International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2024; 9(5): 07-14

www.mathsjournal.com
Received: 16-07-2024
Accepted: 20-08-2024

Dr. Junali Hazarika
Department of Statistics,
THB College, Karchantola
Jamugurihat, Sonitpur, Assam,
India

Corresponding Author:
Dr. Junali Hazarika
Department of Statistics,
THB College, Karchantola
Jamugurihat, Sonitpur, Assam,
India

Size-biased new Quasi-Poisson-Lindley distribution: Properties and applications

Dr. Junali Hazarika

DOI: https://dx.doi.org/10.22271/maths.2024.v9.i5a.1791

Abstract
The aim of this paper is to deal with size biased new quasi-Poisson-Lindley distribution by size biasing new quasi-Poisson-Lindley distribution. Various Statistical properties *viz.* shape of the probability function, moments, recurrence relations have been studied. Estimation of parameters are done by the method of moments and the model is illustrated by using three real data sets. The chi-square goodness of fit has been studied and the fit is compared with that obtained by using the other distributions.

Keywords: Probability function, moments, recurrence relations, method of moments, goodness of fit

Introduction

Size-biased distributions are a special case of weighted distributions. Fisher [2] introduced weighted distributions to model ascertainment biased and was later formalized by Rao [5]. It has been observed that in many situations the experimenters do not work with the random samples from the population in which they are interested. When an observation is recorded by nature according to certain stochastic model, the recorded observation will not have the original distribution unless every observation is given an equal chance of being recorded. Since the observations are recorded with unequal probability, the resulting sampled distribution does not follow the original distribution.

If a random variable X has distribution $P_0(x;\,\theta)$ with unknown parameter θ, then the corresponding weighted distribution is of the form

$$P(x;\,\theta) = \frac{w(x)P_0(x;\,\theta)}{E[w(x)]},$$

where, $w(x)$ is a non-negative function such that $E[w(x)]$ exists. The weighted distribution with weight $w(x) = x$ is known as size biased distribution having the form,

$$P^*(x;\theta,\alpha) = \frac{xP_0(x;\theta,\alpha)}{E(x)}.$$

In this chapter, a size-biased new quasi-Poisson-Lindley (SBNQPL) distribution has been obtained in section II and has been compared with size-biased quasi-Poisson-Lindley (SBQPL) distribution. In section III graphical representations of SBNQPL distribution have been shown. Statistical properties of the distribution have been studied in section IV. In Section V, the estimation of parameters of SBNQPL distribution has been discussed. Goodness of fit has been discussed in section VI.

The two-parameter size-biased quasi-Poisson-Lindley distribution obtained by Shanker and Mishra (2014) [12] has the probability mass function (pmf) as,

$$P(x;\,\theta,\alpha) = \frac{x\theta^2(\theta x+\theta\alpha+\theta+\alpha)}{(\alpha+2)(\theta+1)^{x+2}}, x = 1,2,\ldots;\ \theta > 0, \alpha > -2.$$

The r^{th} factorial moments have been obtained as,

$$\mu'_{(r)} = \frac{\Gamma(r+1)}{(\alpha+2)\theta^r}[(\alpha+r+2)(r+1)+r\theta(\alpha+r+1)] .$$

The raw moments are

$$\mu'_1 = \frac{\theta(\alpha+2)+2(\alpha+3)}{\theta(\alpha+2)} ,$$

$$\mu'_2 = \frac{\theta^2(\alpha+2)+6\theta(\alpha+3)+6(\alpha+4)}{\theta^2(\alpha+2)} ,$$

$$\mu'_3 = \frac{\theta^3(\alpha+2)+14\theta^2(\alpha+3)+36\theta(\alpha+4)+24(\alpha+5)}{\theta^3(\alpha+2)} ,$$

$$\mu'_4 = \frac{\theta^4(\alpha+2)+30\theta^3(\alpha+3)+126\theta^2(\alpha+4)+240\theta(\alpha+5)+120(\alpha+6)}{\theta^4(\alpha+2)} .$$

