Chapter 3

Size-biased new quasi Poisson-Lindley distribution

3.1 Introduction

Size-biased distributions are a special case of weighted distributions. Fisher [37] introduced weighted distributions to model ascertainment biased and was later formalized by Rao [84]. It has been observed that in many situations the experimenter does 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 3.2 and has been compared with size-biased quasi Poisson-Lindley (SBQPL) distribution. In section 3.3 graphical representations of SBNQPL distribution has been shown. Statistical properties of the distribution have been studied in section 3.4. In Section 3.5, the methods of estimation of parameters of SBNQPL distribution have been discussed. Goodness of fit has been discussed in section 3.6.

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

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

The r^{th} factorial moment has 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 have been obtained from factorial moments as

$$\mu'_1 = \mu'_{(1)} = \frac{\theta(\alpha+2)+2(\alpha+3)}{\theta(\alpha+2)}\ ,$$

$$\mu'_2 = \mu'_1 + \mu'_{(2)} = \frac{\theta^2(\alpha+2)+6\theta(\alpha+3)+6(\alpha+4)}{\theta^2(\alpha+2)}\ ,$$

$$\mu'_3 = \mu'_1 + 3\mu'_{(2)} + \mu'_{(3)} = \frac{\theta^3(\alpha+2)+14\theta^2(\alpha+3)+36\theta(\alpha+4)+24(\alpha+5)}{\theta^3(\alpha+2)}\ ,$$

$$\mu'_4 = \mu'_1 + 7\mu'_{(2)} + 6\mu'_{(3)} + \mu'_{(4)}$$

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

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 probabilities is

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

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

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

3.2 Derivation of SBNQPL distribution

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

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

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

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

where, $P_0(x;\ \theta,\alpha)$ is the pmf 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,

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

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

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

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

$$g(x/\lambda) = \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!};\ x = 1,2,\dots,\lambda > 0 \quad (3.2.1)$$

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

$$f(\lambda;\ \theta,\alpha) = \frac{\lambda\theta^3}{\theta^2+2\alpha}(\theta+\alpha\lambda)e^{-\theta\lambda};\ \lambda > 0,\ \theta > 0, \alpha > 0. \quad (3.2.2)$$

Then, the pmf of size-biased new quasi Poisson-Lindley distribution may be obtained from equation (3.2.1) and (3.2.2) by integrating the mixture model as,

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

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

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

$$= \frac{x\theta^3}{(\theta^2+2\alpha)(1+\theta)^{x+1}}\left(\theta+\frac{\alpha(x+1)}{1+\theta}\right);\ x = 1,2,\dots,\theta > 0, \alpha > 0. \quad (3.2.3)$$

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

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

3.3 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 (3.2.3).

Figure 3.1 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 3.2 represents the pmf plot of SBNQPL distribution for $\alpha = 0.9$ and $\theta = 1.5, 2.0$ and 3.0 respectively.

Figure 3.1 pmf plot of SBNQPL distribution for $\alpha = 0.1$ and $\theta = 1.5, 2.0, 3.0$

