Chapter 6

Two-parameter Quasi Poisson-Lindley Distribution and Its Applications

6.1 Introduction

The two-parameter quasi Poisson-Lindley (QPL) distribution which includes Poisson-Lindley distribution of Sankaran (1970) to model count data as a special case is studied in this chapter. An attempt has been made to investigate certain properties of the quasi Poisson-Lindley distribution like recurrence relations involving probabilities, factorial moments without derivatives, since these forms are easy to handle on computer. The zero-truncated, zero-modified and size-biased forms of the distribution have also been discussed in this chapter.

The organization of the chapter as follows: Section 6.2 deals with the derivation of the QPL distribution. To study the behavior of the QPL distribution with varying values of the two parameters plots the probabilities of the QPL distribution for different values of the parameters in section 6.3. In section 6.4, certain distributional properties of the distribution have been investigated. The problem of parameter estimation is considered in section 6.5. Zero-truncated, zero-modified and size-biased forms of the QPL distribution is discussed in section 6.6. Section 6.7 concerned with the discussion of the application part of the distribution.

6.2 Two-parameter quasi Poisson-Lindley distribution

The two-parameter quasi Poisson-Lindley (QPL) distribution obtained from the Poisson distribution with probability mass function (pmf)

$$g(x;\lambda) = \frac{e^{-\lambda}\lambda^{x}}{x!} \quad , \; x = 0, 1, 2, \dots , \;\; \lambda > 0 \qquad (6.2.1)$$

when its parameter λ follows quasi Lindley distribution of Shanker and Mishra (2013) with probability density function (pdf)

$$f(x;\alpha,\theta) = \frac{\theta}{\alpha+1}(\alpha+\theta x)e^{-\theta x} \;\; , \quad x > 0,\; \theta > 0, \alpha > -1 \qquad (6.2.2)$$

The probability mass function (pmf) of the QPL distribution may be derived as

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

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

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

$$= \frac{\theta}{\alpha+1}\frac{[\alpha+\theta(1+\alpha)+\theta x]}{(1+\theta)^{x+2}} \; ; \;\; x = 0,1,2,\dots \qquad (6.2.3)$$

where $\theta > 0$ and $\alpha > -1$. This resultant distribution is the quasi Poisson-Lindley distribution with parameter α and θ and denote it by QPL (α,θ). It can be seen that the

Poisson-Lindley distribution of Sankaran (1970) is a particular case of the QPL distribution (6.2.3) at $\alpha = \theta$. The model (6.2.3) is a more generalized form and more flexible than the Poisson-Lindley distribution for analyzing different types of count data.

6.3 Certain properties of QPL distribution

This section is devoted to studying certain distributional properties of the QPL distribution, specifically distribution function, recurrence relations of probabilities and factorial moments, moments, coefficient of variation etc.

6.3.1 Cumulative distribution function

Let X is a random variable with QPL (α, θ) distribution. Then the cumulative distribution function of X is given by

$$F_X(x) = \frac{\theta}{(1+\alpha)(1+\theta)^2} \sum_{j=0}^{x} \frac{[\alpha+\theta(1+\alpha)+\theta j]}{(1+\theta)^j} \; ; x = 0,1,2,\dots,\theta > 0, \alpha > -1. \qquad (6.3.1)$$

6.3.2 Probabilities

The probability generating function (pgf) $g(\mathrm{t})$ of the QPL distribution derived from the model (6.2.3) may be written as

$$g(\mathrm{t}) = \frac{\theta}{\alpha+1} \frac{\{\theta(1+\alpha)+\alpha-\alpha t\}}{(1+\theta-t)^2} \qquad (6.3.2)$$

The recurrence relation for probabilities of the QPL distribution may be written as

$$P_{r+1} = \frac{1}{A(r+1)} [\{(2r + 1)\alpha(1 + \theta) + \theta(r + 2)\}P_r - \alpha r P_{r-1}] \, , \; r = 1, 2, \dots \qquad (6.3.3)$$

where, $A = \frac{1}{(1+\theta)(\alpha+\theta+\alpha\theta)}$

$$P_0 = \frac{\theta}{(\alpha+1)} \frac{(\alpha+\theta+\alpha\theta)}{(1+\theta)^2}$$

This is obtained by differentiating (6.3.2) w. r. t. 't' and then equating the coefficient of t^r on both sides. The higher order probabilities can be obtained very easily by putting $r = 1,2,3 \ldots$ in the equation (6.3.3).