The variance has been obtained as

$$\mu_2 = \frac{2\{\theta(\alpha^2+5\alpha+6)+(\alpha^2+6\alpha+6)\}}{\theta^2(\alpha+2)^2} .$$

The probability generating function has been obtained as,

$$g(t) = \frac{\theta^2 t(1+\theta)[(\theta+\alpha+\alpha\theta)(1+\theta-t)+2\theta+\theta^2]}{(\alpha+2)(1+\theta-t)^3}, |t| < 1.$$

The recurrence relation for probability may be written as

$$p_{r+1} = \frac{[\{(1+\theta)^2+2(1+\theta)r\}p_r-(3+2\theta-r)p_{r-1}-p_{r-2}]}{(1+\theta)^2(r+1)},$$

where,

$$p_1 = \frac{\theta^2(\theta\alpha+2\theta+\alpha)}{(\alpha+2)(\theta+1)^3} ,$$

$$p_2 = \frac{2\theta^2(\theta\alpha+3\theta+\alpha)}{(\alpha+2)(\theta+1)^4} .$$

Derivation of the proposed distribution

Let us consider X to be a random variable following NQPL distribution having the probability mass function (pmf),

$$P_0(x;\ \theta,\alpha) = \frac{\theta^2}{(1+\theta)^{x+2}}\left(1+\frac{\theta+\alpha x}{\theta^2+\alpha}\right), x = 0,1,2 \ldots;\ \theta > 0, \alpha > \theta^2.$$

Then, the probability mass function (pmf) of (SBNQPL) distribution may be obtained as

$$P^*(x;\theta,\alpha) = \frac{xP_0(x;\theta,\alpha)}{\mu}, x = 1,2,\ldots\ ;\ \theta > 0, \alpha < -\theta^2,$$

where, $P_0(x;\ \theta,\alpha)$ is the p.m.f of NQPL distribution and $\mu = \frac{\theta^2+2\alpha}{\theta(\theta^2+\alpha)}$, is the mean of the NQPL distribution.

Thus, the pmf of SBNQPL distribution has been obtained as,

$$P^*(x;\theta,\alpha) = \frac{x\theta^3(\theta^2+\alpha)}{(\theta^2+2\alpha)}\left(\frac{\alpha x+\theta^2+\alpha+\theta}{(1+\theta)^{x+2}(\theta^2+\alpha)}\right),$$

$$= \frac{x\theta^3}{(\theta^2+2\alpha)}\left(\frac{\alpha x+\theta^2+\alpha+\theta}{(1+\theta)^{x+2}}\right),$$

$$= \frac{x\theta^3}{(\theta^2+2\alpha)(1+\theta)^{x+1}}\left[\theta+\frac{\alpha(x+1)}{(1+\theta)}\right], x = 1,2,\ldots;\ \theta > 0, \alpha < -\theta^2.$$

SBNQPL distribution may also be obtained from the size-biased Poisson distribution having the probability mass function (pmf) as,

$$g(x/\lambda) = \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}, \lambda > 0, \tag{1}$$

When the parameter λ of size-biased Poisson distribution follows the size-biased new quasi-Lindley distribution having the density function,

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

Then, the pmf of size-biased new quasi-Poisson-Lindley distribution may also be obtained from equation (1) and (2) by integrating the mixture model as,

$$P^*(x;\theta,\alpha)=\int_0^\infty \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}\frac{\lambda\theta^3}{\theta^2+2\alpha}(\theta+\alpha\lambda)e^{-\theta\lambda}d\lambda\ ,$$

$$=\frac{\theta^3}{(\theta^2+2\alpha)(x-1)!}\int_0^\infty e^{-\lambda(1+\theta)}\lambda^x(\theta+\alpha\lambda)\,d\lambda,$$

$$=\frac{\theta^3}{(\theta^2+2\alpha)(x-1)!}\left(\theta\int_0^\infty e^{-\lambda(1+\theta)}\lambda^x d\lambda+\alpha\int_0^\infty e^{-\lambda(1+\theta)}\lambda^{x+1}d\lambda\right),$$

$$=\frac{x\theta^3}{(\theta^2+2\alpha)(1+\theta)^{x+1}}\left(\theta+\frac{\alpha(x+1)}{1+\theta}\right),x=1,2,\dots;\theta>0,\alpha>0 \tag{3}$$

