Journal of Statistics Applications & Probability Letters
An International Journal

http://dx.doi.org/10.18576/jsapl/080306

An Under-Dispersed Discrete Distribution and Its Application

Brijesh P. Singh[1], *Gunjan Singh*[2], *Utpal Dhar Das*[1,*] *and Deepak K. Maurya*[1]

[1]Department of Statistics, Institute of Science, Banaras Hindu University, Varanasi, U.P. India
[2]Department of Statistics, Amity School of Applied Science, Amity University, Lucknow, U.P. India

Received: 26 May 2021, Revised: 2 Jun. 2021, Accepted: 11 Jul. 2021
Published online: 1 Sep. 2021

Abstract: Discrete distribution is normally used to study the count data but most of the time count data are not equi-dispersed but they are under dispersed or over dispersed. To model these type of data we need appropriate probability distributions. Some weighted or mixture distributions are used to analyse the under and over dispersed count data. In this paper, an attempt has been made to propose a new kind of Poisson distribution. Some statistical properties are derived such as probability generating function, moment generating function, characteristics function, cumulant generating function, moments, coefficient of variation, reoccurrence relation and index of dispersion. Parameters are estimated by the method of moments and maximum likelihood estimators. The proposed distribution is applied on real data sets to check the suitability of proposed distribution over some competent distribution. The proposed distribution is found a better choice than others.

Keywords: Under-Dispersed, Poisson distribution, PGF, MGF, CF, CGF, Method of moment, Maximum likelihood estimation.

1 Introduction

Poisson distribution is a popular discrete distribution which is used to explain the count data due to its simplicity. It is an equi-dispersed distribution (its variance and mean are equal), but the count data are not equi-dispersed every time. It may be under dispersed (variance less than the mean) or over dispersed (variance greater than the mean). Fisher [1] introduced the weighted distribution first time to learn the biases among the data sets, which was further formalized by Rao [2] in a unified way to deal with the problem arises when the data obtained from non-experimental, non-replicated and non-random categories. Patil and Rao [3] studied the size-biased distribution which is a weighted distribution with weight function that is may or may not be bounded to unity. Weighted discrete distributions provide more flexible probability models to explain the over/under dispersed data as well as truncated data Patil et al. [4].

There are many researchers who discussed the weighted Poisson distribution in their study for under as well as over dispersed data sets (Efron [5], Cameron and Johansson [6], Ridout and Besbeas [7], Castillo and Pérez-Casany [8,9], Kokonendji et al. [10] and Balakrishnan et al. [11]). Consul and Jain [12] proposed the a new class of generalized Poisson distribution which is the limiting form of generalized negative binomial distribution and its properties and applications are discussed Consul [13]. Johnson et al. [14] also described the generalized Poisson distribution and some of its important properties.

The concept of mixture distribution is also used to explain the count data. Sankaran [15] introduced the Poisson-Lindley distribution (PLD) first time to explain the count data. Ghitany and Al-Mutairi [16] studied the PLD in depth and derive its several important properties. Shanker et al. [17] derived the general expression for r^{th} factorial moment of PLD and applied on count data obtained from biological sciences and PLD gives very close fit than the Poisson distribution. Singh et al. [18] used PLD along with Poisson-exponential and Poisson-Gamma for explaining number of child death data and found PLD is a better model. Generalization of PLD is provided by many of the

* Corresponding author e-mail: utpal.statmath@gmail.com

researchers e.g. Shanker & Mishra [19,20] provided a two parameter PLD and a quasi-PLD. Further Shanker and Tekie [21] developed a new quasi PLD. Zero-truncated Poisson-Lindley distribution introduced by Ghitany et al. [22], to model the count data with excluding zero counts. Ghitany and Mutairi [16] discussed the method of estimation for discrete PLD. Inflated Poisson-Lindley distribution is considered by Borah and Deka [23] and further Singh et al. [24] used this for modelling out migration from the household.

