamma_s^2}$. [6] claimed that this approximation provides satisfactory results when the claim size distribution is inverse Gaussian.

2.3 *IG Approximation*

[3] Developed this approximation by the method of matching the moments, as in the previous two methods. They approximated the random sum *S* by a shifted $IG(a,b)$ distribution. The approximate tail probability can be written as:

$$SF(x) \approx 1 - G_{a,b}(x - x_0), \tag{2.4}$$

where G is the cumulative probability of the shifted IG distribution. The parameters a and b can be written as $a = \frac{3[\mu_2(s)]^2}{\mu_3(s)}$, $b = \frac{\mu_3(s)}{3\mu_2(s)}$, and $x_0 = \mu_s - a$. The authors stated that this approximation is almost as good as the gamma approximation. [3] Also introduced another approximation which is a weighted average of gamma and IG approximations. The computation of the weights require the fourth moment of both approximate distributions.

2.4 *LR Approximation*

[11] Introduced a method based on saddlepoint technique. For continuous distributions of $X_i s$, the approximation to the tail probability of *S* can be written as:

$$SF(x) \approx = \begin{cases} \Phi(-\hat{w}) + \phi(\hat{w})\left\{\frac{1}{\hat{u}} - \frac{1}{\hat{w}}\right\}, & x \neq \mu \\ \frac{1}{2} - \frac{\gamma_s}{6\sqrt{2\pi}}, & x = \mu \end{cases} \tag{2.5}$$

where Φ and ϕ are the respective cumulative distribution function and the probability density function of a standard normal random variable, μ is the mean of the distribution, $\hat{w} = sgn(\hat{t})\sqrt{2\{\hat{t}x - C_S(\hat{t})\}}$, and $\hat{u} = \hat{t}\sqrt{C_S''(\hat{t})}$. The saddlepoint $\hat{t} = \hat{t}(x)$ is the unique solution to the equation $C_S'(\hat{t}) = x$. For more mathematical details of this approximation, we refer the readers to [11].

3. Compound Distributions

Let us discuss the Lindley distribution first. The Lindley distribution was originally introduced by [10] in the context of Bayesian statistics. Many generalizations of the Lindley distribution have been proposed over the last decade. Several authors have examined the properties of these distributions, including [4], [5], [7], [8], and [12]. [7] has discussed various properties of a weighted Lindley distribution and showed that its behavior is often similar to that of gamma, exponential, and Weibull in terms of a probability density function, hazard rate function, and mean residual function. The above three distributions are widely used for modeling severity claims in insurance. Lindley's distribution is positively skewed since the mean of the distribution is always greater than the mode. This property is also a desirable one in modeling insurance data. In this paper, we have considered the two-parameter weighted Lindley distribution discussed by [7] for modeling the severity claims. The claim counts distribution is taken as Poisson, negative binomial, or binomial. Our main focus is to assess the accuracy of the approximations discussed in Section 2 for computing the tail probabilities of aggregate claims. Now, let the random variable X follow a two-parameter weighted Lindley distribution with parameters α and θ. Then, the probability density function of X can be expressed ([7]) as:

$$f(x) \quad = \quad \frac{\theta^{\alpha+1}}{(\alpha+\theta)\Gamma(\alpha)} x^{\alpha-1}(1+x)e^{-\theta x}, x>0, \alpha>0, \theta>0. \tag{3.1}$$

3.1 *Poisson-Lindley Distribution* $(\lambda, \alpha, \theta)$:

In this example, we assume that the random variable N follows a Poisson distribution with parameter λ. Then the cumulant generating function of S is written as:

$$C_S(t) = \lambda\left[\frac{\theta^{\alpha+1}}{(\theta+\alpha)}\frac{(\theta+\alpha-t)}{(\theta-t)^{\alpha+1}} - 1\right], t<\theta \tag{3.2}$$