The general expression for probabilities of the QPL distribution may be written as

$$P_r = \frac{\theta[\alpha+\alpha\theta+(r+1)\theta]}{(1+\alpha)(1+\theta)^{r+2}} \; ; \; r = 1, 2, 3, \ldots. \qquad (6.3.5)$$

6.3.3 Moments

The factorial moment recurrence relation for the QPL distribution may be expressed as

$$\mu'_{(r+1)} = \frac{1}{\theta^2(1+\alpha)}\left[\theta\{(1+2\alpha)r + (2+\alpha)\}\mu'_{(r)} - \alpha r^2 \mu'_{(r-1)}\right] ; r > 1 \qquad (6.3.6)$$

where, $$\mu'_{(1)} = \frac{(\alpha+2)}{\theta(1+\alpha)} \qquad (6.3.7)$$

The factorial moment recurrence relation is obtained by differentiating the factorial moment generating function (fmgf) of the QPL distribution

$$m(t) = \frac{\theta}{\alpha+1} \frac{\{\theta(1+\alpha)+\alpha-\alpha t\}}{(1+\theta-t)^2} \qquad (6.3.8)$$

w.r.t. 't' and then equating the coefficients of $\frac{t^r}{r!}$ on both sides of (6.3.8).

The general expression for r^{th}order factorial moment of the QPL distribution may be written as

$$\mu'_{(r)} = \frac{r!(\alpha+r+1)}{\theta^r(1+\alpha)} ; \; r = 1, 2, \ldots \qquad (6.3.9)$$

After obtaining the first four factorial moments, by substituting $r = 1,2,3, 4$ in (6.3.9) and then using the relationship between factorial moments and moments about origin, the first four moments about the origin of the QPL distribution may be obtained as

$$\mu_1' = \frac{(2+\alpha)}{\theta(1+\alpha)} \text{ (Mean)} \tag{6.3.10}$$

$$\mu_2' = \frac{(2+\alpha)(\theta+2)+2}{\theta^2(1+\alpha)} \tag{6.3.11}$$

$$\mu_3' = \frac{(2+\alpha)\theta(\theta+6)+6(\theta+\alpha+4)}{\theta^3(1+\alpha)} \tag{6.3.12}$$

$$\mu_4' = \frac{(2+\alpha)\theta\left(\theta^2+14\theta+36\right)+2(7\theta^2+36\theta+12\alpha+60)}{\theta^4(1+\alpha)} \tag{6.3.13}$$

By using the identity $\mu_r = E(X-\mu)^r = \sum_{k=0}^{r}\binom{r}{k}\mu_k'(-\mu_1')^{r-k}$, the central moments can be obtained as follows

$$\mu_2 = \frac{(1+\theta)(1+\alpha)(2+\alpha)+\alpha}{\theta^2(1+\alpha)^2} \text{ (Variance)} \tag{6.3.14}$$

$$\mu_3 = \frac{(1+\alpha)(2+\alpha)\left(\theta^2+\theta^2\alpha+3\theta\alpha-6\alpha-18\right)+6(1+\alpha)^2(\theta+\alpha+4)+2(2+\alpha)^3}{\theta^3(1+\alpha)^3} \tag{6.3.15}$$

$$\mu_4 = \frac{(1+\alpha)^2(2+\alpha)\left(\theta^3+\theta^3\alpha+6\theta^2+10\theta^2\alpha+12\theta\alpha-36\theta-24\alpha-96\right)+2(1+\alpha)^3(7\theta^2+36\theta+12\alpha+60) -(2+\alpha)(1+\alpha)\left(12\alpha^2+60\alpha+24-6\theta\alpha^2\right)-3(2+\alpha)^4}{\theta^4(1+\alpha)^4} \tag{6.3.16}$$

