DOI: https://doi.org/10.24297/jam.v22i.9528

On Sum and Geometric Sum of independent New Quasi Lindley Random Variables and its Applications

Alaa R. El-Alosey

Department of Mathematics, Faculty of Science, Tanta University, Tanta, Egypt.

E-mail: alaa.abdelmaksoud@science.tanta.edu.eg

ORCID ID: https://orcid.org/0000-0002-3177-2420

Abstract;

The Laplace transformation method is used to drive the distribution of the sum S_n of n-fixed random variables, which has a new quasi Lindley distribution with two parameters θ and α, NQLD (θ,α). The sum of NQLD (SUNQLD) distribution is obtained in pdf and cdf formats. It is discussed how to calculate the random sum S_N of a random number of NQLD random variables. The random sum of the NQLD distribution's pdf and cdf are calculated. When N has a geometric distribution, the geometric sum of NQLD distribution (GSN QLD) was introduced as an example of a random number of NQLD random variables. For all cases, some statistical measures are determined. The distribution's parameters are estimated using the maximum likelihood method. To test the viability and efficiency of the proposed distributions SNQLD and GSNQLD, lifetime count data sets from acute myeloid leukemia are fitted. The results should become accepted knowledge in the fields of probability theory and its allied sciences. In addition, the histogram, fitted probability density function (pdf), and P-P plots for the analyzed real data set are presented.

Keywords: Random sum; Geometric sum; Sum of random variables; New quasi-Lindley distribution (NQLD); Laplace transformation; Moments.

Mathematics subject classification: 60E05, 62E10.

1. Introduction

The random sum model is one of the standard stochastic models used in several areas of applied probability, such as insurance risk theory and queuing theory.. Bon and Kalashnikov [3] looked at certain bounds for geometric sums of random positive variables, as well as some bounds for left-truncated random positive variables. The calculation of the n-fold convolution of generalized exponential-sum distribution functions was explored by Ma and King [13]. The random sum of exponential distribution mixes was explored by Teamah and El-Alosey [22]. The distributional features of random sum were addressed by Willmot and Cai [23]. The probability generating function of general Bernoulli distributions was discovered by Kolev et al. [11]. For the weighted sum of independent random variables, Sung [21] developed novel complete convergence results. The probability density function of a random sum of mid truncated Lindley distribution was derived by Mohie El-Din et al. [15]. The sums of independent generalized Pareto random variables were studied by Nadarajah et al. [16] with applications to insurance and cat bonds. Girondot and Barry [9] focused on how the distribution of the sum of negative binomial random variables plays a unique function in actuarial science, ecology, and insurance mathematics. With the aid of a quantile-based representation, Midhu et al. [14] were able to approximate the distribution function of the sum of two independent random variables and produce a non-parametric estimator for this distribution function.

Lindley [12] devised a one-parameter Lindley distribution for analyzing lifetime data. Ghitany et al. [8] studied Lindley distribution while taking into account the time it takes for banking customers to be served. The Quasi Lindley distribution was developed by Shanker and Mishra [18] (QLD). Shanker and Mishra [19] developed a two-parameter Lindley distribution and examined its attributes. Another two-parameter Lindley distribution was introduced by Shanker et al. [20]. A novel Quasi Lindley distribution was designed by Shanker and Amanuel [17]. Ganaie et al. [6] proposed a weighted new quasi lindley distribution, which is a new generalization of the two-parameter new quasi Lindley distribution. Chesneau et al. [4] used the same parameter to treat the difference of two independent random variables following the Lindley distribution and the sum of two independent random variables following the Lindley distribution. As a special example of weighted distribution, Ganaie et al. [7] presented a new form of the new quasi Lindley distribution known as the length-biased weighted new quasi Lindley distribution. Amer et al. [2] used the same parameter to introduce the sum and difference of two independent random variables following the Quasi-Lindley distribution. Hassan and Abd-Allah [10] obtained the power quasi Lindley power series by compounding the power quasi Lindley and truncated power series distributions.

Lindley [12] proposed a single-parameter Lindley distribution, with a probability density function (pdf) of

$$f(x;\theta)=\frac{\theta^2}{\theta+1}(1+x)e^{-\theta x} \quad ,x>0,\theta>0 \tag{1}$$

Shanker and Mishra [18] proposed a Quasi Lindley distribution (QLD) with parameters, with the (p.d.f) described by

$$f(x;\theta,\alpha)=\frac{\theta}{\alpha+1}(\alpha+\theta x)e^{-\theta x}\quad ,x>0,\theta>0,\alpha>-1 \qquad (2)$$

In addition, Shanker and Amanuel [17] proposed a new Quasi Lindley distribution (NQLD) with parameters, and its pdf is defined by

$$f(x;\theta,\alpha)=\frac{\theta^2}{\theta^2+\alpha}(\theta+\alpha x)e^{-\theta x}\quad ,x>0,\theta>0,\alpha>-\theta^2 \qquad (3)$$

Which is a mixture of gamma $(2,\theta)$ and exponential distributions with parameter θ.

For a continuous random variable X with pdf $f(x)$, the Laplace transformation of pdf $f(x)$ is defined as:

$$L[f(x)]=f_X^*(t)=E(e^{-tx})$$

The corresponding Laplace transformation of the NQLD is given by

$$L\{f(x)\}=f_X^*(t)=\int_0^\infty \frac{\theta^2}{\theta^2+\alpha}(\theta+\alpha x)e^{-(\theta+t)x}\,dx$$
$$=\frac{\theta^2(\theta^2+\alpha+\theta t)}{(\theta^2+\alpha)(\theta+t)^2} \qquad (4)$$

2. Sum of new Quasi Lindley random variables

In this section we drove the pdf and the cdf of a fixed number n, n is a positive integer number, of the sum, $S_n=\sum_{i=1}^n X_i$, of NQLD random variables (SNQLD). We calculate some statistical measures of S_n.