It has been observed that when $\alpha=\theta$, SBNQPL distribution reduces to size-biased Poisson-Lindley (SBPL) distribution having the pmf as,

$$P^{**}(x;\ \theta)=\frac{\theta^3x(x+\theta+2)}{(\theta+2)(\theta+1)^{x+2}};\theta>0,x=1,2,\dots .$$

Graphical representation of SBNQPL distribution

To study the behavior of SBNQPL distribution for different values of parameter α and θ, the probability for possible values of x are computed using equation (3). Figure 1 represents the pmf plot of SBNQPL distribution for fixed α i.e $\alpha=0.1$ and $\theta=1.5, 2.0$ and 3.0 and Figure 2 represents the pmf plot of SBNQPL distribution for $\alpha=0.9$ and $\theta=1.5, 2.0$ and 3.0 respectively.

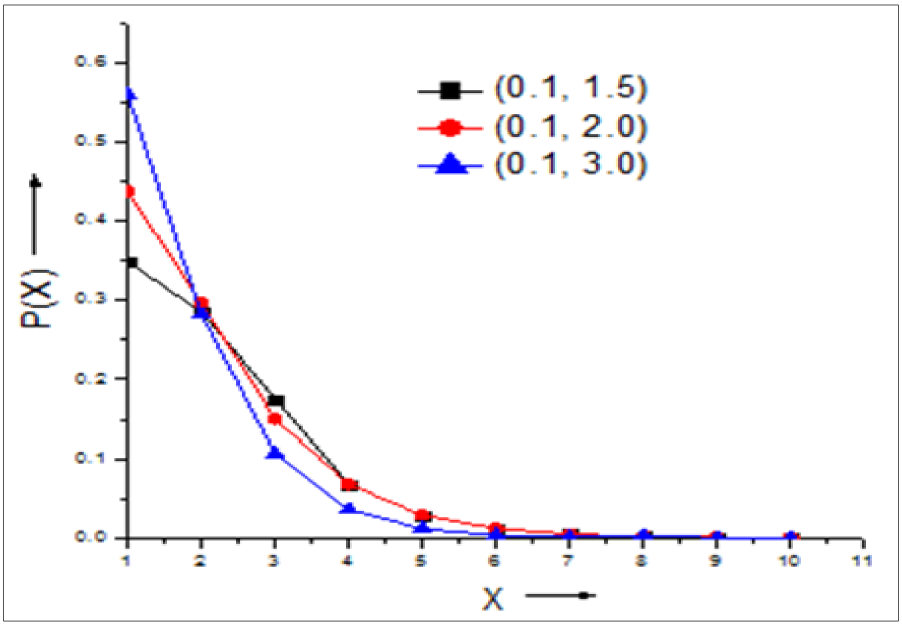

Fig 1: Pmf plot of SBNQPL distribution for $\alpha=0.1$ and $\theta=1.5, 2.0, 3.0$

Statistical properties of SBNQPL distribution

A. Shape of the probability function

We have, $\frac{P(x+1;\,\theta,\alpha)}{P(x;\,\theta,\alpha)}=\left(\frac{1}{1+\theta}\right)\left(1+\frac{1}{x}\right)\left\{1+\frac{\alpha}{(1+\theta)\theta+\alpha(x+1)}\right\}$ which is a decreasing function in $'x'$ and hence SBNQPL distribution is log-concave. Thus, we may conclude that SBNQPL distribution is unimodal and has an increasing failure rate. [Johnson *et al.*] [4].