6.3.4 Coefficients of Skewness and Kurtosis

The other important indices of the shape of the distribution, i.e., the expression for coefficients of skewness and kurtosis of the QPL distribution are given by

$$\sqrt{\beta_1} = \frac{{\mu_3}^2}{{\mu_2}^{3/2}} = \frac{(1+\alpha)(2+\alpha)A+6(1+\alpha)^2B+C}{\sqrt{D^3}} \tag{6.3.17}$$

where $A = (\theta^2 + \theta^2\alpha + 3\theta\alpha - 6\alpha - 18)$

$B = (\theta + \alpha + 4)$

$C = 2(2+\alpha)^3$

$D = (1+\theta)(1+\alpha)(2+\alpha) + \alpha$

and $$\beta_2 = \frac{\mu_4}{{\mu_2}^2} = \frac{(1+\alpha)^2(2+\alpha)E + 2(1+\alpha)^3F - (2+\alpha)(1+\alpha)G - H}{D^2} \quad (6.3.18)$$

where $E = (\theta^3 + \theta^3\alpha + 6\theta^2 + 10\theta^2\alpha + 12\theta\alpha - 36\theta - 24\alpha - 96)$

$F = 7\theta^2 + 36\theta + 12\alpha + 60$

$G = 12\alpha^2 + 60\alpha + 24 - 6\theta\alpha^2$

$H = 3(2+\alpha)^4$

The skewness, $\sqrt{\beta_1}$ and kurtosis, β_2 is increasing function in θ for fixed α.

It can be easily verified that at $\alpha = \theta$, these measures reduces to the respective measures of the Poisson-Lindley distribution of Sankaran (1970).

6.3.5 Coefficient of Variation

The coefficient of variation of the QPL distribution is given by

$$CV = \frac{\sigma}{\mu} = \sqrt{\frac{(1+\theta)(\alpha+1)}{\alpha+2} + \frac{\alpha}{(\alpha+2)^2}} = \sqrt{(1+\theta) - \frac{2}{(\alpha+2)^2}} \quad (6.3.19)$$

It has been seen that the coefficient of variation (CV) of the QPL distribution increases if the values of θ and α increases in equation (6.3.19). For varying values of θ and α, the corresponding changes in the values of CV of the QPL distribution have been shown in **Table 6.1.**

Table 6.1 The oefficient of variation of QPL distribution for different values of θ and α.

$\theta \rightarrow$ / $\alpha \downarrow$	0.2	0.8	1.4	2.0	2.6	3.2
0.1	0.739	1.160	1.395	1.596	1.774	1.936
0.3	0.907	1.192	1.422	1.619	1.795	1.955
0.5	0.938	1.217	1.442	1.637	1.811	1.970
0.8	0.972	1.243	1.465	1.657	1.829	1.986
1.0	0.989	1.256	1.476	1.667	1.838	1.994
1.2	1.002	1.267	1.485	1.675	1.845	2.001

It is observed from the above table that, the CV of the QPL distribution is increased as the values of θ is increased when α is fixed. Similarly, when θ is fixed and α is increased the CV is also increased. It may also be noted that for a small changes in the values of θ there are significant differences in the values of CV . Hence, θ is the most significant parameter of this distribution.

Fig. 6.1 The CV increases as θ increases for different values of α for the QPL distribution.