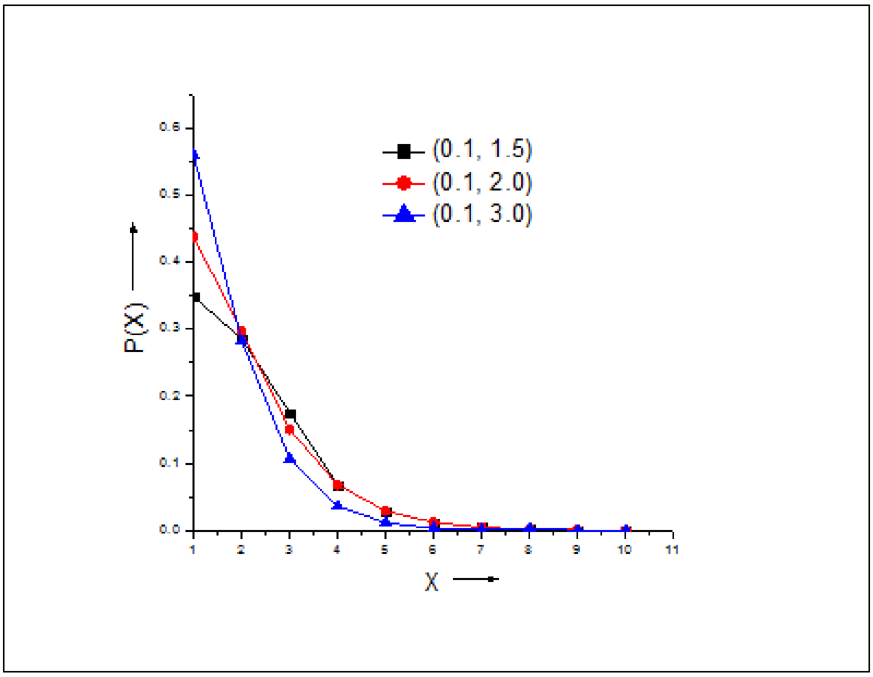
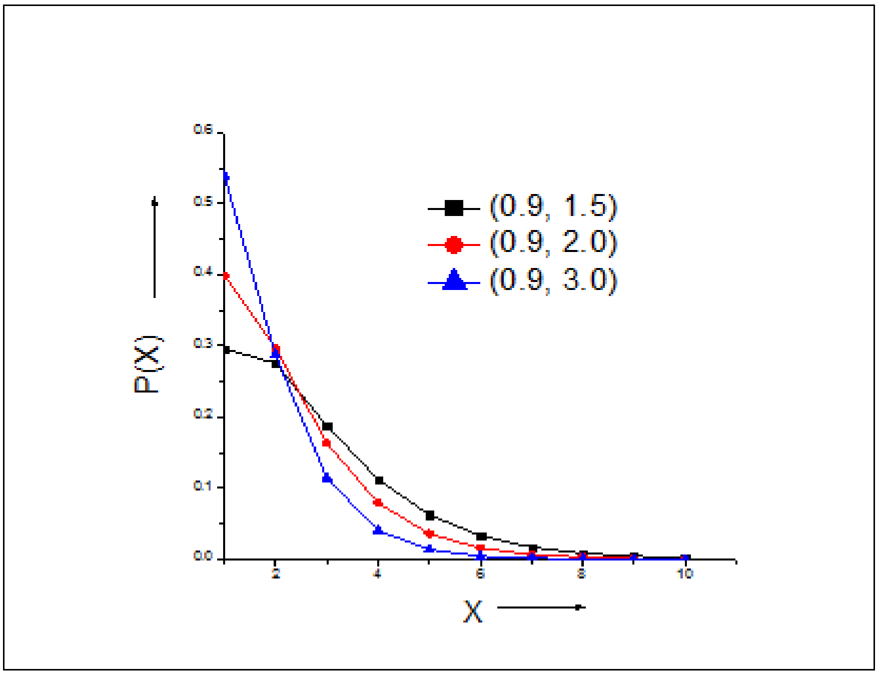


Figure 3.2 pmf plot of SBNQPL distribution for $\alpha = 0.9$ and $\theta = 1.5, 2.0, 3.0$



3.4 Statistical properties of SBNQPL distribution

Some properties of SBNQPL distribution have been studied in this section.

3.4.1 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. [56]]

3.4.2 Moments and related measures

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

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

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

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

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,

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

$$= \frac{\Gamma(r+1)}{\theta^r(\theta^2+2\alpha)}\left[(r+1)(\theta^2+\alpha r+2\alpha+r\alpha\theta)+r\theta^3\right]. \qquad (3.4.1)$$

Thus from equation (3.4.1) the r^{th} factorial moments have been 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,

$$\mu_2 = \mu'_2 - {\mu'_1}^2,$$

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

3.4.3 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 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.

