

An Alternative to Intervened Poisson Distribution for Prevalence Reduction

Brijesh P. Singh

Assistant Professor, Faculty of Commerce & DST-CIMS, Banaras Hindu University, Varanasi.

Shweta Dixit and Upasana Shukla

SRF, Department of Statistics, Institute of Science, Banaras Hindu University, Varanasi.

Abstract

In this paper an attempt has been made to propose an alternative of truncated and intervened Poisson distribution, having two parameters and named as Bounded Poisson (BP) distribution. To estimate the parameters, method of moment and first cell frequency method have been used. To check the suitability of the model it has been applied on real data set used by Dahiya et al. (1973). Proposed model provides a good fitting to the data under consideration.

Introduction

Many authors have worked on the poisson distribution and the readers are referred to Johnson and Kotz (1969) and Haight (1965). The Poisson distribution is an appropriate mathematical model for studying count data model like haemocytometer counts of yeast cells per square, the number of noxious weed seed per unit of field and the number of defects of per unit of a manufactured product, number of births/deaths in certain duration (Cohen, 1954) and therefore biologists, agronomists, quality control engineers, social scientists, medical scientists etc have been in attraction towards this distribution. It is widely used in medical field to explain the occurrence rare events. Cohen (1960) introduced positive Poisson distribution to describe a chance mechanism whose observational apparatus (i.e. diagnosis) becomes active only when at least one event occurs. Further (Singh, 1978) obtained a numerical example to illustrate the statistical application of the positive Poisson distribution in such situations. Consul (1989)

Corresponding author: Brijesh P. Singh, Assistant Professor, Faculty of Commerce & DST-CIMS, Banaras Hindu University, Varanasi.

and Consul et al. (1989) discussed about generalized Poisson distribution and its truncated version.

A new type of distribution in which the idea of intervention has been incorporated has received much attention in the literature. These types of distributions provide information on the effectiveness of various preventive actions taken in several areas of scientific research. (Shanmugam, 1985) introduced an intervened Poisson distribution (*IPD*) as a replacement for positive Poisson distribution where some intervention process changes the mean of rare events. An advantage of the IPD is that it provides information on how effective various preventive actions taken by health service agents, where positive Poisson fails. The IPD is widely used in several areas such as reliability analysis, queuing problems, epidemiological studies, etc and has been further studied by Shanmugam (1992); Huang and Fung (1989) and Dhanavanthan (1998, 2000). Scollnik (2006) developed the intervened generalized Poisson distribution and Kumar and Shibu (2011) considered a modified version of IPD. The advantage of this distribution over the IPD is that it stretches the probability in all directions so that clustering of probabilities at initial values of operating mechanism is overlooked. In the present paper we derive a new distribution which is two parameter extension of a ZTPD and we termed it as Bounded Poisson distribution (*BPD*). The proposed distribution is equally efficient as IPD. To estimate parameters we have used moment estimation procedure as well as method of first moment and first cell frequency.

Model

Let Y be a random variable which denotes the number of active cholera cases in a household. Here we are not considering $Y=0$ because medical assistance should be provided only active cases. Thus Y is a random variable which denotes number of cholera cases in a household and follows zero truncated Poisson distribution as given below:

$$P(Y = y) = \frac{(e^\theta - 1)^{-1} \theta^y}{y!} ; y = 1, 2, 3, \dots$$

where $\theta > 0$ is called *incidence parameter*. After applying health programs to prevent increasing cholera patients, the incidence parameter θ reduces by ϕ amount resulting in parameter $(\theta - \phi)$ where $0 \leq \phi \leq \theta$. This ϕ

is called as reduction parameter as it tells the amount of reduction in incidence parameter after implication of health program. If ϕ equals to 0 it means health programs are not useful as they are not controlling the spreading of disease, whereas it ϕ equals θ it means the health program is 100 percent effective as the disease is completely controlled in the population. Thus Z is a random variable representing the number of cholera patients in a household after applying health program and its distribution is given by

$$P(Z = z) = \frac{e^{-(\theta-\phi)} (\theta-\phi)^z}{z!}; Z = 0, 1, 2, 3, \dots$$

Thus we find Y and Z are stochastically independent random variables. Further we assume that we have only data of number of cholera patients in a household after the applying health program. Thus the final distribution becomes $X=Y+Z$ and we call it as Bounded Poisson distribution. The density function of BPD is given below:

$$X = Y + Z$$

$$P(X = x) = \sum_{l=0}^{x-1} P(Y = x-l).P(Z = l | Y = x-l)$$

$$P(x) = (e^\theta - 1)^{-1} e^{-(\theta-\phi)} \sum_{l=0}^{x-1} \frac{\theta^l}{(x-l)!} \cdot \frac{(\theta-\phi)^l}{l!}$$

Thus the probability density function of X is given by

$$p(x) = \frac{(2\theta - \phi)^x - (\theta - \phi)^x}{(e^\theta - 1)e^{(\theta-\phi)} x!}; x = 1, 2, 3, \dots \quad (1)$$

If $\phi = 0$ then the above probability distribution becomes a sum of simple poisson and truncated poisson distribution. When $\phi = \theta$ then the distribution becomes truncated poisson distribution.