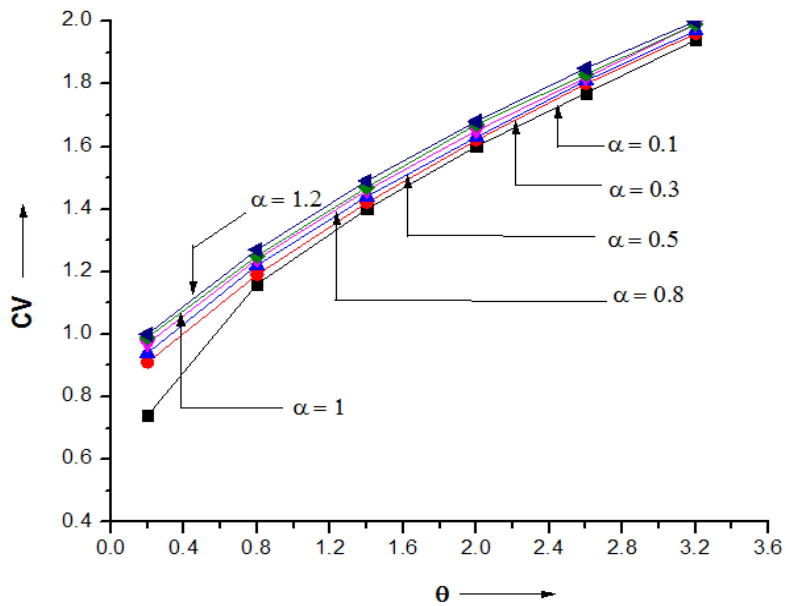
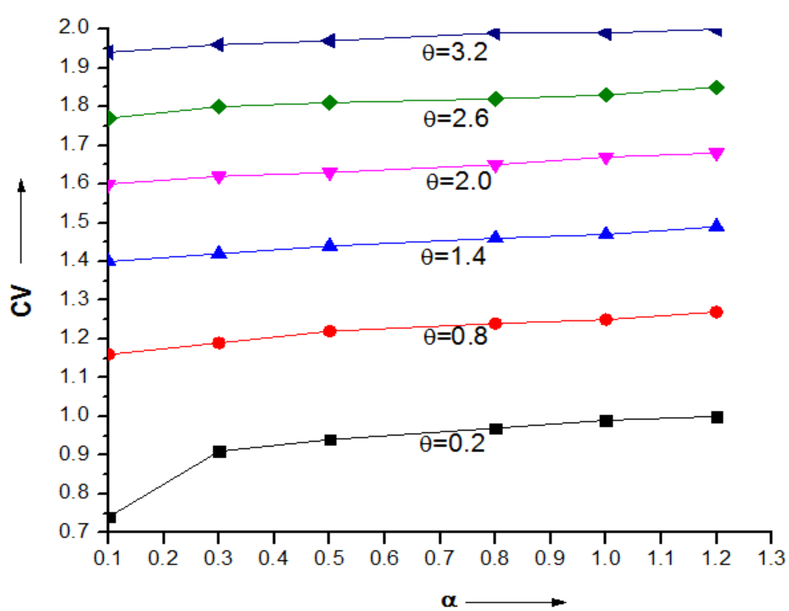


Fig. 6.2 The CV increases slowly as α increases for different values of θ for the QPL distribution.



From the above graphical representations, it is observed that there is a little significant difference in the values of the CV, when θ is increasing but for small changes in the values of α there is a significant difference in the values of CV of the QPL distribution. Hence, θ is the more sensitive parameter than α of the QPL distribution.

6.4 Estimation of Parameters

In this section, for estimating the parameters of the QPL distribution different estimation methods, via method of moments and method of maximum likelihood have been investigated.

6.4.1 Method of moments

The QPL distribution has two parameters to be estimated and so its first two moments are required to get the estimates of its parameters by the method of moments.