A size-biased Poisson-Lindley distribution (SBPLD), developed by Ghitany and Al-Mutairi [25] by compounding size biased Poisson distribution with Lindley distribution, SBPLD is the size-biased version of the PLD. Shanker et al. [26] studied the SBPLD in detail and found that SBPLD is a suitable model for thunderstorms data. Mahmoudi and Zakerzadeh [27] provided the generalized Poisson-Lindley distribution (GPLD), which is a mixture of Poisson and two-parameter generalized Lindley distribution (GLD), proposed by Zakerzadeh and Dolati [28], and generalized size-biased Poisson-Lindley distribution (GSBPLD) provided by Shankar and Shukla [29]. GSBPLD is the size biased version of generalized Poisson-Lindley distribution (GPLD), to explain the count data excluding zero counts due to have enough flexibility in two-parameter GSBPLD than the one-parameter SBPLD. Singh et al. [30] introduced a generalized version of Lindley type distribution that can be used with Poisson distribution. In this paper an attempt has been made to develop a new kind of Poisson distribution and to explore its various statistical properties.

2 Proposed Distribution

The probability density function of proposed distribution i.e. Under-dispersed Poisson distribution (UDPD-I) is given by

$$p(x;\lambda,\theta)=\frac{e^{-\lambda}\lambda^{x-1}(\lambda+\theta x)}{(1+\theta)x!}\quad ;\theta>0,\lambda>0\quad \&\quad x=0,1,2,.... \tag{1}$$

In fact, this is a mixture of two distributions such as Poisson distribution and size-biased Poisson distribution which is given as

$$p(x)=\alpha p_1(x)+(1-\alpha)p_2(x),\quad \text{where}\quad \alpha=\frac{1}{1+\theta}$$
$$p_1(x)=\frac{e^{-\lambda}\lambda^{x}}{x!},\quad \lambda>0,x=0,1,2,..\quad \text{and}\quad p_2(x)=\frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!},\lambda>0,x=1,2,.. \tag{2}$$

If θ is zero then the UDPD-I converted into simple Poisson distribution. If θ is increasing the weight of size-biased Poisson distribution is increasing.

Weighted distributions provide flexible probability models for studying over/under-dispersed data as studied by Patil and Rao [3]. A weighted distribution is defined as follows

$$P^{*}(x,\lambda)=\frac{w(x)P(x_i;\lambda)}{E[w(x)]};\quad x\in\mathbb{R},\lambda>0. \tag{3}$$

Where $w(x)$ is the weight function, $P(x_i;\lambda)$ is the parent distribution. The probability distribution given in (1) can be obtained as a weighted Poisson distribution where weight function $\left(1+\frac{\theta x}{\lambda}\right)$. If θ is equal to λ then the UDPD-I becomes a single parameter weighted Poisson distribution i.e. UDPD-II with weight function is $(1+x)$ given in (4)

$$p(x;\lambda)=\frac{e^{-\lambda}\lambda^{x}(1+x)}{(1+\lambda)x!}\quad ;\lambda>0\quad \&\quad x=0,1,2,.... \tag{4}$$

The plot of probability mass function of UDPD-I is given in Fig 1. It has been observed that for a fix value of θ, if the value of λ is increasing, the shape of PMF become flat. Also it has been observed that for fix value of λ , if the value of θ is increasing, the probability of non happening of the event is decreasing and mode of the distribution is increasing.