2.1. The pdf and the cdf of the SNQLD random variables

In the following theorem drove the pdf and the cdf of S_n by using the Laplace transformation method.

Theorem 2.1

Let $X_i, i=1,2,\dots,n$ be independent and identically distributed random variables

having the NQLD distribution, the pdf of the sum $S_n=\sum_{i=1}^n X_i$ is given by:

$$f_{S_n}(s)=\sum_{i=0}^n \binom{n}{i}\frac{\alpha^i\theta^{3n-i}}{(\theta^2+\alpha)^n}\frac{e^{-\theta s}\,s^{n+i-1}}{\Gamma(n+i)}\ ,s>0,\theta>0,\alpha>-\theta^2$$

and the cdf of the sum $S_n=\sum_{i=1}^n X_i$ is:

$$F_{S_n}(u)=\sum_{i=0}^n \binom{n}{i}\frac{\alpha^i\theta^{2(n-i)}}{\Gamma(n+i)(\theta^2+\alpha)^n}\gamma(n+i,u)$$

where $\gamma(n+i,u)=\int_0^u \theta^{n+i}e^{-\theta s}\,s^{n+i-1}ds$, which is incomplete gamma function.

Proof.

Let $X_i, i=1,2,\dots,n$ be independent and identically distributed random variables having the NQLD distribution with Laplace transformation given in relation(4)

Let $S_n=\sum_{i=1}^n X_i$, where n is a non-negative integer number. Then, the Laplace

$$L[f_{S_n}(s)]=f_{S_n}^*(t)=\prod_{i=1}^n f_{X_i}^*(t)=\{f_{X_i}^*(t)\}^n=\left[\frac{\theta^2(\theta^2+\alpha+\theta t)}{(\theta^2+\alpha)(\theta+t)^2}\right]^n$$
$$=\frac{\theta^{2n}}{(\theta^2+\alpha)^n}\left[\frac{\alpha+\theta(\theta+t)}{(\theta+t)^2}\right]^n=\frac{\theta^{2n}}{(\theta^2+\alpha)^n}\left[\frac{\alpha}{(\theta+t)^2}+\frac{\theta}{\theta+t}\right]^n$$
$$=\frac{\theta^{2n}}{(\theta^2+\alpha)^n}\left[\sum_{i=0}^n \binom{n}{i}\frac{\alpha^i}{(\theta+t)^{2i}}\frac{\theta^{n-i}}{(\theta+t)^{n-i}}\right]$$

where $\binom{n}{i}=\frac{n!}{i!(n-i)!}$

Thus the Laplace transformation of S_n is

$$f_{S_n}^*(t)=\frac{\theta^{2n}}{(\theta^2+\alpha)^n}\sum_{i=0}^n \binom{n}{i}\alpha^i\theta^{n-i}\frac{1}{(\theta+t)^{n+i}} \qquad (5)$$

By using the inverse Laplace transformation in relation (5), the pdf of S_n is given by

$$f_{S_n}(s) = L^{-1}\left[f^*_{S_n}(t)\right]$$

therefore,

$$f_{S_n}(s) = \sum_{i=0}^{n} \binom{n}{i} \frac{\alpha^i \theta^{3n-i}}{(\theta^2+\alpha)^n} \frac{e^{-\theta s}\, s^{n+i-1}}{\Gamma(n+i)} \quad , s > 0, \theta > 0, \alpha > -\theta^2 \tag{6}$$

The cdf of the sum S_n can be calculated by the relation

$$F_{S_n}(u) = \int_0^u f_{S_n}(s)\, ds$$

therefore,

$$F_{S_n}(u) = \sum_{i=0}^{n} \binom{n}{i} \frac{\alpha^i \theta^{3n-i}}{\Gamma(n+i)(\theta^2+\alpha)^n} \int_0^u e^{-\theta s}\, s^{n+i-1} ds$$

$$= \sum_{i=0}^{n} \binom{n}{i} \frac{\alpha^i \theta^{2(n-i)}}{\Gamma(n+i)(\theta^2+\alpha)^n} \gamma(n+i, u) \tag{7}$$

where $\gamma(n+i, u) = \int_0^u \theta^{n+i} e^{-\theta s}\, s^{n+i-1} ds$, which is incomplete gamma function.

2.2. The moments

The r^th moment about the origin of the sum S_n has been obtained as

$$\mu^*_r = E(s^r) = \sum_{i=0}^{n} \binom{n}{i} \frac{\alpha^i \theta^{3n-i}}{\Gamma(n+i)(\theta^2+\alpha)^n} \int_0^\infty e^{-\theta s}\, s^{n+i+r-1} ds$$

$$= \sum_{i=0}^{n} \binom{n}{i} \frac{\Gamma(n+i+r)\alpha^i \theta^{2(n-i)-r}}{\Gamma(n+i)(\theta^2+\alpha)^n}$$