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

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

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

$$= \frac{\theta^3 t[\theta(\theta+1-t)+2\alpha]}{(\theta^2+2\alpha)(1+\theta-t)^3}, \qquad |t| < 1 \ . \tag{3.4.2}$$

Equating the coefficient of $'t^r'$ on both sides of equation (3.4.2), the expression for recurrence relation for probabilities of SBNQPL distribution has been obtained as,

$$p_r = \frac{\{3(\theta+1)^2 \, p_{r-1} - 3(\theta+1)p_{r-2} + p_{r-3}\}}{(\theta+1)^3}, r > 3 \tag{3.4.3}$$

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 (3.4.3) by substituting

$r = 4, 5, \ldots$ etc.

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

The moment generating function may be written as

$$m(t) = \frac{\theta^3 e^t[\theta(\theta+1-e^t)+2\alpha]}{(\theta^2+2\alpha)(1+\theta-e^t)^3}$$

3.4.4 Recurrence relation for factorial moment generating function of SBNQPL distribution

The factorial moment generating function (fmgf) of SBNQPL distribution may be obtained from the probability generating function (pgf) as,

$$G(t) = g(1+t)\ ,$$

$$G(t) = \frac{\theta^3(1+t)[\theta(\theta-t)+2\alpha]}{(\theta^2+2\alpha)(\theta-t)^3}. \tag{3.4.3}$$

Expanding equation (3.4.3) and then equating the co-efficient of $\frac{t^r}{r!}$ we may obtain the recursive expression for factorial moment generating (fmgf) function as,

$$\mu'_{(r)} = \frac{3r\mu'_{(r-1)} - 3\theta r\mu'_{(r-2)} + r(r-1)(r-2)\mu'_{(r-3)}}{\theta^3}, r > 3$$

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

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

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

It has been observed that the factorial moment generating function of SBNQPL distribution reduces to SBPL distribution when $\alpha = \theta$.

3.4.5 Index of dispersion and coefficient of variation

The index of dispersion defined as the ratio of variance to mean is a measure to determine whether a distribution is over-dispersed, equi-dispersed or under-dispersed. It is denoted by γ and may be written as,

$$\text{I. D} = \frac{\sigma^2}{\mu} = \frac{2\theta^5+10\alpha\theta^3+2\theta^4+12\theta^2\alpha+12\alpha^2\theta+12\alpha^2}{\theta(\theta^2+2\alpha)(\theta^3+2\theta^2+2\alpha\theta+6\alpha)}.$$

Coefficient of variation denoted by CV may be defined as the ratio of standard deviation to mean. It is invariant of change of scale but not of origin and is given as,

$$\mathrm{CV} = \frac{\sigma}{\mu} = \frac{\sqrt{2\theta^5+10\alpha\theta^3+2\theta^4+12\theta^2\alpha+12\alpha^2\theta+12\alpha^2}}{(\theta^3+2\theta^2+2\alpha\theta+6\alpha)}.$$

3.5 Estimation of parameters of SBNQPL distribution

3.5.1 Method of moments

The first two moments are required to estimate the parameters by the method of moments. Thus we have,

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

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

Now let, $\frac{(\mu_2'-1)-3(\mu_1'-1)}{(\mu_1'-1)^2} = \frac{3(\theta^2+4\alpha)(\theta^2+2\alpha)}{2(\theta^2+3\alpha)^2} = k$ (say). (3.5.1)

Substituting $\alpha = b\theta^2$ in (3.5.1) a quadratic equation in *b* has been obtained as

$$(24-18k)b^2 + 6(3-2k)b + (3-2k) = 0 \qquad (3.5.2)$$

Replacing the first two population moment by the respective sample moments in equation (3.5.1) the estimate of k may be obtained. Substituting the estimate of k in (3.5.2) the estimate $\hat{b}$ can be obtained from the quadratic equation (3.5.2).

The mean of SBNQPL distribution may be given as $\bar{x} = 1 + \frac{2(\theta^2+3\alpha)}{\theta(\theta^2+2\alpha)}$. (3.5.3)

Substituting $\alpha = b\theta^2$, in the above expression (3.5.3) we have

$$\hat{\theta} = \frac{2(1+3b)}{(1+2b)(\bar{x}-1)},$$

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

where $\bar{x}-1 = \frac{2(1+3b)}{\theta(1+2b)}$ and $\hat{b}$ may be obtained from equation (3.5.2).