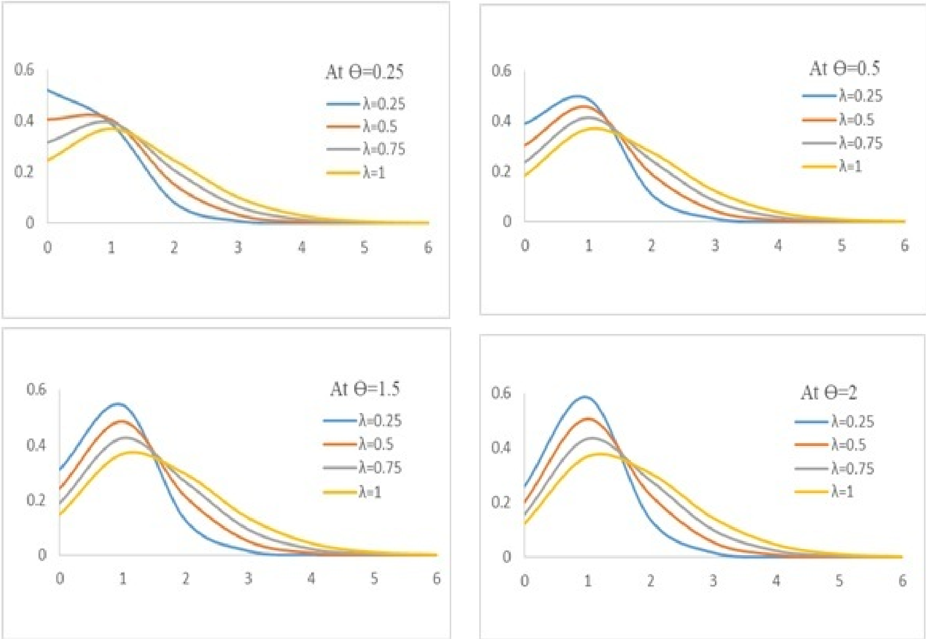

Fig. 1: PMF plot (smooth) of proposed distribution UDPD-I for different values of θ and λ.

The cumulative mass function of UDPD-I can be written as

$$F(x)=\sum_{u=0}^{x}p(u;\lambda)=\frac{e^{-\lambda}}{(1+\theta)}\sum_{u=0}^{x}\frac{\lambda^{u}}{u!}+\frac{\theta e^{-\lambda}}{(1+\theta)}\sum_{u=1}^{x}\frac{\lambda^{u-1}}{(u-1)!} \tag{5}$$

Hazard function of the UDPD-I is given by

$$h(x)=\frac{f(x)}{1-F(x)}=\frac{\frac{e^{-\lambda}\lambda^{x-1}(\lambda+\theta x)}{(1+\theta)x!}}{1-\frac{e^{-\lambda}}{(1+\theta)}\sum\limits_{u=0}^{x}\frac{\lambda^{u}}{u!}+\frac{\theta e^{-\lambda}}{(1+\theta)}\sum\limits_{u=1}^{x}\frac{\lambda^{u-1}}{(u-1)!}} \tag{6}$$

3 Statistical Properties of Proposed Distribution UDPD-I

3.1 Probability Generating Function:

The probability generating function of the UDPD-I is obtained as

$$\begin{aligned}P_x(s)=E(s^X)&=\sum_{x=0}^{\infty}s^x\frac{e^{-\lambda}\lambda^x}{(1+\theta)x!}+\sum_{x=1}^{\infty}s^x\frac{\theta e^{-\lambda}\lambda^{x-1}}{(1+\theta)(x-1)!}\\&=\frac{e^{-\lambda}}{(1+\theta)}\left(1+\frac{s\lambda}{1!}+\frac{s^2\lambda^2}{2!}+\ldots\right)+\frac{\theta e^{-\lambda}}{(1+\theta)}\left(s+\frac{s^2\lambda}{1!}+\frac{s^3\lambda^2}{2!}+\ldots\right)\\&=\frac{1}{(1+\theta)}e^{\lambda(s-1)}+\frac{\theta}{(1+\theta)}se^{\lambda(s-1)}=\frac{(1+\theta s)}{(1+\theta)}e^{\lambda(s-1)}\end{aligned} \tag{7}$$

3.2 Moment Generating Function

The moment generating function of the UDPD-I is obtained as

$$M_x(t)=E(e^{tX})=\sum_{x=0}^{\infty}e^{tx}p(x;\lambda)=\sum_{x=0}^{\infty}e^{tx}\frac{e^{-\lambda}\lambda^x}{(1+\theta)x!}+\sum_{x=1}^{\infty}e^{tx}\frac{\theta e^{-\lambda}\lambda^{x-1}}{(1+\theta)(x-1)!} \tag{8}$$

After solving the above equation (8), we get

$$M_x(t) = \frac{(1+\theta e^t)}{(1+\theta)} e^{\lambda(e^t-1)} \tag{9}$$

3.3 Cumulant Generating Function

The cumulant generating function of the UDPD-I is given by

$$\begin{aligned}\kappa_x(t) &= \log M_x(t) = \log\left[\frac{(1+\theta e^t)}{(1+\theta)} e^{\lambda(e^t-1)}\right] \\ &= \lambda(e^t-1) + \log\left(1+\theta e^t\right) - \log(1+\theta)\end{aligned} \tag{10}$$