Therefore, the mean is defined when $r = 1$ as

$$\mu = \mu^*_1 = E(S) = \sum_{i=0}^{n} \binom{n}{i} \frac{(n+i)!\ \alpha^i \theta^{2(n-i)-1}}{(n+i-1)!\,(\theta^2+\alpha)^n}$$

$$= \frac{n\,(\theta^2+2\alpha)}{\theta(\theta^2+\alpha)} \tag{8}$$

The second moment , when $r = 2$, about the origin is

$$\mu^*_2 = E(S^2) = \sum_{i=0}^{n} \binom{n}{i} \frac{\Gamma(n+i+2)\alpha^i \theta^{2(n-i)-2}}{\Gamma(n+i)(\theta^2+\alpha)^n}$$

$$= \frac{2n\alpha^2 + 4n^2\alpha\theta^2 + n\theta^4 + 4n^2\alpha^2 + 4n\alpha\theta^2 + n^2\theta^4}{\theta^2(\theta^2+\alpha)^2}$$

Thus the variance of the sum S_n is given by

$$V(S) = \mu^*_2 - (\mu^*_1)^2 = \frac{2n\alpha^2 + 4n^2\alpha\theta^2 + n\theta^4}{\theta^2(\theta^2+\alpha)^2} \tag{9}$$

By using equations (8) and (9) the mean and variance of the SNQLD distributions for different combinations of $\alpha, \theta,$ and n are computed in Table 1. It can be seen that both the mean and the variance diminishing by expanding the parameter θ but when the parameter αincrements the mean diminishes and the variance isn't as in Table 1.

Table 1 : Mean and Variance of the SNQLD distributions for different values

n	α / θ	2 Mean	2 Variance	10 Mean	10 Variance	20 Mean	20 Variance
2	2	1.33333	1.222222	0.34286	1.367347	0.18333	1.263889
	10	1.01961	0.022314	0.21818	0.030083	0.11667	0.037222
	20	1.00498	0.005149	0.20488	0.005717	0.10476	0.006372
10	2	6.66667	23.88889	1.71429	23.16327	0.91667	17.43056
	10	5.09804	0.173087	1.09091	0.414876	0.58333	0.630556
	20	5.02488	0.029703	1.02439	0.04762	0.52381	0.068141
50	2	33.3333	563.8889	8.5714	523.9796	4.5833	364.9306
	10	25.4902	2.403306	5.4545	8.68595	2.9167	14.26389
	20	25.1244	0.247525	5.1220	0.71401	2.6190	1.247732

2.3. The moment generating function

The moment generating function of the sum S_n is given by

$$M_{S_n}(t) = E(e^{ts}) = \sum_{i=0}^{n} \binom{n}{i} \frac{\alpha^i \theta^{3n-i}}{\Gamma(n+i)(\theta^2+\alpha)^n} \int_0^\infty e^{-(\theta-t)s}\, s^{n+i-1} ds$$

$$= \frac{\theta^{2n}(\alpha+\theta(\theta-t))^n}{(\theta^2+\alpha)^n(\theta-t)^{2n}} \quad (10)$$

The mean (first moment) of the SNQLD distribution can be calculated using eq.(10) as follows:

$$\mu_1 = \mu = E(X) = \frac{dM_X(t)}{dt}|_{t=0} = \frac{n\,(\theta^2+2\alpha)}{\theta(\theta^2+\alpha)}$$

The 2nd moment about the origin is

$$\mu_2 = E(X^2) = \frac{dM_X^2(t)}{dt^2}|_{t=0} = \frac{2n\alpha^2+4n^2\alpha\theta^2+n\theta^4+4n^2\alpha^2+4n\alpha\theta^2+n^2\theta^4}{\theta^2(\theta^2+\alpha)^2}$$

As a result the GSNQLD distribution's variance is given by

$$\sigma^2 = \mu_2 - \mu_1^2 = \frac{2n\alpha^2+4n^2\alpha\theta^2+n\theta^4}{\theta^2(\theta^2+\alpha)^2}$$

3. Random Sum of new Quasi Lindley random variables.

In this section, we drove the pdf and the cdf of the random sum, $S_N = \sum_{i=1}^{N} X_i$, where *N* is a non-negative discrete random variable, of NQLD random variables. We calculate some of the statistical measures of S_N.

3.1. The pdf and the cdf of the random sum of NQLD random variables

The following theorem drove the pdf and the cdf of *N*, where *N* is a non-negative discrete random variable, of NQLD random variables by using the Laplace transformation method.

Theorem 3.1

Let $X_i, i = 1,2,\dots,N$, where N is a non-negative discrete random variable, be independent and identically distributed random variables having the NQLD distribution, the pdf of the random sum $S_N = \sum_{i=1}^{N} X_i, S_0 = 0$ is

$$f_{S_N}(s) = \sum_{n=0}^{\infty} \sum_{i=0}^{n} g(n) \binom{n}{i} \frac{\alpha^i \theta^{3n-i}}{(\theta^2+\alpha)^n} \frac{e^{-\theta s}\, s^{n+i-1}}{\Gamma(n+i)}, s>0, \theta>0, \alpha>-\theta^2$$

and the cdf of the sum $S_N = \sum_{i=1}^{N} X_i$ is:

$$F_{S_N}(u) = \sum_{n=0}^{\infty} \sum_{i=0}^{n} g(n) \binom{n}{i} \frac{\alpha^i \theta^{2(n-i)}}{\Gamma(n+i)(\theta^2+\alpha)^n} \gamma(n+i,u), u>0, \theta>0, \alpha>-\theta^2$$

where $\gamma(n+i,u) = \int_0^u \theta^{n+i} e^{-\theta s}\, s^{n+i-1} ds$, which is incomplete gamma function.

Proof.