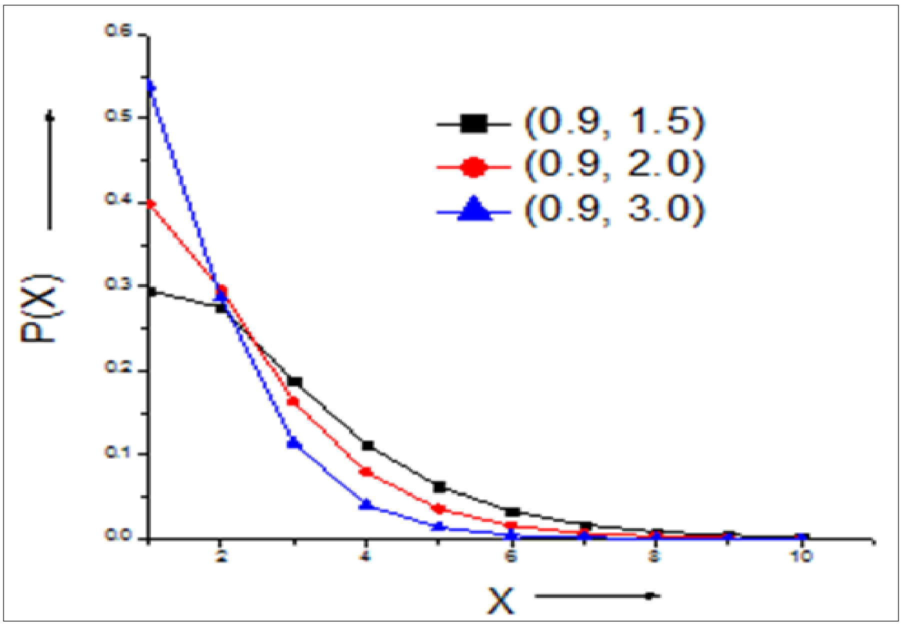

Fig 2: Pmf plot of SBNQPL distribution for $\alpha = 0.9$ and $\theta = 1.5, 2.0, 3.0$

B. Moments and related measures

The r^{th} factorial moment of SBNQPL distribution may be obtained as,

$\mu'_{(r)} = E\left[E\left(X^{(r)}|\lambda\right)\right]$ where, $X^{(r)} = X(X-1)(X-2)\ldots(X-r+1)$

$$= \int_0^\infty \sum_{x=0}^\infty x^{(r)} \frac{e^{-\lambda}\lambda^x}{(x-1)!} \frac{\lambda\theta^3}{(\theta^2+2\alpha)} (\theta+\alpha\lambda)e^{-\theta\lambda} d\lambda\,,$$

$$= \int_0^\infty \lambda^{r-1} \sum_{x=r}^\infty x \frac{e^{-\lambda}\lambda^{x-r}}{(x-r)!} \frac{\lambda\theta^{`3}}{(\theta^2+2\alpha)} (\theta+\alpha\lambda)e^{-\theta\lambda} d\lambda.$$

Taking $'x+r'$ in place of x, we get,

$$\mu'_{(r)} = \int_0^\infty \lambda^{r-1} \left(\sum_{x=0}^\infty (x+r) \frac{e^{-\lambda}\lambda^x}{x!}\right) \frac{\lambda\theta^3}{(\theta^2+2\alpha)} (\theta+\alpha\lambda)e^{-\theta\lambda} d\lambda\,.$$

The expression in the bracket is $(\lambda+r)$ and hence we have,

$$\mu'_{(r)} = \frac{\theta^3}{(\theta^2+2\alpha)} \int_0^\infty (\lambda+r)\lambda^r(\theta+\alpha\lambda)e^{-\theta\lambda} d\lambda\,.$$