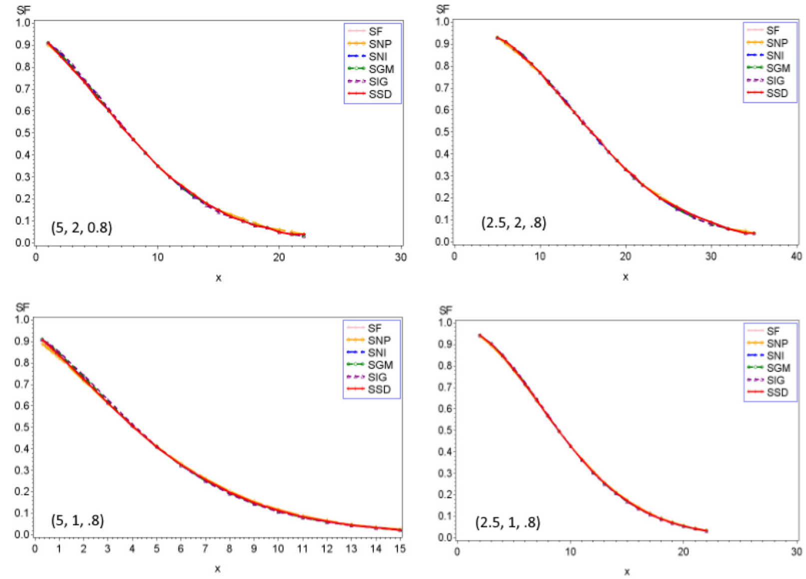

Figure 1

Tail Probabilities: PN-Lindley $(\lambda, \alpha, \theta)$

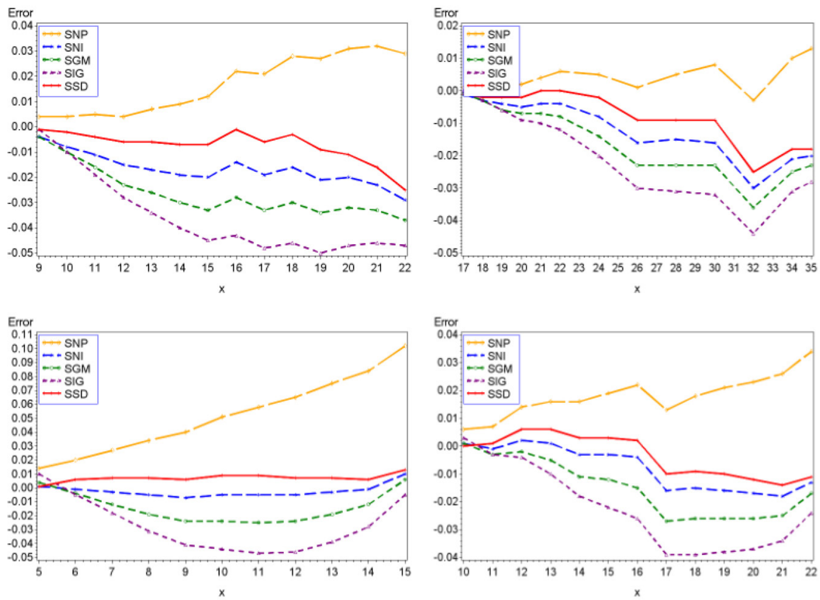

Figure 2

Relative Errors: PN-Lindley $(\lambda, \alpha, \theta)$

The saddlepoint $\hat{t}$ can be obtained by solving the equation $\lambda\delta(\theta+\alpha+1-\hat{t}) = x(\theta-\hat{t})^{\alpha+2}$, where $\delta = \left(\frac{\alpha}{\theta+\alpha}\right)\theta^{\alpha+1}$. From (3.2), we obtain the first three central moments which are given as:

$$\mu_s = \lambda\left[\frac{\alpha(\alpha+\theta+1)}{\theta(\theta+\alpha)}\right]$$

$$\mu_2(s) = \lambda\left[\frac{\alpha(\alpha+1)(\alpha+\theta+2)}{\theta^2(\theta+\alpha)}\right]$$

$$\mu_3(s) = \lambda\left[\frac{\alpha(\alpha+1)(\alpha+2)(\alpha+\theta+3)}{\theta^3(\theta+\alpha)}\right]$$

Figures 1 and 2 display the tail probabilities (SF) for different parameter combinations, and the relative errors for those combinations. For all cases, the approximation (2.5) gives excellent results.