After solving the above equation (10), we get

$$\begin{aligned}\kappa_x(t) = \lambda\left(t+\frac{t^2}{2!}+\frac{t^3}{3!}+\cdots\right) + \theta\left(1+t+\frac{t^2}{2!}+\frac{t^3}{3!}+\cdots\right) - \frac{\theta^2}{2}\left(1+2t+\frac{(2t)^2}{2!}+\frac{(2t)^3}{3!}+\cdots\right) \\ +\frac{\theta^3}{3}\left(1+3t+\frac{(3t)^2}{2!}+\frac{(3t)^3}{3!}+\cdots\right) - \frac{\theta^4}{4}\left(1+4t+\frac{(4t)^2}{2!}+\frac{(4t)^3}{3!}+\cdots\right)+\cdots-\left(\theta-\frac{\theta^2}{2}+\frac{\theta^3}{3}-\cdots\right)\end{aligned} \tag{11}$$

First Cumulant κ_1 is mean, which is given by the coefficient of t in equation number (11).

$$\begin{aligned}\kappa_1 = \mu_1' = \lambda + \left(\theta-\theta^2+\theta^3-\theta^4+\cdots\right) = \lambda + \theta(1+\theta)^{-1} \\ \text{Therefore} \quad E(X) = \kappa_1 = \lambda + \frac{\theta}{1+\theta}\end{aligned} \tag{12}$$

The second cumulant κ_2 is varience μ_2, which is given by the coefficient of $\frac{t^2}{2!}$ in equation number (11)

$$\kappa_2 = V(X) = \mu_2 = \lambda + \frac{\theta}{(1+\theta)^2} \tag{13}$$

$$\kappa_3 = \mu_3 = \lambda + \frac{\theta(1-\theta)}{(1+\theta)^3} \tag{14}$$

$$\kappa_4 = \lambda + \frac{\theta(1-\theta)^2 - 2\theta^2}{(1+\theta)^4} \tag{15}$$

$$\mu_4 = \kappa_4 + 3\kappa_2^2 = \lambda + 3\lambda\left[\lambda + \frac{2\theta}{(1+\theta)^2}\right] + \frac{\theta}{(1+\theta)^2}\left[1 - \frac{3\theta}{(1+\theta)^2}\right] \tag{16}$$

The coefficient of variation is obtained as

$$CV = \frac{SD}{Mean} = \frac{\sqrt{\lambda + \frac{\theta}{(1+\theta)^2}}}{\lambda + \frac{\theta}{1+\theta}} = \frac{\sqrt{\lambda(1+\theta)^2+\theta}}{\lambda(1+\theta)+\theta} \tag{17}$$

The coefficient of skewness and kurtosis is obtained as

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{\left(\lambda + \frac{\theta(1-\theta)}{(1+\theta)^3}\right)^2}{\left(\lambda + \frac{\theta}{(1+\theta)^2}\right)^3} = \frac{\left[\lambda(1+\theta)^3 + \theta(1-\theta)\right]^2}{\left[\lambda(1+\theta)^2+\theta\right]^3} \tag{18}$$

and

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{\lambda + 3\lambda\left[\lambda + \frac{2\theta}{(1+\theta)^2}\right] + \frac{\theta}{(1+\theta)^2}\left[1 - \frac{3\theta}{(1+\theta)^2}\right]}{\left[\lambda + \frac{\theta}{(1+\theta)^2}\right]^2} = \frac{(1+\theta)^2\left[\lambda(1+3\lambda)(1+\theta)^2 + \theta(1+6\lambda)\right] - 3\theta^2}{\left[\lambda(1+\theta)^2+\theta\right]^2} \tag{19}$$

Now we have Fisher's Index of dispersion is as follows

$$\gamma = \frac{Varience}{Mean} = \frac{\lambda + \frac{\theta}{(1+\theta)^2}}{\lambda + \frac{\theta}{1+\theta}} = \frac{\lambda(1+\theta)^2 + \theta}{(1+\theta)\left[\lambda(1+\theta)+\theta\right]} \tag{20}$$

From the above expression (20) it is clear that Fisher's index of dispersion is less than 1 means the distribution is under-dispersed for every value of λ and θ.