Hence the solution of the gamma integral will be,

$$\mu'_{(r)} = \frac{\theta^3}{(\theta^2+2\alpha)} \int_0^\infty (\lambda^{r+1} + r\lambda^r)(\theta+\alpha\lambda)e^{-\theta\lambda} d\lambda,$$

$$= \frac{\Gamma(r+1)}{\theta^r(\theta^2+2\alpha)} [(r+1)(\theta^2+\alpha r+2\alpha+r\alpha\theta)+r\theta^3]. \tag{4}$$

Thus, from equation (4) the r^{th} factorial moments may be obtained by substituting

$r = 1,2,3$ and 4 as,

$$\mu'_{(1)} = \frac{\theta^3+2\theta^2+6\alpha+2\alpha\theta}{\theta(\theta^2+2\alpha)},$$

$$\mu'_{(2)} = \frac{2(2\theta^3+3\theta^2+12\alpha+6\alpha\theta)}{\theta^2(\theta^2+2\alpha)},$$

$$\mu'_{(3)} = \frac{6(3\theta^3+4\theta^2+20\alpha+12\alpha\theta)}{\theta^3(\theta^2+2\alpha)},$$

$$\mu'_{(4)} = \frac{24(4\theta^3+5\theta^2+30\alpha+20\alpha\theta)}{\theta^4(\theta^2+2\alpha)}.$$

Now, using the relationship between raw and factorial moments we have obtained the raw moments as,

$$\mu_1' = \frac{\theta^3+2\theta^2+6\alpha+2\alpha\theta}{\theta(\theta^2+2\alpha)},$$

$$\mu_2' = \frac{\theta^4+6\theta^3+6\theta^2+18\alpha\theta+2\alpha\theta^2+24\alpha}{\theta^2(\theta^2+2\alpha)},$$

$$\mu_3' = \frac{\theta^5+14\theta^4+36\theta^3+24\theta^2+144\alpha\theta+42\alpha\theta^2+2\alpha\theta^3+120\alpha}{\theta^3(\theta^2+2\alpha)},$$

The variance of SBNQPL distribution may be obtained as,

$$\mu_2 = \mu_2' - \mu_1'^2,$$

$$\mu_2 = \frac{2\theta^5+10\alpha\theta^3+2\theta^4+12\theta^2\alpha+12\alpha^2\theta+12\alpha^2}{\theta^2(\theta^2+2\alpha)^2}$$

C. Recurrence Relations for Probabilities of SBNQPL distribution

The probability generating function which is considered as a useful tool for dealing with discrete random variables has been used to generate all the probabilities of the distribution.

The probability generating function (pgf) of size biased new quasi-Poisson-Lindley distribution may be obtained as,

$$g(t) = E(t^x),$$

$$= \sum_{x=1}^{\infty} t^x P^*(x),$$

where, $P^*(x)$ is the pmf of SBNQPL distribution.

$$= \sum_{x=1}^{\infty} t^x \frac{x\theta^3}{(\theta^2+2\alpha)(1+\theta)^{x+1}}\left(\theta + \frac{\alpha(x+1)}{1+\theta}\right),$$

$$= \frac{\theta^3}{(\theta^2+2\alpha)}\left\{\sum_{x=1}^{\infty} t^x \frac{x}{(1+\theta)^{x+1}}\left(\theta + \frac{\alpha(x+1)}{1+\theta}\right)\right\},$$

$$= \frac{\theta^3\left\{(\theta(1+\theta)+\alpha)\sum_{x=1}^{\infty} x\left(\frac{t}{1+\theta}\right)^x + \alpha\sum_{x=1}^{\infty} x^2\left(\frac{t}{1+\theta}\right)^x\right\}}{(\theta^2+2\alpha)(1+\theta)^2},$$

$$= \frac{\theta^3 t[\theta(\theta+1-t)+2\alpha]}{(\theta^2+2\alpha)(1+\theta-t)^3}, |t| < 1. \quad (5)$$