From equations (6.3.10) and (6.3.11), we have

$$\frac{\mu_2' - \mu_1'}{(\mu_1')^2} = \frac{2(2+\alpha)(1+\alpha)+2(1+\alpha)}{(2+\alpha)^2} = K(say) \qquad (6.4.1)$$

which gives a quadratic equation in α as

$$g(\alpha) = (2-K)\alpha^2 + (8-2K)\alpha + (6-4K) = 0 \qquad (6.4.2)$$

Replacing the first two population moments by the respective sample moments in (6.4.1) an estimate k of K can be obtained and using it in (6.4.2), the estimate $\hat{\alpha}$ of α can be obtained. Again, substituting the value of $\hat{\alpha}$ in (6.3.10) and replacing the population mean by the sample mean $\bar{x}$, the estimate of θ is obtained as

$$\hat{\theta} = \frac{2+\alpha}{\bar{x}(1+\alpha)} \qquad (6.4.3)$$

6.4.2 Method of maximum likelihood

Let $(x_1, x_2, \dots, x_n)$ be a random sample of size n from the QPL distribution (6.2.3) and let f_x be the observed frequency in the sample corresponding to $X = x$ $(x = 1, 2, \dots, k)$ such that $\sum_{x=1}^{k} f_x = n$, where k is the largest observed value having non-zero frequency. The likelihood function, L of the QPL distribution is given by

$$\mathrm{L} = \left(\frac{\theta}{1+\alpha}\right)^n \frac{1}{(1+\theta)^{\sum_{x=1}^{k}(x+2)f_x}} \prod_{x=1}^{k} [\alpha + \theta(1+\alpha) + \theta x]^{f_x} \qquad (6.4.4)$$

and the log likelihood function as

$$logL = nlog\left(\frac{\theta}{1+\alpha}\right) - \sum_{x=1}^{k} f_x\,(x+2)\log(1+\theta) + \sum_{x=1}^{k} f_x\,log[\alpha + \theta(1+\alpha) + \theta x] \qquad (6.4.5)$$

The two log likelihood equations are thus obtained as

$$\frac{\partial logL}{\partial\theta} = \frac{n}{\theta} - \sum_{x=1}^{k} \frac{(x+2)f_x}{1+\theta} + \sum_{x=1}^{k} \frac{f_x[1+\alpha+x]}{[\alpha+\theta(1+\alpha)+\theta x]} = 0 \qquad (6.4.6)$$

$$\frac{\partial logL}{\partial\alpha} = -\frac{n}{1+\alpha} + \sum_{x=1}^{k} \frac{f_x(1+\theta)}{[\alpha+\theta(1+\alpha)+\theta x]} = 0 \qquad (6.4.7)$$

The two equations (6.4.4) & (6.4.5) are difficult to solve directly. However, the Fisher's scoring method can be applied to solve these equations.

For this, we have

$$\frac{\partial^2 logL}{\partial\theta^2} = -\frac{n}{\theta^2} + \sum_{x=1}^{k} \frac{(x+2)f_x}{(1+\theta)^2} - \sum_{x=1}^{k} \frac{f_x[1+\alpha+x]^2}{[\alpha+\theta(1+\alpha)+\theta x]^2} \qquad (6.4.8)$$

$$\frac{\partial^2 logL}{\partial\theta\partial\alpha} = -\sum_{x=1}^{k} \frac{f_x(1+x)}{[\alpha+\theta(1+\alpha)+\theta x]^2} \qquad (6.4.9)$$

$$\frac{\partial^2 logL}{\partial \alpha^2} = \frac{n}{(1+\alpha)^2} - \sum_{x=1}^{k} \frac{f_x(1+\theta)^2}{[\alpha+\theta(1+\alpha)+\theta x]^2} \qquad (6.4.10)$$