3.2 *NB-Lindley Distribution* $(r, \beta, \alpha, \theta)$:

In this example, we assume that the distribution of the random variable N is a negative binomial with parameters r and β. The *c.g.f.* of S can be written as:

$$C_S(t) = -r\, ln\left[1-\beta\left(\frac{\eta(\theta+\alpha-t)}{(\theta-t)^{\alpha+1}}-1\right)\right], \tag{3.3}$$

where $\eta = \frac{\theta^{\alpha+1}}{(\theta+\alpha)}$. The saddlepoint $\hat{t}$ is the solution of

$$\frac{r\beta\eta\alpha(\theta+\alpha+1-\hat{t})}{1-\beta\left(\frac{\eta(\theta+\alpha-t)}{(\theta-t)^{\alpha+1}}-1\right)} = x(\theta-\hat{t})^{(\alpha+2)}. \tag{3.4}$$

Function in (3.3) yields the following moments:

$$\mu_s = \frac{r\beta\alpha}{\theta(\theta+\alpha)}(\theta+\alpha+1)$$

$$\mu_2(s) = \frac{r\beta\alpha}{\theta^2(\theta+\alpha)^2}[\Delta_1+\beta\alpha(\theta+\alpha+1)^2]$$

$$\mu_3(s) = \frac{r\beta\alpha}{\theta^3(\theta+\alpha)^3}[\Delta_2+3\beta\Delta_3+2\beta^2\alpha^2(\theta+\alpha+1)^3],$$

where $\Delta_1 = (\alpha+1)(\theta+\alpha)(\theta+\alpha+2)$, $\Delta_2 = (\alpha+1)(\alpha+2)(\theta+\alpha)^2(\theta+\alpha+3)$ and $\Delta_3 = \alpha(\alpha+1)(\theta+\alpha)(\theta+\alpha+1)(\theta+\alpha+2)$. Figures 3 and 4 present the tail probabilities and the corresponding relative errors. The outcomes in Figures reveal that the LR approximation given in (2.5) outperforms the other approximations.

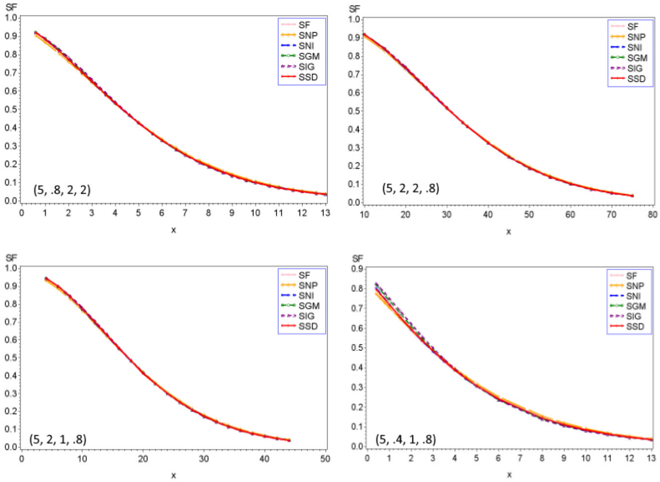

Figure 3

Tail Probabilities: NB-Lindley $(r, \beta, \alpha, \theta)$

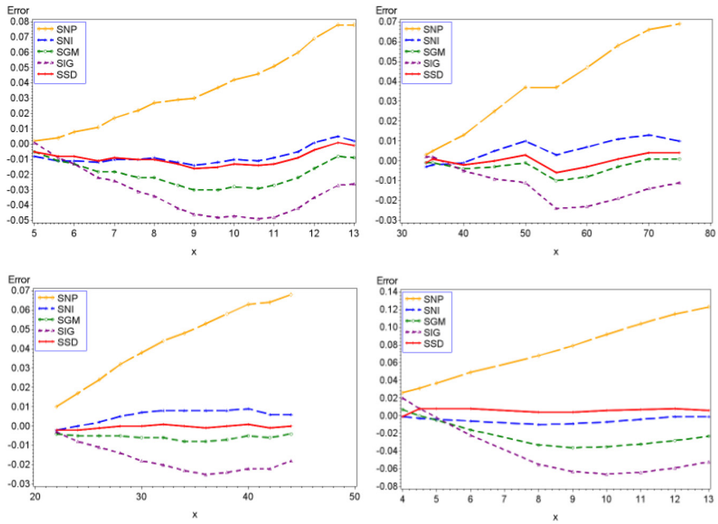

Figure 4

Relative Errors: NB-Lindley $(r, \beta, \alpha, \theta)$

3.3 *Binomial-Lindley Distribution* (m, q, α, θ):