Equating the coefficient of $'t^r'$ on both sides of equation (5), the expression for recurrence relation for probabilities of SBNQPL distribution has been obtained as,

$$p_r = \frac{\{3(\theta+1)^2 p_{r-1} - 3(\theta+1)p_{r-2} + p_{r-3}\}}{(\theta+1)^3}, r > 2 \; (() \quad (6)$$

where $p_1 = \frac{\theta^3(\theta^2+\theta+2\alpha)}{(1+\theta)^3(\theta^2+2\alpha)}$,

$$p_2 = \frac{2\theta^3(\theta^2+\theta+3\alpha)}{(1+\theta)^4(\theta^2+2\alpha)},$$

and, $p_3 = \frac{3\theta^3(\theta^2+\theta+4\alpha)}{(1+\theta)^5(\theta^2+2\alpha)}$.

The higher probabilities may be obtained from equation (6) substituting by $r = 4, 5, \ldots$ etc.

It has been observed that when $\alpha = \theta$ SBNQPL distribution reduces to SBPL distribution. and

$$\hat{\alpha} = b\hat{\theta}^2 = \frac{4\hat{b}\left(1+3\hat{b}\right)^2}{\left(1+2\hat{b}\right)^2(\bar{x}-1)^2},$$

where $\bar{x} - 1 = \frac{2(1+3b)}{\theta(1+2b)}$ and $\hat{b}$ may be obtained from equation (8).

D. Maximum Likelihood Estimates

Let $x_1, x_2, \ldots\ldots, x_n$ be a random sample of size n from the SBNQPL distribution and let f_x be the observed frequency in the sample corresponding to $X = x \; (x = 1,2,\ldots\ldots,k)$ such that $\sum_{x=1}^{k} f_x = n$.

The Likelihood function L of the new quasi-Poisson Lindley distribution may be given as

$$L = \prod_{x=1}^{k} f(x;\theta,\alpha),$$

$$= \left(\frac{\theta^3}{\theta^2+\alpha}\right)^n \frac{1}{(1+\theta)^{\sum_{x=1}^{k}(x+2)f_x}} \prod_{x=1}^{n}[\alpha x^2 + x(\theta^2+\theta+\alpha)]^{f_x}.$$

The log likelihood function becomes

$$\log L = n\log\left(\frac{\theta^3}{\theta^2+\alpha}\right) - \sum_{x=1}^{k} f_x(x+2)\log(1+\theta) + \sum_{x=1}^{k} f_x \log[\alpha x^2 + x(\theta^2+\theta+\alpha)].$$

Then the log likelihood equations are obtained as

$$\frac{\partial \log L}{\partial \alpha} = -\frac{n}{\theta^2+\alpha} + \sum_{x=1}^{k} \frac{(x^2+x)f_x}{[\alpha x^2 + x(\theta^2+\theta+\alpha)]} = 0\ ,$$

And

$$\frac{\partial \log L}{\partial \theta} = \frac{3n}{\theta} - \frac{2n\theta}{\theta^2+\alpha} - \sum_{x=1}^{k}\frac{(x+2)f_x}{1+\theta} + \sum_{x=1}^{k}\frac{x(2\theta+1)f_x}{[\alpha x^2+x(\theta^2+\theta+\alpha)]} = 0\ .$$

The above two equations could be solved using Fisher's scoring method. We have

$$\frac{\partial^2 \log L}{\partial \alpha^2} = \frac{n}{(\theta^2+\alpha)^2} - \sum_{x=1}^{k}\frac{(x^2+x)^2 f_x}{[\alpha x^2+x(\theta^2+\theta+\alpha)]^2}\ ,$$

$$\frac{\partial^2 \log L}{\partial \alpha \partial \theta} = \frac{2n\theta}{(\theta^2+\alpha)^2} - \sum_{x=1}^{k}\frac{x(x^2+x)(2\theta+1)f_x}{[\alpha x^2+x(\theta^2+\theta+\alpha)]^2}\ ,$$