Using (6.4.6 – 6.4.10), we have

$$\begin{bmatrix} \frac{\partial^2 logL}{\partial \theta^2} & \frac{\partial^2 logL}{\partial \theta \partial \alpha} \\ \frac{\partial^2 logL}{\partial \theta \partial \alpha} & \frac{\partial^2 logL}{\partial \alpha^2} \end{bmatrix}_{\substack{\hat{\theta}=\theta_0 \\ \hat{\alpha}=\alpha_0}} \begin{bmatrix} \hat{\theta} - \theta_0 \\ \hat{\alpha} - \alpha_0 \end{bmatrix} = \begin{bmatrix} \frac{\partial logL}{\partial \theta} \\ \frac{\partial logL}{\partial \alpha} \end{bmatrix}_{\substack{\hat{\theta}=\theta_0 \\ \hat{\alpha}=\alpha_0}} \qquad (6.4.11)$$

which can be solved for the estimates, $\hat{\theta}$ and $\hat{\alpha}$, where θ_0 and α_0 are the initial values of θ and α respectively. These equations are solved numerically and iteratively till sufficiently close estimates $\hat{\theta}$ and $\hat{\alpha}$ are obtained.

6.5 Zero-truncated, zero-modified and size-biased versions

In this section, zero-truncated, zero-modified and size-biased version of the QPL distribution have been stated and also try to derive certain recurrence relations of probabilities and factorial moments of the size-biased QPL distribution.

6.5.1 Zero-truncated quasi Poisson-Lindley distribution

In the most common form of truncation, the zeroes are not recorded. In this case, the zero-truncated distributions can be used as a distribution for the sizes of groups. This situation occurs in applications such as the number of claims per claimant, the number of occupants per car. The probability mass function of the zero-truncated quasi Poisson-Lindley (ZTQPL) distribution considering by its zero-truncated from is given by

$$P_T(x;\ \alpha,\theta) = \frac{\theta[\alpha+\theta(1+\alpha)+\theta x]}{(1+\theta)^x[\alpha+\theta(\alpha+2)+1]}\ ; x = 1,2,3,\dots,\theta > 0,\alpha > -1 \qquad (6.5.1)$$

Simply denote it by ZTQPL (α , θ). A briefly discussion about ZTQPL distribution for analyzing different types of count data has been given in the next chapter.

Remark 6.1 In (6.5.1) if $\alpha \rightarrow \theta$ then the zero-truncated Poisson-Lindley distribution of Ghitany et al. (2008) is obtained as a limiting form of the ZTQPL distribution.

6.5.2 Zero-modified quasi Poisson-Lindley distribution

As mentioned in chapter 4, the zero-modified distributions are widely used in many fields because of the low frequency of the events. The major motivation force behind the development of zero-modified distribution is that, many distributions obtained in the course of experimental investigations often have an excess frequency of the observed event at zero point.

A combination of the original distribution with pmf $P_x(\alpha,\theta)\,; x = 0,1,2,\ldots$ (6.2.1) together with the degenerate distribution with all probabilities concentrated at the origin gives a zero-modified quasi Poisson-Lindley(ZMQPL) distribution with pmf

$$P[X = 0] = w + (1 - w)P_0 \qquad (6.5.2)$$

$$P[X = x] = (1 - w)P_x \;;\; x \geq 1 \qquad (6.5.3)$$

where w is a parameter assuming arbitrary value in the interval $0 < w < 1$. It is also possible to take < 0 , provided $w + (1 - w)P_0 \geq 0$ [cf. Johnson et al. (2005)] and P_0 be the zero-order probability and P_x be the pmf of the QPL distribution.

The probability generating function of the ZMQPL distribution may be written as

$$G(t) = w + (1 - w)g(t) \qquad (6.5.4)$$

where $g(t)$ be the probability generating function of the QPL distribution and $0 < w < 1\,;\ \theta > 0, \alpha > -1$ be the parameters of the distribution.

6.5.3 Size-biased quasi Poisson-Lindley distribution

Size-biased distributions arise in practice when observations from a sample are recorded with unequal probabilities, having probability proportional to some measure of unit size already discussed about this type of distribution in chapter 2 and 3. Fisher (1934) first introduced these distributions to model ascertainment bias which were later formalized by Rao (1965) in a unifying theory.