If we put $\left(1 - \frac{\phi}{\theta}\right) = \rho$ in above probability model then we the distribution proposed by Shanmugam (1985).

Estimation

In the present paper we have considered two different estimation procedures and these are discussed below in details.

Method of Moments

In the present paper we have considered method of moment estimation procedure. The moments are calculated as follows

$$E(x) = \frac{\theta e^\theta}{(e^\theta - 1)} + (\theta - \phi) \quad (2)$$

$$E(x^2) = \frac{1}{(e^\theta - 1)} \left[(2\theta - \phi)^2 e^\theta - (\theta - \phi)^2 \right] + \frac{\theta e^\theta}{(e^\theta - 1)} + (\theta - \phi) \quad (3)$$

With the help of above equations we get a function of θ as

$$\Rightarrow E(x) - Var(x) - \frac{\theta^2 e^\theta}{(e^\theta - 1)^2} = 0 \quad (4)$$

So the value of parameter θ can be obtained by solving above equation. Once value of parameter θ is estimated ϕ can be estimate using first moment given in equation 2. From the equation 4 it is clear that the variance of this distribution is lesser than the mean thus this distribution is under dispersed, although the under dispersion is significant only for small values of θ , as θ increases variance approaches to the mean of the distribution.

Estimation of Parameters by First Cell Frequency

Putting $x=1$ in eq. (1) we get

$$p(1) = \frac{\theta e^\phi}{e^\theta (e^\theta - 1)}$$

Since we know first moment i.e. mean

$$E(x) = \frac{\theta e^\theta}{(e^\theta - 1)} + (\theta - \phi) \Rightarrow \phi = \frac{\theta e^\theta}{(e^\theta - 1)} + \theta - E(x)$$

Therefore putting the value of ϕ in $p(1)$, we get

$$\begin{aligned} \Rightarrow p(1) &= \frac{\theta e^{\left\{\frac{\theta e^\theta}{(e^\theta - 1)} + \theta - E(x)\right\}}}{e^\theta (e^\theta - 1)} = \frac{\theta \left\{e^{\frac{\theta e^\theta}{(e^\theta - 1)}} e^\theta e^{-E(x)}\right\}}{e^\theta (e^\theta - 1)} = \frac{\theta \left\{e^{\frac{\theta}{(1 - e^{-\theta})}} e^{-E(x)}\right\}}{(e^\theta - 1)} \\ \Rightarrow p(1) - \frac{\theta \left\{e^{\frac{\theta}{(1 - e^{-\theta})}} e^{-E(x)}\right\}}{(e^\theta - 1)} &= 0 \end{aligned}$$

After solving this equation we can have the estimate of θ . Then ϕ can be obtained using either $p(1)$ or $E(x)$.

Application of the Model

To see the effect of model we have used data from Dahiya et al. (1973) which was also used by Shanmugam (1985). In that paper data on epidemic of cholera in a village in India has been reported. In this table, x represents the number of cholera cases in a house and f_x represents the number of houses with x cases of cholera. For the sample size $n=55$, the mean and variance, are 1.56 and 0.58 respectively.

Table 1: Expected frequencies due to various models

x	1	2	3	4+	Chi sq	Estimate of parameters
Observed f_x	32	16	6	1		
ZTPD	32.63	15.72	5.03	1.62	0.44	$\theta=0.967$
IPD	31.71	17.16	4.94	1.19	0.34	$\theta=0.480$ $\rho=0.627$
BPD (MM)	31.73	17.14	4.95	1.18	0.33	$\theta=0.490$ $\phi=0.190$
BPD (FCF)	31.92	16.86	4.96	1.26	0.32	$\theta=0.620$ $\phi=0.402$

Table 2: Expected frequency due to proposed models for the total number of any abortion to the female in Uttar Pradesh

x	1	2	3	4+	Chi sq	Estimate of parameters
Observed f_x	33	14	4	1		
BPD (MM)	32.77	14.50	3.84	0.89	0.039	$\theta=0.680$ $\phi=0.577$
BPD (FCF)	32.68	14.68	3.80	0.84	0.076	$\theta=0.610$ $\phi=0.466$