Let $X_i, i = 1,2, \dots, N$, be independent and identically distributed random variables having the NQLD distribution with Laplace transform given in relation (4)

Let $S_N = \sum_{i=1}^{N} X_i$, where N is a non-negative discrete random variable with probability mass function $g(n) = P(N = n)$, and probability generating function of N is given by

$$P_N(\theta) = \sum_{n=0}^{\infty} \theta^n g(n)$$

Then, the Laplace transformation of the random sum S_N is

$$L\{f_{S_N}(s)\} = f_{S_N}^*(t) = P_N\{f_{X_i}^*(t)\}$$

$$= \sum_{n=0}^{\infty} \{f_{X_i}^*(t)\}^n g(n)$$

$$= \sum_{n=0}^{\infty} \left\{\frac{\theta^2(\theta^2 + \alpha + \theta t)}{(\theta^2 + \alpha)(\theta + t)^2}\right\}^n g(n)$$

$$= \sum_{n=0}^{\infty} \frac{g(n)\theta^{2n}}{(\theta^2 + \alpha)^n} \sum_{i=0}^{n} \binom{n}{i} \alpha^i \theta^{n-i} \frac{1}{(\theta + t)^{n+i}}$$

Thus the Laplace transformation of S_N is

$$f_{S_N}^*(t) = \sum_{n=0}^{\infty} \sum_{i=0}^{n} \frac{g(n)\theta^{2n}}{(\theta^2 + \alpha)^n} \binom{n}{i} \alpha^i \theta^{n-i} \frac{1}{(\theta + t)^{n+i}} \qquad (11)$$

By using the inverse Laplace transformation in relation (11), the p.d.f of S_N is given by

$$f_{S_N}(s) = L^{-1}\{ f_{S_N}^*(t)\}$$

therefore,

$$f_{S_N}(s) = \sum_{n=0}^{\infty} \sum_{i=0}^{n} g(n) \binom{n}{i} \frac{\alpha^i \theta^{3n-i}}{(\theta^2 + \alpha)^n} \frac{e^{-\theta s} s^{n+i-1}}{\Gamma(n+i)}, s > 0, \theta > 0, \alpha > -\theta^2 \qquad (12)$$

The cdf of the sum S_N can be calculated by the relation

$$F_{S_N}(u) = \int_0^u f_{S_N}(s)\, ds$$

therefore,

$$F_{S_n}(u) = \sum_{n=0}^{\infty} \sum_{i=0}^{n} g(n) \binom{n}{i} \frac{\alpha^i \theta^{3n-i}}{\Gamma(n+i)(\theta^2 + \alpha)^n} \int_0^u e^{-\theta s} s^{n+i-1} ds$$

$$= \sum_{n=0}^{\infty} \sum_{i=0}^{n} g(n) \binom{n}{i} \frac{\alpha^i \theta^{2(n-i)}}{\Gamma(n+i)(\theta^2 + \alpha)^n} \gamma(n+i, u) \qquad (13)$$

where $u > 0, \theta > 0, \alpha > -\theta^2$ and $\gamma(n+i, u) = \int_0^u \theta^{n+i} e^{-\theta s} s^{n+i-1} ds$, which is incomplete gamma function.

3.2. The moments

The r[th] moment about origin of the sum S_N has been obtained as

$$\mu_r^* = E(s^r) = \sum_{n=0}^{\infty} g(n) \sum_{i=0}^{n} \binom{n}{i} \frac{\alpha^i \theta^{3n-i}}{\Gamma(n+i)(\theta^2 + \alpha)^n} \int_0^{\infty} e^{-\theta s} s^{n+i+r-1} ds$$

$$= \sum_{n=0}^{\infty} g(n) \sum_{i=0}^{n} \binom{n}{i} \frac{\Gamma(n+i+r)\alpha^i \theta^{2(n-i)-r}}{\Gamma(n+i)(\theta^2 + \alpha)^n}$$

Therefore, the mean is defined when $r = 1$ as

$$\mu = \mu_1^* = \sum_{n=0}^{\infty} g(n) \sum_{i=0}^{n} \binom{n}{i} \frac{(n+i)!\ \alpha^i \theta^{2(n-i)-1}}{(n+i-1)!\,(\theta^2+\alpha)^n}$$
$$= \sum_{n=0}^{\infty} n\, g(n) \frac{(\theta^2+2\alpha)}{\theta(\theta^2+\alpha)} = \frac{(\theta^2+2\alpha)}{\theta(\theta^2+\alpha)} E(N)$$

3.3. The moment generating function

The moment generating function of the random sum S_N is given by

$$M_{S_N}(t) = E(e^{ts}) = \sum_{n=0}^{\infty} g(n) \sum_{i=0}^{n} \binom{n}{i} \frac{\alpha^i \theta^{3n-i}}{\Gamma(n+i)(\theta^2+\alpha)^n} \int_0^{\infty} e^{-(\theta-t)s} s^{n+i-1} ds$$
$$= \sum_{n=0}^{\infty} g(n) \frac{\theta^{2n}(\alpha+\theta(\theta-t))^n}{(\theta^2+\alpha)^n(\theta-t)^{2n}}$$

4. Geometric Sum of new Quasi Lindley random variables.

In this section, we introduce geometric sum of NQLD distribution, GSNQLD as an example of the random sum of NQLD distribution