Shanker and Mishra (2013c) proposed size-biased quasi Poisson-Lindley distribution. They consider the size-biased form of quasi Poisson-Lindley (SBQPL) distribution and obtained the pmf of the distribution as

$$P_{S_1}(x;\theta,\alpha) = \frac{xP(x;\theta,\alpha)}{\mu} = \frac{\theta^2}{\alpha+2}\frac{x(\theta x+\theta\alpha+\theta+\alpha)}{(1+\theta)^{x+2}}\,;\, x = 1, 2, \ldots;\ \theta > 0, \alpha > -2 \qquad (6.5.4)$$

where, $\mu = \frac{\alpha+2}{\theta(\alpha+1)}$ is the mean of quasi Poisson-Lindley distribution. They also obtained the raw moments of the SBQPL distribution and also discussed maximum likelihood method and the method of moments to estimate the parameters of this distribution. It is also noted that, the size-biased Poisson-Lindley (SBPL) distribution of Ghitany and Al-Mutairi (2008) already discussed in chapter 2 may be obtained putting $\alpha = \theta$ in (6.5.4) as a special case of the SBQPL distribution.

The probability generating function (pgf) of the SBQPL (α, θ) distribution may be written as

$$g(t) = \frac{\theta^2}{(\alpha+2)}\frac{t(\alpha+\alpha\theta+2\theta-\alpha t)}{(1+\theta-t)^3}\,;\, \theta > 0, \alpha > -2 \qquad (6.5.5)$$

The pgf of the SBPL distribution of Ghitany and Al-Mutairi (2008) may be obtained as a particular case by putting $\alpha = \theta$ in (6.5.5) [see chapter 2].

The recurrence relation for probabilities of the SBQPL distribution is given by

$$P_r = \frac{1}{(1+\theta)^3}\left[3(1+\theta)^2 P_{r-1} - 3(1+\theta)P_{r-2} + P_{r-3}\right]; \quad r > 3 \tag{6.5.7}$$

where, $P_1 = \frac{\theta^2}{(\alpha+2)}\frac{(\alpha+\alpha\theta+2\theta)}{(1+\theta)^3}$

$P_2 = \frac{\theta^2}{(\alpha+2)}\frac{2(\alpha+\alpha\theta+3\theta)}{(1+\theta)^4}$

$P_3 = \frac{\theta^2}{(\alpha+2)}\frac{3(\alpha+\alpha\theta+4\theta)}{(1+\theta)^5}$

be the first three probabilities of the SBQPL distribution.

The general form of r^{th} order probability is written as

$$P_r = \frac{\theta^2}{(\alpha+2)}\frac{r[\alpha+\alpha\theta+(r+1)\theta]}{(1+\theta)^{r+2}}\ ; r = 1,2,3,\ldots \tag{6.5.8}$$

The higher order probabilities of the SBQPL distribution can be obtained very easily by using equation (6.5.8).

Differentiating the factorial moment generating function of the SBQPL distribution

$$m(t) = \frac{\theta^2}{(\alpha+2)}\frac{(1+t)(\alpha\theta+2\theta-\alpha t)}{(\theta-t)^3}\ ;\ \theta > 0, \alpha > -2 \tag{6.5.9}$$

and then equating the coefficients of $\frac{t^r}{r!}$ on both sides, the recurrence relation for factorial moments of the SBQPL distribution may be obtained as

$$\mu'_{(r)} = \frac{r(r-1)}{\theta^3}\left[3\theta^2\mu'_{(r-1)} - (r-2)\mu'_{(r-2)} + (r-2)(r-3)\mu'_{(r-3)}\right];\ r > 3. \tag{6.5.10}$$

where, $\mu'_{(1)} = \frac{1![2(\alpha+3)+\theta(\alpha+2)]}{\theta(\alpha+2)}$

$$\mu'_{(2)} = \frac{2![3(\alpha+4)+2\theta(\alpha+3)]}{\theta^2(\alpha+2)}$$

$$\mu'_{(3)} = \frac{3![4(\alpha+5)+3\theta(\alpha+4)]}{\theta^3(\alpha+2)}$$

Now using the relationship between factorial moments and moments about origin, the first four moments about origin of the SBQPL distribution were obtained very easily. Shanker and Mishra (2013c) obtained the more general form of r^{th} order factorial moment of SBQPL distribution and obtained the first four moments about origin.