$$\frac{\partial^2 \log L}{\partial \theta^2} = -\frac{3n}{\theta^2} + \frac{4n\theta}{(\theta^2+\alpha)^2} + \sum_{x=1}^{k}\frac{(x+2)f_x}{(1+\theta)^2} + \sum_{x=1}^{k}\frac{(2\alpha x^2 - 2x\theta^2 - 2\theta x + 2\alpha x - 1)f_x}{[\alpha x^2+x(\theta^2+\theta+\alpha)]^2},$$

The following equations can be solved for $\hat{\theta}$ and $\hat{\alpha}$ iteratively till sufficiently close values of θ and α are obtained. Thus, we have

$$\begin{bmatrix} \frac{\partial^2 \log L}{\partial \alpha^2} & \frac{\partial^2 \log L}{\partial \alpha \partial \theta} \\ \frac{\partial^2 \log L}{\partial \alpha \partial \theta} & \frac{\partial^2 \log L}{\partial \theta^2} \end{bmatrix} \begin{bmatrix} \hat{\theta} - \theta_0 \\ \hat{\alpha} - \alpha_0 \end{bmatrix} = \begin{bmatrix} \frac{\partial \log L}{\partial \alpha} \\ \frac{\partial \log L}{\partial \theta} \end{bmatrix}_{\substack{\hat{\theta}=\theta_0 \\ \hat{\alpha}=\alpha_0}}$$

Where $\theta_0$and α_0 are the initials value of θ and α respectively.

Goodness of fit

To illustrate the applications and to justify suitability of SBNQPL distribution, the SBNQPL distribution has been fitted to some published datasets for which various distributions were fitted earlier by different author. In fitting of probability distributions, the estimation of parameters plays a very important role. The parameters of SBNQPL distribution have been estimated by the method of moment. The parameters are then used to obtain the expected frequencies which are later used to calculate the χ^2 value. The SBNQPL distribution is then compared with other distributions as measured by χ^2 criterion.

Here, the SBNQPL distribution has been fitted to three data sets and the distribution has been compared with size-biased Poisson-Lindley, size-biased quasi-Poisson-Lindley.

Table 1: Observed and expected frequency of SBNQPL distribution for the counts of people in public places on a spring afternoon in Portland. [data from James] [3]

Size of group	Observed frequencies	Expected frequencies		
		SBPL	**SBQPL**	**SBNQPL**
1	1486	1532.5	1485.4	1484.00
2	694	630.6	697.3	699.41
3	195	191.9	189.7	191.73
4	37	51.3	40.6	38.78
5	10	12.8	8.3	7.73
6	1	3.9	1.7	1.35
Total	2423	2423	2423	2423
Parameters		$\hat{\theta} = 4.5082$	$\hat{\alpha} = -0.7916$ $\hat{\theta} = 7.14227$	$\hat{\alpha} = -63.6352$ $\hat{\theta} = 7.2136$
χ^2		13.766	0.710	0.588
D.F		4	3	3
p-value		0.01	0.871	0.899

The SBNQPL distribution has also been fitted to data reported by James [3] in Table 1 which is regarding the distribution for the Counts of people in public places on a spring afternoon in Portland. The parameters are obtained by the method of moment as $\hat{\alpha} = -63.6352$ and $\hat{\theta} = 7.2136$. The derived distribution has been compared with SBPL and SBNQPL distributions.