The proposed model is close to the real life problems and the estimate of ϕ provides us idea about the amount of the reduction in incidence. Table 1 show that the expected frequencies of various model such as zero truncated Poisson, intervened Poisson and proposed model. Expected frequency for proposed model is obtained by two methods i.e. method of moments and method of first cell frequency and first moment. Both methods provide reasonable values for the observed values. The estimate of incidence of cholera is 0.49 however after intervention of the health program it becomes 0.3 due to method of moment but it is 0.62 and 0.22 in case of method of first cell frequency and first moment. To check the suitability of the model another data has been taken from second round of National Family Health Survey for the number of any type of abortion (spontaneous and induced) in Uttar Pradesh a traditional state of India where abortion is treated as social taboos. Table 2 reveals that the total number of any abortion to the female (age ≥ 35 years and age at marriage ≥ 20) of Uttar Pradesh with average number of abortion is 1.48 and variance is 0.52. Chi square value point out that model is suitable for the data under consideration. Major drawback of the model that this is very sensitive for the variance of the data and also we cannot know how much proportion of the population is facilitating by the intervention program thus there is a need to develop a model keeping this into mind.

References

- [1]. Cohen, A. C. (1954). Estimation of the Poisson Parameter from Truncated Samples and from Censored Samples. *Journal of the American Statistical Association*, 49, 158-168.
- [2]. Cohen, A. C. (1960). Estimating parameters in a conditional Poisson distribution. *Biometrics*, 16, 203-211.
- [3]. Consul, P. C. (1989). Generalized Poisson Distributions: Properties and Applications. Marcel Dekker Inc. New York.
- [4]. Consul, P. C. and Famoye, F. (1989). The truncated generalized Poisson distribution and its estimation. *Communication in Statistics-Theory and Methods*, 18(10): 3635-3648.
- [5]. Dahiya, R. C. and Gross, A. J. (1973). Estimating the zero class from a truncated Poisson sample. *Journal of the American Statistical Association*, 68, 731-733.
- [6]. Dhanavanthan, P. (1998). Compound intervened Poisson distribution. *Biometrical Journal*, 40, 641-646.
- [7]. Dhanavanthan, P. (2000). Estimation of the parameters of compound intervened Poisson distribution. *Biometrical Journal*, 42, 315-320.
- [8]. Haight, F (1965). Handbook of the Poisson distribution. John Wiley and Sons, New York.
- [9]. Huang, M. and Fung, K.Y. (1989). Intervened truncated Poisson distribution. *Sankhya*, Series B, 51, 302-310.
- [10]. Johnson, N. L.; Kemp, A.W. and Kotz, S. (2005). Univariate Discrete Distributions. John Wiley and Sons, New York.
- [11]. Kumar, C. S. and Shibu, D. S. (2011). Modified intervened Poisson distribution. *Statistica*, 71, 489-499.
- [12]. Shanmugam, R. (1985). An intervened Poisson distribution and its medical application. *Biometrics*, 41, 1025-1029.
- [13]. Shanmugam R. (1992). An inferential procedure for the Poisson intervention parameter. *Biometrics*, 48, 559-565.
- [14]. Singh, J. (1978). A characterization of positive Poisson distribution and its application. *SIAM Journal of Applied Mathematics*, 34, 545-48.
- [15]. Scollnik, D. P. M. (2006). On the intervened generalized Poisson distribution. *Communication in Statistics-Theory and Methods*, 35:6, 953-963.

Appendix

The moments are calculated as follows

$$E(x) = \sum_{x=1}^{\infty} x p(x) = \sum_{x=1}^{\infty} x \frac{(2\theta - \phi)^x - (\theta - \phi)^x}{(e^\theta - 1)e^{(\theta - \phi)} x!} = \frac{1}{(e^\theta - 1)e^{(\theta - \phi)}} \sum_{x=1}^{\infty} \frac{(2\theta - \phi)^x - (\theta - \phi)^x}{(x-1)!}$$

$$E(x) = \frac{1}{(e^\theta - 1)e^{(\theta - \phi)}} \left[\frac{(2\theta - \phi)^1 - (\theta - \phi)^1}{0!} + \frac{(2\theta - \phi)^2 - (\theta - \phi)^2}{1!} + \frac{(2\theta - \phi)^3 - (\theta - \phi)^3}{2!} + \dots \right]$$