3.4 Characteristics Function

Characteristics function can be obtain by replacing for in moment generating function given in equation number (9), we get

$$\Phi_x(t) = \frac{(1+\theta e^{it})}{(1+\theta)} e^{\lambda(e^{it}-1)} \tag{21}$$

Raw moments are computed from the moment generating function given in equation (9) by differentiating with respect to t and equating with zero as follows

$$\mu_r' = \frac{\partial^r}{\partial t^r} M_X(t)|_{t=0}$$

$$\mu_1' = \lambda + \frac{\theta}{1+\theta}; \quad \mu_2' = \lambda^2 + \frac{\lambda(3\theta+1)+\theta}{1+\theta}$$

$$\mu_3' = \lambda^3 + \frac{2\lambda^2(2\theta+1)+\lambda(9\theta+2)+\theta}{1+\theta}; \quad \mu_4' = \lambda^4 + \frac{\lambda^3(8\theta+5)+\lambda^2(21\theta+6)+\lambda(21\theta+3)+\theta}{1+\theta} \tag{22}$$

3.5 Reoccurrence Relation

For obtaining the probability of different value of x we need reoccurrence relation. The reoccurrence relation of the UDPD-I is as follows

$$p(x+1) = \frac{\lambda\left[\lambda+\theta(x+1)\right]}{\left[(\lambda+\theta x)(x+1)\right]} p(x) \tag{23}$$

and $p(0) = \frac{e^{-\lambda}}{(1+\theta)}$; once $p(0)$ is obtained, we can easily obtain $p(1), p(2), p(3), \cdots$ with the help of (23) and so on.

4 Entropy

Entropy is a measure of uncertainty. It is easily seen that the entropy Δ associated with UDPD-I is given by

$$\begin{aligned}
\Delta &= -\sum_{k=0}^{\infty} p(k;\theta,\lambda)\log p(k;\theta,\lambda) = -\sum_{k=0}^{\infty} p(k;\theta,\lambda)\log\left[\frac{e^{-\lambda}}{1+\theta}\frac{\lambda^{k-1}(\lambda+\theta k)}{k!}\right] \\
&= -\log\left(\frac{e^{-\lambda}}{1+\theta}\right) - \log\lambda\sum_{k=0}^{\infty}(k-1)p(k;\theta,\lambda) - \sum_{k=0}^{\infty}\log(\lambda+\theta k)p(k;\theta,\lambda) + \sum_{k=0}^{\infty}\log(k!)p(k;\theta,\lambda) \\
&= -\log\left(\frac{e^{-\lambda}}{1+\theta}\right) - \left(\lambda - \frac{1}{1+\theta}\right)\log\lambda - \left(\frac{e^{-\lambda}}{1+\theta}\right)\Omega_{k,n}(\lambda,\theta) + \left(\frac{e^{-\lambda}}{1+\theta}\right)\Phi_k(\lambda,\theta)
\end{aligned} \tag{24}$$

where $\Omega_{k,n}(\lambda,\theta) = \sum_{k=0}^{\infty}\sum_{n=0}^{\infty}\frac{(-1)^{n-1}}{n}\left(\frac{\theta k}{\lambda}\right)^n\frac{\lambda^{k-1}(\lambda+\theta k)}{k!}$ and $\Phi_k(\lambda,\theta) = \sum_{k=0}^{\infty}\frac{\lambda^{k-1}(\lambda+\theta k)}{k!}\log(k+1)!$

5 Estimation of Parameters

5.1 Method of Moments

The UDPD-I has two parameters then at least two moments are required to get the estimates of the parameters. From equation (12) and (13) we have

$$V(X) = E(X) - \frac{\theta}{1+\theta} + \frac{\theta}{(1+\theta)^2} = E(X) - \frac{\theta^2}{(1+\theta)^2}$$

$$\frac{\theta^2}{(1+\theta)^2} = E(X) - V(X) = K \quad \text{(say)} \tag{25}$$

From the above equation (25) we find that the mean is greater than the variance because $\frac{\theta^2}{(1+\theta)^2} > 0$ because $\theta > 0$ i.e. the UDPD-I is under-dispersed.