In this example, the random variable N follows a binomial distribution with parameters m and q. The *c.g.f.* of S is given as:

$$C_S(t) = m\, ln\left[1 + q\left(\frac{\eta(\theta + \alpha - t)}{(\theta - t)^{\alpha+1}} - 1\right)\right]. \tag{3.5}$$

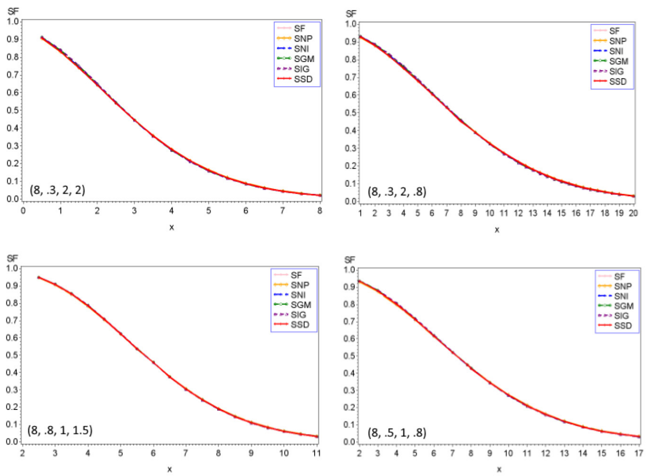

Figure 5

Tail Probabilities: BM-Lindley (m, q, α, θ)

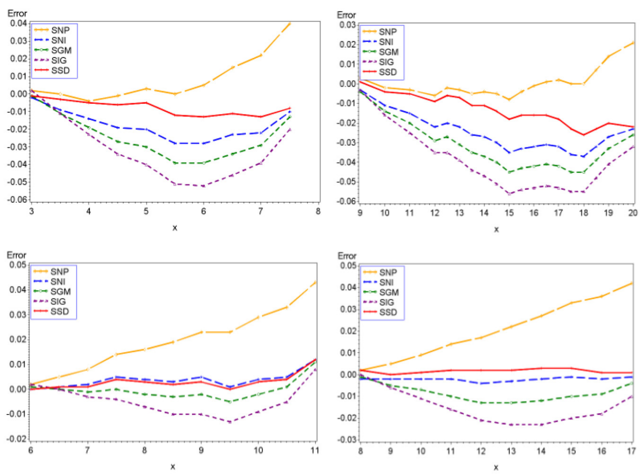

Figure 6

Relative Errors: BM-Lindley (m, q, α, θ)

The saddlepoint $\hat{t}$ can be obtained numerically from

$$\frac{mq\eta(\theta+\alpha+1-\hat{t})(\theta-\hat{t})^{-(\alpha+2)}}{1+q\left(\frac{\eta(\theta+\alpha-t)}{(\theta-t)^{\alpha+1}}-1\right)}=\frac{x}{\alpha}. \tag{3.6}$$

The required moments obtained from (3.5) are as follows:

$$\mu_s=\frac{mq\alpha}{\theta(\theta+\alpha)}(\theta+\alpha+1)$$

$$\mu_2(s)=\frac{mq\alpha}{\theta^2(\theta+\alpha)^2}[\Delta_1-q\alpha(\theta+\alpha+1)^2]$$

$$\mu_3(s)=\frac{mq\alpha}{\theta^3(\theta+\alpha)^3}[\Delta_2+2q^2\alpha^2(\theta+\alpha+1)^3-3q\Delta_3],$$

where Δ_1, Δ_2 and Δ_3 are defined as in subsection 3.2. Figures 5 and 6 depict the tail probabilities and the corresponding relative errors for various combinations of parameter values. The outcomes are the same as in the previous two applications.