Let $X_i, i = 1,2,\dots,N,$ be independent and identically distributed random variables having the NQLD distribution, and N has a geometric distribution with parameter p and probability mass function

$$g(n;p) = p(1-p)^{n-1}, \qquad n = 1,2,\dots, 0 \le p \le 1$$

By using relations (9) & (10), the pdf of the GSNQLD random variables is given by

$$f_{S_N}(s) = \sum_{n=1}^{\infty} \sum_{i=0}^{n} p(1-p)^{n-1} \binom{n}{i} \frac{\alpha^i \theta^{3n-i}}{(\theta^2+\alpha)^n} \frac{e^{-\theta s} s^{n+i-1}}{\Gamma(n+i)} \tag{14}$$

where $s > 0, \theta > 0, \alpha > \theta^2, 0 < p < 1$

also the cdf of the GSNQLD random variables is given by

$$F_{S_N}(u) = \sum_{n=1}^{\infty} \sum_{i=0}^{n} p(1-p)^{n-1} \binom{n}{i} \frac{\alpha^i \theta^{2(n-i)}}{\Gamma(n+i)(\theta^2+\alpha)^n} \gamma(n+i,u) \tag{15}$$

where $u > 0, \theta > 0, \alpha > -\theta^2, 0 < p < 1$ and $\gamma(n+i,u) = \int_0^u \theta^{n+i} e^{-\theta s} s^{n+i-1} ds$, which is incomplete gamma function.

4.1. The moments

The rth moment about origin of the GSNQLD random variables has been obtained as

$$\mu_r^* = E(s^r) = \sum_{n=1}^{\infty} p(1-p)^{n-1} \sum_{i=0}^{n} \binom{n}{i} \frac{\Gamma(n+i+r)\alpha^i \theta^{2(n-i)-r}}{\Gamma(n+i)(\theta^2+\alpha)^n}$$

Therefore, the mean is defined when $r = 1$ as

$$\mu = \mu_1^* = \sum_{n=1}^{\infty} p(1-p)^{n-1} \sum_{i=0}^{n} \binom{n}{i} \frac{(n+i)!\ \alpha^i \theta^{2(n-i)-1}}{(n+i-1)!\,(\theta^2+\alpha)^n}$$
$$= \sum_{n=1}^{\infty} n\, p(1-p)^{n-1} \frac{(\theta^2+2\alpha)}{\theta(\theta^2+\alpha)} = \frac{(1-p)(\theta^2+2\alpha)}{\theta p(\theta^2+\alpha)} \tag{16}$$

The second moment , when $r = 2$, about the origin is

$$\mu_2^* = E(S^2) = \sum_{n=1}^{\infty} p(1-p)^{n-1} \sum_{i=0}^{n} \binom{n}{i} \frac{\Gamma(n+i+2)\alpha^i \theta^{2(n-i)-2}}{\Gamma(n+i)(\theta^2+\alpha)^n}$$
$$= \frac{2(4\alpha^2 - 5p\alpha^2 + p^2\alpha^2 + 4(1-p)\alpha\theta^2 + (1-p)\theta^4)}{p^2\theta^2(\theta^2+\alpha)^2}$$

Thus the variance of the sum S_N is given by

$$V(Y) = \mu_2^* - (\mu_1^*)^2 = \frac{\theta^4 + 6\alpha\theta^2 + 7\alpha^2 - 4p\alpha\theta^2 - 8p\alpha^2 + p^2\alpha^2 - p^2\theta^4 - 2\alpha p^2\theta^2}{p^2\theta^2(\theta^2+\alpha)^2} \quad (17)$$

The mean and variance of the GSNQLD for different combinations of $\alpha, \theta,$ and p are computed in Table 2. It is obvious that when $\alpha, \theta,$ and p grow up, both the mean and variance dropdown. As a result, the three parameters of the suggested distribution can be utilized to fit most count over-dispersion data sets.

Table 2: Mean and Variance of the GSNQLD distribution for different values

α		2		10		20	
p	θ	Mean	Variance	Mean	Variance	Mean	Variance
	2	6.0000	477.000	7.7143	1819.29	8.2500	3555.000
0.1	10	0.9176	190.864	0.9818	235.473	1.0500	293.400
	20	0.4522	182.704	0.4610	193.599	0.4714	207.386
	2	0.6667	10.3333	0.0512	4.3012	0.9167	72.333
0.5	10	0.1020	4.2408	0.1091	5.2181	0.1167	6.467
	20	0.0502	4.0601	0.0512	4.3012	0.0524	4.605
	2	0.0741	0.6214	0.0952	2.14286	0.1019	4.0535
0.9	10	0.0113	0.2617	0.0121	0.3212	0.0130	0.3959
	20	0.0056	0.2506	0.0057	0.2654	0.0058	0.2840

4.2. The moment generating function

The moment generating function of the GSNQLD random variables is given by

$$M_{S_N}(t) = \sum_{n=1}^{\infty} p(1-p)^{n-1}\frac{\theta^{2n}\big(\alpha+\theta(\theta-t)\big)^n}{(\theta^2+\alpha)^n(\theta-t)^{2n}} = \sum_{n=1}^{\infty} p\left[\frac{\theta^2(1-p)\big(\alpha+\theta(\theta-t)\big)}{(\theta^2+\alpha)(\theta-t)^2}\right]^n$$

$$= \frac{p(\theta^2+\alpha)(\theta-t)^2}{t^2\alpha - 2t\alpha\theta^2 + t^2\theta^2 + p\alpha\theta^2 - t\,\theta^3 - pt\theta^3 + p\theta^4} \quad (19)$$