3.5.2 Maximum Likelihood Estimates

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

The Likelihood function L for the vector of parameter $\Theta = (\theta, \alpha)^T$of the size biased new quasi Poisson-Lindley distribution may be written as

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

The log likelihood function becomes

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

Then the log likelihood equations are obtained as

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

and

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

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

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

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

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

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.

3.6 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 authors. In fitting of probability distribution, the estimation of parameters plays a very important role. The parameters of SBNQPL distribution has been estimated by the method of maximum likelihood. The parameters are then used to obtain the expected frequencies 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 and size-biased quasi Poisson-Lindley distributions.

Table 3.1: Observed and expected frequency of SBNQPL distribution for the counts of people in public places on a spring afternoon in Portland. [data from James [55]]

Size of group	Observed frequencies	Expected frequencies		
		SBPL	SBQPL	SBNQPL
1	1486	1532.5	1534.4	1517.9
2	694	630.6	625.1	648.0
3	195	191.9	191.9	193.4
4	37	51.3	52.1	49.1
5	10	12.8	13.3	11.3
6	1	3.9	3.5	2.5
Total	2423	2423	2423	2423
Parameters		$\hat{\theta} = 4.5082$	$\hat{\alpha} = 24.4597$ $\hat{\theta} = 4.0458$	$\hat{\alpha} = 40.3068$ $\hat{\theta} = 5.3352$
	χ^2	13.766	16.1	7.97
	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 [55] in Table 3.1 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 maximum likelihood as $\hat{\alpha} = 40.3068$and $\hat{\theta} = 5.3352$. The derived distribution has been compared with SBPL and SBNQPL distributions.

Table 3.2: 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 [24]]

Size of Groups	Observed frequencies	Expected frequencies		
		SBPL	SBQPL	SBNQPL
1	316	323.0	323.8	320.1
2	141	132.5	128	136.8
3	44	40.2	40.1	42.5
4	5	10.7	10.9	10.3
5	4	2.0	2.8	2.4
Total	510	510	510	510
Parameters		$\hat{\theta} = 4.5212$	$\hat{\alpha} = 29.5164$ $\hat{\theta} = 4.0477$	$\hat{\alpha} = 35.0710$ $\hat{\theta} = 5.3016$
χ^2		3.021	5.49	4.03
d.f		3	2	2
$p -$ value		0.40	0.72	0.96

In Table 3.2 the SBNQPL distribution has been fitted to data reported by Coleman and James [24] 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 3.3: Observed and expected frequency of SBNQPL distribution for the counts of Play Groups-Eugene, Spring Public Playground D. [data from Sminoff [98]]

Size of Groups	Observed Frequency	Expected frequency		
		SBPL	SBQPL	SBNQPL
1	305	314.4	314.19	310.97
2	144	134.4	127.12	138.0
3	50	42.5	42.8	43.0
4	5	11.8	9.2	9.0
5	2	3.1	1.9	2.7
6	1	0.8	0.4	0.6
Total	507	507	507	507
Parameters		$\hat{\theta}$= 4.3179	$\hat{\theta}$ = 3.9563 $\hat{\alpha}$ = 14.2299	$\hat{\theta}$ = 5.0189 $\hat{\alpha}$ = 30.7657
χ^2		6.415	6.5443	4.2595
d.f		2	1	1
p-value		0.09	0.12	0.18

Table 3.3 is regarding the sets of data reported by Simonoff [98] 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.

3.7 Conclusion:

The observed frequencies and expected frequencies of SBPL, SBQPL and SBNQPL distribution for the three data sets have been shown in Table 3.1, Table 3.2 and Table 3.3 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 3.1, Table 3.2 and Table 3.3 the SBNQPL distribution provide a closer fit than SBPL and SBQPL distribution.