Table 2: Observed and expected frequency of SBNQPL distribution for the counts of shopping groups-Eugene, spring, Department Store and Public Market. [data from Coleman and James] [1]

Size of Groups	Observed frequencies	Expected frequencies		
		SBPL	SBQPL	SBNQPL
1	316	323.0	315.7	315.52
2	141	132.5	142.9	143.07
3	44	40.2	40.5	42.48
4	5	10.7	8.9	7.10
5	4	3.6	2.0	1.84
Total	510	510	510	510
Parameters		$\hat{\theta} = 4.5224$	$\hat{\alpha} = -0.52229$ $\hat{\theta} = 6.57872$	$\hat{\alpha} = -81.0372$ $\hat{\theta} = 6.6080$
χ^2		3.021	0.659	0.0854
D.F		3	2	2
$p-$ value		0.40	0.72	0.96

In Table 2 the SBNQPL distribution has been fitted to data reported by Coleman and James [1] for the Counts of Shopping Groups-Eugene, spring, Department Store and Public Market. From the data set we have obtained the expected frequency and χ^2 value for goodness of fit.

Table 3: Observed and expected frequency of SBNQPL distribution for the counts of Play Groups-Eugene, spring, Department Store and Public Market. [data from Smin off] [13]

Size of Groups	Observed Frequency	Expected frequency		
		SBPL	SBQPL	SBNQPL
1	305	314.4	304.2	304.1
2	144	134.4	148.5	149.0
3	50	42.5	42.8	43.0
4	5	11.8	9.2	9.0
5	2	3.1	1.9	1.9
6	1	0.8	0.4	0.4
Total	507	507	507	507
Parameters		$\hat{\theta}$= 4.3179	$\hat{\theta}$ = 6.74857 $\hat{\alpha}$ = -0.76547	$\hat{\theta}$ = 6.8123 $\hat{\alpha}$ = -58.6404
χ^2		6.415	2.53	1.78
D.F		2	1	1
p-value		0.09	0.12	0.18

Table 3 is regarding the sets of data reported by Simonoff [13] for counts of Play Groups-Eugene, Spring Public Playground D where there are six groups and the observed frequency corresponding to the groups are given.

Conclusion

The observed frequencies and expected frequencies of SBPL, SBQPL and SBNQPL distribution for the three data sets have been shown in Table 1, Table 2 and Table 3 respectively for its comparison. Comparing the observed and expected frequencies the χ^2 values and p-values have been calculated for testing the goodness of fit.

It has been observed that based on chi-square value and p-value from Table 1, Table 2, and Table 3 the SBNQPL distribution provide a closer fit than SBPL and SBQPL distribution.

References

1. Coleman JS, James J. The equilibrium size distribution of freely forming groups. Sociometry. 1961;24:36-45.
2. Fisher RA. The effects of methods of ascertainment upon the estimation of frequencies. Annals of Eugenics. 1934;6:13-25.
3. James J. The distribution of free-forming small group size. American Sociological Review. 1953;18:569-570.
4. Johnson NL, Kemp AW, Kotz S. Univariate discrete distributions. Hoboken, NJ: John Wiley & Sons; c2005.
5. Rao CR. On discrete distributions arising out of methods of ascertainment. Sankhya, A. 1965;27:311-324.
6. Sankaran M. The discrete Poisson-Lindley distribution. Biometrics. 1970;26:145-149.
7. Shanker R, Ghebretsadik AH. A new quasi-Lindley distribution. International Journal of Statistics and Systems. 2013;8(2):143-156.
8. Shanker R, Mishra A. A quasi-Lindley distribution. African Journal of Mathematics and Computer Science Research. 2013;6(4):64-71.
9. Shanker R, Mishra A. A two-parameter Lindley distribution. Statistics in Transition new series. 2013;14(1):45-56.

10. Shanker R, Sharma S, Shanker R. A two-parameter Lindley distribution for modeling waiting and survival times data. Applied Mathematics. 2013;4:363-368.
11. Shanker R, Mishra A. On a size-biased quasi-Poisson-Lindley distribution. International Journal of Probability and Statistics. 2013;2(2):28-34.
12. Shanker R, Mishra A. A two-parameter Poisson-Lindley distribution. International Journal of Statistics and Systems. 2014;9(1):79-85.
13. Simonoff JS. Analyzing categorical data. New York: Springer; c2003.