The mean of the GSNQLD distribution can be calculated using eq.(19) as follows

$$\mu_1 = \mu = E(X) = \frac{dM_X(t)}{dt}\Big|_{t=0} = \frac{(1-p)(\theta^2+2\alpha)}{\theta p(\theta^2+\alpha)}$$

The 2nd moment about the origin is

$$\mu_2 = E(X^2) = \frac{dM_X^2(t)}{dt^2}\Big|_{t=0} = \frac{2(4\alpha^2 - 5p\alpha^2 + p^2\alpha^2 + 4(1-p)\alpha\theta^2 + (1-p)\theta^4)}{p^2\theta^2(\theta^2+\alpha)^2}$$

As a result the GSNQLD distribution's variance is given by

$$\sigma^2 = \mu_2 - \mu_1^2 = \frac{\theta^4 + 6\alpha\theta^2 + 7\alpha^2 - 4p\alpha\theta^2 - 8p\alpha^2 + p^2\alpha^2 - p^2\theta^4 - 2\alpha p^2\theta^2}{p^2\theta^2(\theta^2+\alpha)^2}$$

5. Application

This section compared the SNQLD and GSNQLD distributions to various competing distributions, as Sum of Two Lindley Distributions (STLD) in Chesneau et al. [4], XLindley (XLD) in Chouia and Zeghdoudi [5], Lindley (LD), and quasi Lindley (QLD), and discussed the SNQLD distribution's flexibility in fitting real data sets. The following symmetric set of data, see Afify et al. [1], is the survival times (in months) of 20 acute myeloid leukemia patients

2.226, 2.113, 3.631, 2.473, 2.720, 2.050, 2.061, 3.915, 0.871, 1.548,

2.746, 1.972, 2.265, 1.200, 2.967, 2.808, 1.079, 2.353, 0.726, 1.958

We used maximum likelihood estimator method to determine the estimated parameters of all distributions. For best model selection, we used well-known statistics such as Akaike information criteria (AIC), corrected Akaike information criteria (CAIC), Bayesian information criteria (BIC), Hannan-Quinn information criteria (HQIC), Anderson Darling statistic (ADS), Cramer-Von-Messes statistic (CVMS), Kolmogorov-Smirnov statistics (KSS), with its p-value (PKSS). It is known that the best model is the one with the smallest value of these statistics except PKSS, the best model which has the large value. The numerical values of these statistics and maximum likelihood estimates (along with their standard error (SE)) are presented in Table 3. From Table 3, we conclude that our proposed model has the lowest AIC, CAIC, BIC, HQIC, ADS, CVMS, and KSS values, and has the biggest value of PKSS between all compared models values, then the SNQLD distribution is the best model for fitting the analyzed real data set.

Table 3: Numerical values for analyzing the real data set.

Model	Est. parameters (SEs)	AIC	CAIC	BIC	HQIC	ADS	CVMS	KSS	PKSS
SNQLD	$\hat{\theta} = 4.57855\ (1.23781)$ $\hat{\alpha} = 4.5331 \times 10^{-7}\ (5.4701)$	56.5801	57.2859	58.5715	56.9688	0.951909	0.0816767	0.157953	0.70058
GSNQLD	$\hat{\theta} = 4.52827\ (0.947173)$ $\hat{\alpha} = 3.67809 \times 10^{12}\ (8.87867 \times 10^{17})$ $\hat{p} = 0.168206\ (0.0439796)$	83.7924	85.2924	86.7796	84.3755	3.48394	0.668079	0.35805	0.011856
STLD	$\hat{\theta} = 1.31181\ (0.160119)$	58.5697	58.7919	59.5654	58.7641	1.38245	0.253768	0.264546	0.121668
XLD	$\hat{\theta} = 0.624873\ (0.106456)$	70.8166	71.0388	71.8123	71.011	3.096	0.596629	0.319467	0.0337345
LD	$\hat{\alpha} = 0.723507\ (0.119796)$	67.9643	68.1865	68.96	68.1587	2.70357	0.512308	0.308123	0.0448493
QLD	$\hat{\alpha} = -0.297521\ (0.256468)$ $\hat{\theta} = 0.570898\ (0.163838)$	71.584	72.2899	73.5755	71.9728	10.4327	2.07451	0.527021	0.0000299227

Figure 1, we introduced the fitted PDFs for all compared models and histogram for the analyzed real data set.