Also, we have from equation (25)

$$\theta^2(K-1) + 2K\theta + K = 0 \quad \text{where} \quad K = E(X) - V(X) \tag{26}$$

The above equation (26) is a quadratic equation of θ and can be estimated easily. The which can be obtained by substituting the value of $\hat{\theta}$ in equation (27).

$$\hat{\lambda} = E(X) - \frac{\hat{\theta}}{1+\hat{\theta}} \tag{27}$$

5.2 Maximum Likelihood Estimation

Let $X = (x_1, x_2, x_3, \cdots, x_n)$ be a random sample of size n from the UDPD-I(λ, θ) distribution. Thus from the equation (1) the likelihood function, L of the UDPD-I can be written as

$$L = \prod_{i=1}^{n} p(x_i; \lambda, \theta) = \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{x_i - 1}(\lambda + \theta x_i)}{(1+\theta)x_i!} = \frac{e^{-n\lambda}}{(1+\theta)^n}\lambda^{\sum_{i=0}^{n}(x_i-1)} \prod_{i=1}^{n}\left(\frac{\lambda+\theta x_i}{x_i!}\right) \tag{28}$$

Now, the log likelihood function is given by

$$\log L = -n\lambda - n\log(1+\theta) + \log\lambda \sum_{i=1}^{n}(x_i - 1) + \sum_{i=1}^{n}\log\left(\frac{\lambda+\theta x_i}{x_i!}\right) \tag{29}$$

Differentiating the log likelihood function (29) partially with respect to λ and θ equating with zero then we get

$$\frac{\partial}{\partial\lambda}\log L = -n + \frac{1}{\lambda}\sum_{i=1}^{n}(x_i - 1) + \sum_{i=1}^{n}\frac{1}{(\lambda+\theta x_i)} = 0$$

$$n\left(\frac{\lambda+1}{\lambda}\right) = \frac{n\bar{x}}{\lambda} + \sum_{i=1}^{n}\frac{1}{(\lambda+\theta x_i)} \tag{30}$$

$$\frac{\partial}{\partial\theta}\log L = -\frac{n}{1+\theta} + \sum_{i=1}^{n}\left(\frac{x_i}{\lambda+\theta x_i}\right) = 0$$

$$\frac{n}{1+\theta} = \sum_{i=1}^{n}\left(\frac{x_i}{\lambda+\theta x_i}\right) \tag{31}$$

The above equations (30) and (31) are non-linear equations and cannot be solved analytically, so we can obtain the solutions of these equations by using existing iterative procedures.

6 Applications

The application of UDPD-I and II is discussed with the some real data sets. The first data set is relating to the number of outbreaks of strikes in the UK coal mining industry in successive four week periods in the years 1948-1959, Kendall [31] and second date set is the data of number of sperm in the egg related to fertilization in a sea-urchin's egg by Morgan [32]. The third data set is about the word length in the Turkish poem Gidisat by Ercüment Behzat Lâv available in Wimmer et al. [33]. Here, the count for the response x is treated as $(x-1)$. These data sets exhibit under dispersion with a Fisher's index of dispersion less than one. The suitability of UDPD-I and II is compared fitted along with Poisson distribution, Com-Poisson distribution given by Efron [5] and generalized Poisson distribution proposed by Consul and Jain [12]. From Table 1, we found that the UDPD-I and II show a significant departure from the Poisson and COM-Poisson distribution and a considerable improvement from Generalized Poisson distribution. The Chi-square value for the UDPD-I and II is lower among all the distribution discussed. Table 2, also represent that the UDPD-I and II gives better fit than other alternative distributions and has smallest Chi-square. Expected frequencies given by UDPD-I and II are very close to observed one than expected frequencies given by other alternative distributions. From the Table 3 it is clear that the UDPD-I and II are a competent competitor of the other distributions considered here.

Table 1: Expected frequencies and value of chi-square with degree of freedom from fitted distributions to the number outbreak of strikes data of Kendall [31].

Number of outbreaks	Observed frequencies	Distributions				
		Poisson	COM-Poisson	GPD	UDPD-I	UDPD-II
0	46	57.76	68.33	49.22	47.20	52.39
1	76	57.39	59.35	66.70	71.74	64.23
2	24	28.51	22.33	32.57	29.24	29.53
3	9	9.44	5.15	6.90	6.64	8.05
4	1	2.90	0.84	0.61	1.18	1.80
Total	156	156.00	156.00	156.00	156.00	156.00
$\hat{\lambda}$		0.9936	0.8899	-0.1610	0.4872	0.613
$\hat{\theta}$		-	1.207	1.154	1.026	-
χ^2		9.59	14.78	4.59	1.83	3.97
Degree of freedom		2	1	1	1	2
$p-$value		0.008	0.000	0.032	0.176	0.137

Table 2: Expected frequencies and value of chi-square with degree of freedom from fitted distributions to the number of sperm in egg data from R. W. Morgan [32].