4. Conclusion

The purpose of this paper is to assess the accuracy of the approximation, introduced by [11], given in 2.5 with the other four approximations discussed in Section 2 for computing the tail probabilities of the distribution of aggregate claims. The approximate probabilities have been calculated for various compound distributions and compared with the exact values in terms of relative errors. Based on the results in Figures 1 to 6, one can state that the LR approximation controls the relative errors uniformly over a whole range of values of S, followed by the improved NP approximation in all examples considered. The Gamma and the IG approximations behave similarly. The enhanced NP approximation given in (2.2) produces much smaller errors than the traditional NP approximation provided in (2.1). Suppose the claim amounts follow a two-parameter weighted Lindley distribution. In that case, the LR approximation or the improved normal power approximation is simple and easy to compute the tail probabilities with great accuracy, requiring only the lowest three moments of the random variable S.

References

[1] Alhejaili, A. D. and Abd-Elfattah, E. F. Saddlepoint Approximations for Stopped-Sum Distributions. *Communications in Statistics: Theory and Methods*, 42: 3735-3743 (2013).

[2] Bohman, H. and Esscher, F. Studies in Risk Theory with Numerical Illustrations. *Scandinavian Actuarial Journal*, 46: 173-225 (1963).

[3] Chaubey, Y. P., Garrido, J. and Trudeau, S. On the Computation of Aggregate Claims Distributions: Some New Approximations. *Insurance: Mathematics and Economics*, 23: 215-230 (1998).

[4] Chesneau, C., Tomy, L. and Jose, M. Wrapped Modified Lindley Distribution. *Journal of Statistics and Management Systems*, 24: 1025-1040 (2021).

[5] Chesneau, C., Tomy, L. and Gillariose, J. A New Modified Lindley Distribution with Properties and Applications. *Journal of Statistics and Management Systems*, 24: 1383-1403 (2021).

[6] Gendron, M. and Crepeau, H. On the Computation of the aggregate Claim Distribution when claims are inverse Gaussian. *Insurance: Mathematics and Economics*, 8: 251-258 (1989).

[7] Ghitany, M. E., Alqallaf, F., Al-Mutairi, D. K. and Husain, H. A. A two-parameter weighted Lindley distribution and its applications to survival data. *Mathematics and Computers in Simulation*, 81: 1190-1201 (2011).

[8] Ghitany, M. E., Song, P. and Wang, S. New Modified Moment Estimators for the Two-Parameter Weighted Lindley Distribution. *Journal of Statistical Computation and Simulation*, 87: 3225-3240 (2017).

[9] Klugman, S. A., Panjer, H. H. and Willmot, G. E. *Loss Models: From Data to Decisions*, John Wiley, New York (2012).

[10] Lindley, D. V. Fiducial distributions and Bayes' theorem. *Journal of the Royal Statistical Society: Series A*, 20: 102-107 (1958).

[11] Lugannani, R. and Rice, S. O. Saddlepoint Approximations for the Distribution of the Sum of Independent Random Variables. *Adv. Appl. Probab.*, 12: 475-490 (1980).

[12] Mazucheli, J., Louzada, F. and Ghitany, M. E. Comparison of Estimation Methods for the Parameters of the Weighted Lindley Distribution. *Applied Mathematics and Computation*, 220: 463-471 (2013).

[13] Panjer, H. H. Recursive Evaluation of a Family of Compound Distributions. *ASTIN Bulletin*, 12: 22-26 (1981).

[14] Pesonen, E. NP-Approximation of Risk Processes. *Scandinavian Actuarial Journal*, 3: 63-69 (1969).

[15] Ramsay, C. M. A Note on the Normal Power Approximation. *ASTIN Bulletin*, 21: 147-150 (1991).

[16] Reijnen, R., Albers, W. and Kallenberg, W. C. M. Approximations for stop-loss re-insurance premiums. *Insurance: Mathematics and Economics*, 36: 237-250 (2005).

[17] Seri, R. and Choirat, C. Comparison of approximations for compound Poisson processes. *ASTIN Bulletin*, 45: 601-637 (2015).

[18] Thiagarajah, K. Approximate Tail Probabilities of Aggregate Claims. *International Journal of Statistics and Economics*, 18: 1-11 (2017).

[19] Thiagarajah, K. Comparison: Some Approximations for the Aggregate Claims Distributions. *Variance*, 13: 190-206 (2021).

[20] Thiagarajah, K. Some Approximate Tail Probabilities for a Compound Logarithmic Distribution. *International Journal of Statistics and Economics* 23 : 1-10 (2022).

Revised September, 2021

Revised March, 2022