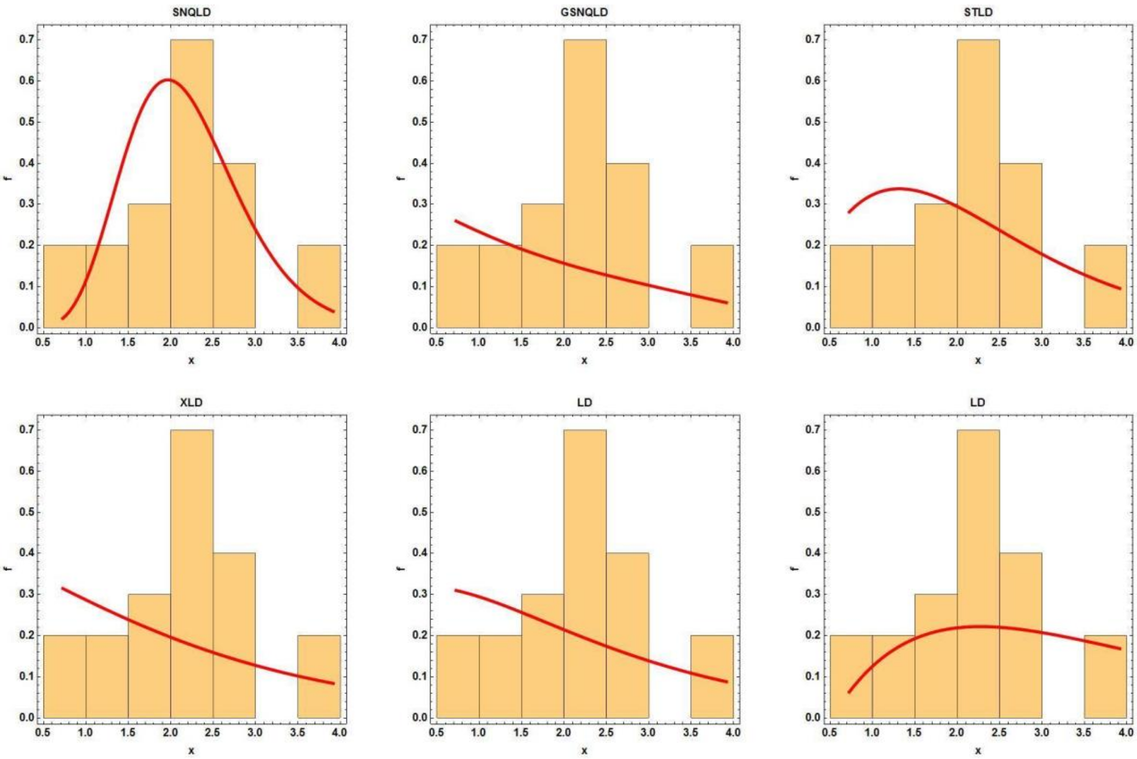

Figure 1: Estimated pdf of the SNQLD, GSNQLD and other compared models for the acute myeloid leukemia real data set.

Figure 2, the P-P plots for all compared models are presented.

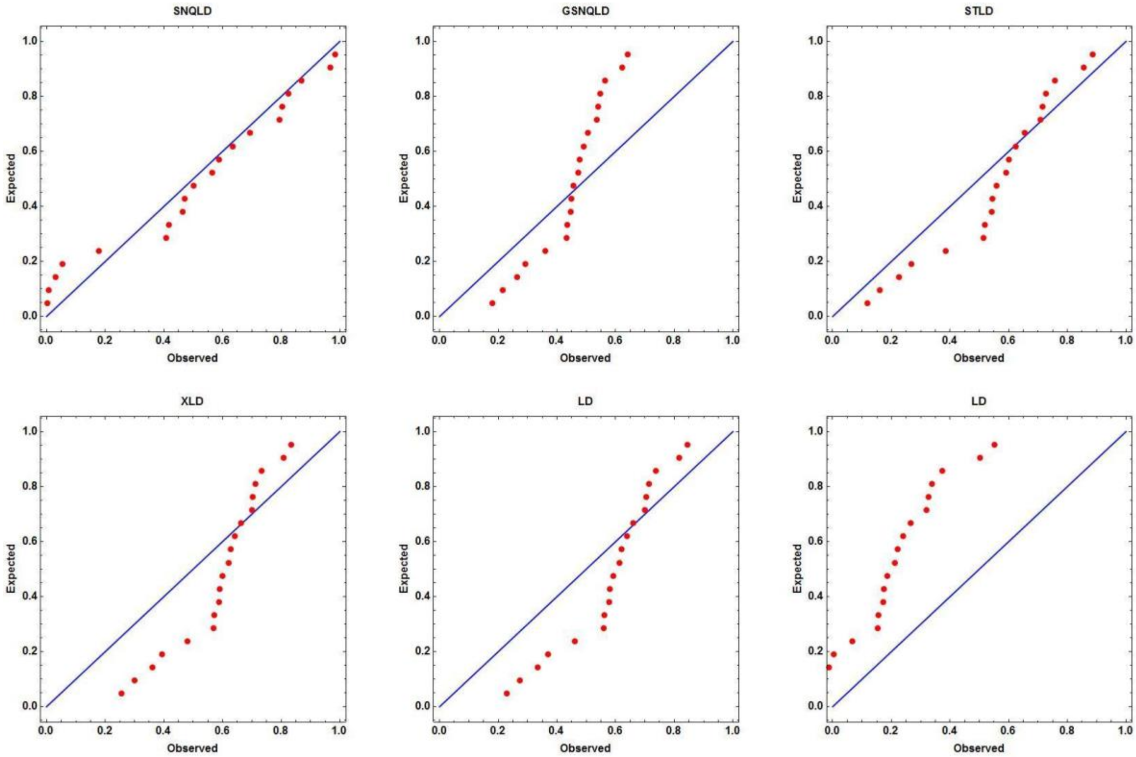

Figure 2: P-P plot of the SNQLD, GSNQLD and other compared models for the acute myeloid leukemia real data set.

From figures 1 and 2, the SNQLD distribution is the best model for fitting the analyzed real data set.

6. Conclusion

The pdf and the cdf for the SNQLD and GSNQLD distributions are obtained by using the Laplace transformation. For all cases, some statistical measures are determined. The distribution's parameters are estimated using the maximum likelihood method. The distribution's parameters are estimated using the maximum likelihood method. To test the viability and efficiency of the proposed distributions SNQLD and GSNQLD, lifetime count data sets from acute myeloid leukemia are fitted. The results should become accepted knowledge in the fields of probability theory and its allied sciences. In addition, the histogram, fitted probability density function (pdf), and P-P plots for the analyzed real data set are indicated that the SNQLD is the best model.