Number of sperm in egg	Observed frequencies	Distributions				
		Poisson	COM-Poisson	GPD	UDPD-I	UDPD-II
0	28	37.32	50.51	29.02	27.29	35.12
1	44	28.46	24.39	40.93	45.06	31.68
2	7	10.85	4.60	10.04	7.05	10.72
3	1	3.37	0.50	0.01	0.60	2.49
Total	80	80.00	80.00	80.00	80.00	80.00
$\hat{\lambda}$		0.7625	0.5845	-0.3299	0.1646	0.451
$\hat{\theta}$		-	1.356	1.014	1.487	-
χ^2		13.53	27.45	0.71	0.06	8.285
Degree of freedom		1	-	-	-	1
$p-$value		0.000	-	-	-	0.004

Table 3: Expected frequencies and value of chi-square with degree of freedom from fitted distributions to the number of word length in Turkish poem, Wimmer et al. [33].

Number of word length	Observed frequencies	Distributions				
		Poisson	COM-Poisson	GPD	UDPD-I	UDPD-II
1	64	80.67	59.69	61.96	61.09	65.29
2	131	127.94	141.87	135.34	143.83	139.72
3	122	101.46	118.70	121.25	114.39	112.12
4	61	53.64	53.92	57.43	52.67	53.32
5	13	21.27	15.88	15.46	16.90	17.83
≥ 6	3	9.02	3.94	2.56	5.12	5.73
Total	394	394.00	394.00	394.00	394.00	394.00
$\hat{\lambda}$		2.586	2.377	1.850	1.013	1.070
$\hat{\theta}$		-	1.506	-0.166	1.341	-
χ^2		15.93	2.91	0.66	4.89	5.15
Degree of freedom		4	2	2	3	4
$p-$value		0.000	0.234	0.719	0.181	0.272

7 Conclusions

This paper studied a well known and widely used Poisson distribution and weighted as well as mixture distribution. The various statistical properties of the UDPD-I is discussed and also applied on real data sets. It performs better than various other distributions for two data sets and for one data set UDPD-I shows that this may be used as an alternative. Generalization proposed here is simple and can be used to handle various real data sets with complex structure. If both the parameters are same as λ then the UDPD-I become a single parameter distribution UDPD-II with average $\frac{\lambda(2+\lambda)}{(1+\lambda)}$ is again useful to model a phenomenon.

Acknowledgement

Authors are grateful to the anonymous referees for their constructing suggestions to improve the quality of paper.

References

[1] R. A. Fisher, The effect of methods of ascertainment upon the estimation of frequencies, Annals of eugenics, **6**(1), 13-25, (1934).
[2] C. R. Rao, On discrete distributions arising out of methods of ascertainment, Sankhyā: The Indian Journal of Statistics, Series A, 311-324, (1965).
[3] G. P. Patil and C. R. Rao, Weighted distributions and size-biased sampling with applications to wildlife populations and human families, Biometrics, **34**(2), 179-189, (1978).
[4] G. P. Patil, C. R. Rao and M. V. Ratnaparkhi, On discrete weighted distributions and their use in model choice for observed data, Communications in Statistics-Theory and Methods, **15**(3), 907-918, (1986).
[5] B. Efron, Double exponential families and their use in generalized linear regression, Journal of the American Statistical Association, **81**(395), 709-721, (1986).
[6] A. C. Cameron and P. Johansson, Count data regression using series expansions: with applications, Journal of Applied Econometrics, **12**(3), 203-223, (1997).
[7] M. S. Ridout and P. Besbeas, An empirical model for underdispersed count data, Statistical Modelling, **4**(1), 77-89, (2004).
[8] J. D. Castillo and M. Pérez-Casany, Weighted Poisson distributions for overdispersion and underdispersion situations, Annals of the Institute of Statistical Mathematics, **50**(3), 567-585, (1998).
[9] J. D. Castillo and M. Pérez-Casany, Overdispersed and underdispersed Poisson generalizations, Journal of Statistical Planning and Inference, **134**(2), 486-500, (2005).
[10] C. C. Kokonendji, D. Mizére and N. Balakrishnan, Connections of the Poisson weight function to overdispersion and underdispersion, Journal of Statistical Planning and Inference, **138**(5), 1287-1296, (2008).
[11] N. Balakrishnan, and T. J. Kozubowski, A class of weighted Poisson processes, Statistics & probability letters, **78**(15), 2346-2352, (2008).
[12] P. C. Consul and G. C. Jain, A generalization of the Poisson distribution, Technometrics, **15**(4), 791-799, (1973).
[13] P. C. Consul, Generalized Poisson distribution: properties and applications, Marcel Dekker, New York, (1989).
[14] N. L. Johnson, S. Kotz and A. W. Kemp, Univariate discrete distributions, 2nd ed. John Wiley & Sons Inc; (1992).