Hence, the variance of the SBQPL distribution may be written as follows

$$\mu_2 = \frac{2\theta(\alpha+2)(\alpha+3)+6(\alpha+2)(\alpha+4)-4(\alpha+3)^2}{\theta^2(\alpha+2)^2} \tag{6.5.11}$$

Remark 6.2 The variance of the SBPL distribution [already discussed in chapter 2] may be obtained from the SBQPL as a particular case, putting $\alpha = \theta$ in (6.5.11).

6.6 Fitting of quasi Poisson-Lindley distribution and applications

The well-known one-parameter Poisson-Lindley distribution is the most commonly used probability distribution in different fields such as biology & ecology, social-information, genetic and so on. The QPL distribution has only two parameters and more generalized form of the Poisson-Lindley distribution. For which it is also applicable in these fields.

To see the applicability and suitability of the QPL distribution to applied in various fields we have tried to fitting this distribution to some published data sets to which earlier the Poisson-Lindley distribution has been fitted by Sankaran (1970) and others.

Two sets of real data have been considered. The first set of data represents the problem of mistakes in copying groups of random digits [data from Kemp and Kemp (1965)] for which single parameter Poisson-Lindley distribution was fitted by Sankaran (1970) showed in **Table 6.2** and the second data set on Pyrausta nibilalis [data from Beall (1940)] for which Poisson-Lindley distribution was fitted by Borah and Deka Nath (2001a) showed in **Table 6.3**.

Table 6.2 and **Table 6.3** give the comparison of observed and expected frequencies for the Poisson-Lindley and QPL distributions. Also, the χ^2 – statistic and estimated parameters are presented in these tables. The present two-parameter QPL distribution appears to give a satisfactory fit in both the cases, whereas the Poisson-Lindley distribution does not.

It is clear from the tables that the χ^2-values of the test for fitting the QPL distribution to the two sets of data shows that, the null hypothesis H_0 ("distribution of the data is QPL") cannot be rejected; indeed, the close agreement between the observed and expected frequencies and provides a "good fit" to these two sets errors and biological data.

Table 6.2 Fitting of the QPL distribution to data of mistakes in copying groups of random digits. [data from Kemp and Kemp (1965)]

No. of errors per group	Observed frequencies	Expected frequencies	
		Poisson-Lindley Sankaran (1970) $\hat{\theta}$ =1.743	Fitted QPL (MoM) $\hat{\theta}$=2.332 $\hat{\alpha}$ = 0.210
0	35	33.1	34.3
1	11	15.3	13.8
2	8	6.8	7.2
3	4	2.9	3.4
4	2	1.2	1.3
Total	60	60.0	60.0
χ^2		2.20	0.952
$d.f$		1	1
$p-value$		0.14	0.33

Note: QPL: Quasi Poisson-Lindley distribution.

MoM: Method of moments.

Table 6.3 Fitting of the QPL distribution to Pyrausta nubilslis data [data from Beall (1940)]

No. of Insects	Observed frequencies	Expected frequencies	
		Poisson-Lindley Borah and Deka Nath (2001a) $\hat{\theta}$=1.808	Fitted QPL(MoM) $\hat{\theta}$ =2.247 $\hat{\alpha}$ = 0.459
0	33	34.08	32.0
1	12	11.23	13.8
2	6	5.61	6.0
3	3	2.71	2.5
4	1	1.28	1.0
5	1	1.09	0.7
χ^2		0.653	0.328
$d.f$		1	1
$p-value$		0.42	0.57

Note: QPL: Quasi Poisson-Lindley distribution.

MoM: Method of moments.

****