Availability of data and materials

The datasets used and/or analyzed during the current study available from the author on reasonable request.

Conflicts of Interest

There are no conflicts of interest declared by the author for the publication of this paper.

References

1. A. Z. Afify, Z. M. Nofal, and N. S. Butt, Transmuted complementary Weibull geometric distribution, Pak. J. Stat. Oper. Res., 10 (2014), pp. 435-454. https://doi.org/10.18187/pjsor.v10i4.836
2. Y. M. Amer, D. H. Abdel Hady, and R. M. Shalabi On a Sum and Difference of Two Quasi Lindley Distributions: Theory and Applications, American Journal of Applied Mathematics and Statistics 9(1) (2021), pp. 12-23. https://doi.org/10.12691/ajams-9-1-3
3. L. Bon, and V. Kalashnikov, Some estimates of geometric sums, Statistics & Probability Letters, 55 (2001), pp. 89-97. https://doi.org/10.1016/S0167-7152(01)00136-5
4. C. Chesneau, L. Tomy, and J. Gillariose, On a Sum and Difference of Two Lindley Distributions: Theory and Applications, Revstat - Statistical Journal 18(5) (2020), pp. 673-695.
5. S. Chouia, and H. Zeghdoudi, The XLindley Distribution: Properties and Application. Journal of Statistical Theory and Applications, 20(2) (2021), pp. 318–327. https://doi.org/10.2991/jsta.d.210607.001
6. R. A.Ganaie, V. Rajagopalan, and A. A. Rather, Weighted new quasi Lindley distribution with Properties and Applications, Journal of Xi'an University of Architecture Technology, 12(2) (2020), pp. 561 – 575.

7. R. A. Ganaie, V. Rajagopalan, and A. A. Rather, Length biased Weighted New Quasi Lindley Distribution: Statistical Properties and Applications, Pak. j. stat. oper. res., 17(1) (2021), pp. 123-136. https://doi.org/10.18187/pjsor.v17i1.3084
8. M. E. Ghitany, B. Atieh, and S. Nadarajah, Lindley Distribution and Its Application, Mathematics and Computers in Simulation, 78 (2008), pp. 493-506. https://doi.org/10.1016/j.matcom.2007.06.007
9. M. Girondot, and J. Barry, Computation of the Distribution of the Sum of Independent Negative Binomial Random Variables, Math. Comput. Appl. 28(3) (2023), No. 63. https://doi.org/10.3390/mca28030063
10. A. S. Hassan, , and M. Abd-Allah, Power Quasi Lindley Power Series Class of Distributions: Theory and Applications, Thailand Statistician. 21(2) (2023), pp. 314-336.
11. N. Kolev, E. T.Kolkovska, and J. A.Lopez-Mimbela, Joint probability generating function for a vector of arbitrary indicator variables Journal of Computational and Applied Mathematics, 186 (2006), pp. 89-98. https://doi.org/10.1016/j.cam.2005.03.066
12. D.V. Lindley, Fiducial distributions and Bayes' theorem, Journal of the Royal Statistical Society, Series B, 20 (1958), pp. 102-107. https://doi.org/10.1111/j.2517-6161.1958.tb00278.x
13. N. Y. Ma, and R. P. King, The n-fold convolution of generalized exponential-sum distribution functions, Applied Mathematics and Computation, 142 (2003), pp. 23-33. https://doi.org/10.1016/S0096-3003(02)00281-3
14. N. N. Midhu, I. Dewan, K. K. Sudheesh, and E. P. Sreedevi, On approximation and estimation of distribution function of sum of independent random variables, Stat Papers (2023). https://doi.org/10.1007/s00362-023-01413-4
15. M. M. Mohie El-Din, A. A. Teamah, A. M. Salem, and A. T. Abd El-Bar, Random sum of mid truncated Lindley distribution, Journal of Advanced Research in Statistics and Probability, 2(1) (2010), pp. 27 – 36.
16. S. Nadarajah, Y. Zhang, and T. K. Pog´any, On sums of independent generalized Pareto random variables with applications to insurance and and cat bonds, Probability in the Engineering and Informational Sciences, 32 (2017), pp. 296–305. https://doi.org/10.1017/S0269964817000055
17. R. Shanker, and A. G. Amanuel, A new quasi Lindley distribution, International Journal of Statistics and Systems, 8(2) (2013), pp. 143 – 156.
18. R. Shanker, and A. A. Mishra, Quasi Lindley Distribution, African Journal of Mathematics and Computer Science Research, 6(4) (2013), pp. 64 – 71.
19. R. Shanker, and A. A. Mishra, two-parameter Lindley distribution, Statistics in Transition-new series, 14(1) (2013), pp. 45- 56.
20. R. Shanker, S. Sharma, and R. Shanker, A two-parameter Lindley distribution for modeling waiting and survival times data, Applied Mathematics, 4 (2013), pp. 363 – 368. https://doi.org/10.4236/am.2013.42056
21. S. H. Sung, Complete convergence for weighted sums of random variables, Statistics & Probability Letters. 77 (2007), pp. 303-311. https://doi.org/10.1016/j.spl.2006.07.010
22. A. A. Teamah, and A. R. El-Alosey, Random sum of mixtures of exponential distributions. Journal of Applied Mathematics, 16 (2004), pp. 237-247.
23. G. E. Willmot, and J. Cai, Monotonicity and aging properties of random sums, Statistics & Probability Letters. 73 (2005), pp. 381-392. https://doi.org/10.1016/j.spl.2005.04.013