[15] M. Sankaran, The discrete Poisson-Lindley distribution, Biometrics, **26**(1), 145-149, (1970).
[16] M. E. Ghitany and D. K. Al-Mutairi, Estimation methods for the discrete Poisson?Lindley distribution, Journal of Statistical Computation and Simulation, **79**(1), 1-9, (2009).
[17] R. Shanker, F. Hagos, S. Sujatha and Y. Abrehe, On zero-truncation of Poisson and Poisson-Lindley distributions and their applications, Biometrics & Biostatistics International Journal, **2**(6), 1-14, (2015).
[18] B. P. Singh, S. Dixit and T. K. Roy, On some compounded Poisson distributions with applications to the pattern of number of child deaths. Journal of Probability and Statistical Sciences, **14**(2), 203-209, (2016).
[19] R. Shanker and A. Mishra, A quasi Poisson-Lindley distribution, Journal of Indian Statistical Association, **54**(1&2), 113-125, (2016).
[20] R. Shanker and A. Mishra, A two-parameter Poisson-Lindley distribution, International journal of Statistics and Systems, **9**(1), 79-85, (2014).
[21] R. Shanker and A. L. Tekie, A new quasi Poisson-Lindley distribution, International Journal of Statistics and Systems, **9**(1), 87-94, (2014).
[22] M. E. Ghitany, D. K. Al-Mutairi and S. Nadarajah, Zero-truncated Poisson-Lindley distribution and its application, Mathematics and Computers in Simulation, **79**(3), 279-287, (2008).
[23] M. Borah and N. Deka, A study on the inflated Poisson Lindley distribution. Journal of the Indian Society of Agricultural Statistics, **54**(3), 317-323, (2001).
[24] B. P. Singh, N. K. Singh and S. Dixit, Estimation of parameters of inflated Poisson-Lindley distribution for adult out-migration, Journal of Institute of Science and Technology, Tribhuvan University, Nepal, **20**(2), 6-10, (2015).
[25] M. E. Ghitany and D. K. Al-Mutairi, Size-biased Poisson-Lindley distribution and its application, METRON-International Journal of Statistics, **66**(3), 299-311, (2008).
[26] R. Shanker, F. Hagos and Y. Abrehe, On size biased Poisson Lindley distribution and its applications to model thunderstorms, American Journal of Mathematics and Statistics, **5**(6), 354-360, (2015).
[27] E. Mahmoudi and H. Zakerzadeh, Generalized Poisson-Lindley distribution, Communications in Statistics-Theory and Methods, **39**(10), 1785-1798, (2010).
[28] H. Zakerzadeh and A. Dolati, Generalized Lindley distribution, 13-25, (2009).
[29] R. Shanker and K. K. Shukla, A generalized size-biased Poisson-Lindley distribution and its applications to model size distribution of freely-forming small group, Journal of Scientific Research, **10**(2), 145-157, (2018).
[30] B. P. Singh, S. Singh, and U. D. Das, A general class of new continuous mixture distribution and application, Journal of Mathematical and Computional Science, **11**(1), 585-602, (2020).
[31] M. G. Kendall, Natural law in social sciences, Journal of the Royal Statistical Society, Series A, 124-119, (1961).
[32] R. W. Morgan, Some stochastic models to describe the fertilization of an egg, Journal of the Royal Statistical Society: Series C (Applied Statistics), **24**(1), 137-138, (1975).
[33] G. Wimmer, R. Köhler, R. Grotjahn and G. Altmann, Towards a theory of word length distribution, Journal of Quantitative Linguistics, **1**(1), 98-106, (1994).