$$E(x) = \frac{1}{(e^\theta - 1)e^{(\theta - \phi)}} \left[\left\{ \frac{(2\theta - \phi)^1}{0!} + \frac{(2\theta - \phi)^2}{1!} + \frac{(2\theta - \phi)^3}{2!} + \dots \right\} - \left\{ \frac{(\theta - \phi)^1}{0!} + \frac{(\theta - \phi)^2}{1!} + \frac{(\theta - \phi)^3}{2!} + \dots \right\} \right]$$

$$E(x) = \frac{1}{(e^\theta - 1)e^{(\theta - \phi)}} \left[(2\theta - \phi) \left\{ 1 + \frac{(2\theta - \phi)^1}{1!} + \frac{(2\theta - \phi)^2}{2!} + \dots \right\} - (\theta - \phi) \left\{ 1 + \frac{(\theta - \phi)^1}{1!} + \frac{(\theta - \phi)^2}{2!} + \dots \right\} \right]$$

$$E(x) = \frac{1}{(e^\theta - 1)e^{(\theta - \phi)}} \left[(2\theta - \phi)e^{(2\theta - \phi)} - (\theta - \phi)e^{(\theta - \phi)} \right]$$

$$\text{Thus } E(x) = \frac{1}{(e^\theta - 1)} \left[(2\theta e^\theta - \theta) - \phi(e^\theta - 1) \right] = \frac{(2\theta e^\theta - \theta) - \phi}{(e^\theta - 1)}$$

$$E(x) = \theta \left[\frac{e^\theta + e^\theta - 1}{(e^\theta - 1)} \right] - \phi \Rightarrow E(x) = \frac{\theta e^\theta}{(e^\theta - 1)} + (\theta - \phi)$$

Now the second moment is

$$E(x^2) = \sum_{x=1}^{\infty} x^2 p(x) = \sum_{x=1}^{\infty} x^2 \frac{(2\theta - \phi)^x - (\theta - \phi)^x}{(e^\theta - 1)e^{(\theta - \phi)} x!}$$

$$E(x^2) = \frac{1}{(e^\theta - 1)e^{(\theta - \phi)}} \sum_{x=1}^{\infty} \{x(x-1) + x\} \frac{(2\theta - \phi)^x - (\theta - \phi)^x}{x!}$$

$$E(x^2) = \frac{1}{(e^\theta - 1)e^{(\theta - \phi)}} \sum_{x=1}^{\infty} x(x-1) \frac{(2\theta - \phi)^x - (\theta - \phi)^x}{x!} + \frac{1}{(e^\theta - 1)e^{(\theta - \phi)}} \sum_{x=1}^{\infty} x \frac{(2\theta - \phi)^x - (\theta - \phi)^x}{x!}$$

$$E(x^2) = \frac{1}{(e^\theta - 1)e^{(\theta - \phi)}} \sum_{x=2}^{\infty} \frac{(2\theta - \phi)^x - (\theta - \phi)^x}{(x-2)!} + E(x)$$

$$E(x^2) = I + E(x)$$

$$I = \frac{1}{(e^\theta - 1)e^{(\theta - \phi)}} \sum_{x=2}^{\infty} \frac{(2\theta - \phi)^x - (\theta - \phi)^x}{(x-2)!}$$

$$I = \frac{1}{(e^\theta - 1)e^{(\theta - \phi)}} \left\{ \frac{(2\theta - \phi)^2 - (\theta - \phi)^2}{0!} + \frac{(2\theta - \phi)^3 - (\theta - \phi)^3}{1!} + \frac{(2\theta - \phi)^4 - (\theta - \phi)^4}{2!} + \dots \right\}$$

$$I = \frac{1}{(e^\theta - 1)e^{(\theta - \phi)}} \left[\left\{ \frac{(2\theta - \phi)^2}{0!} + \frac{(2\theta - \phi)^3}{1!} + \frac{(2\theta - \phi)^4}{2!} + \dots \right\} - \left\{ \frac{(\theta - \phi)^2}{0!} + \frac{(\theta - \phi)^3}{1!} + \frac{(\theta - \phi)^4}{2!} + \dots \right\} \right]$$

$$I = \frac{1}{(e^\theta - 1)e^{(\theta - \phi)}} \left[\left\{ \frac{(2\theta - \phi)^2}{0!} + \frac{(2\theta - \phi)^3}{1!} + \frac{(2\theta - \phi)^4}{2!} + \dots \right\} - \left\{ \frac{(\theta - \phi)^2}{0!} + \frac{(\theta - \phi)^3}{1!} + \frac{(\theta - \phi)^4}{2!} + \dots \right\} \right]$$

$$I = \frac{1}{(e^\theta - 1)e^{(\theta - \phi)}} \left[(2\theta - \phi)^2 \left\{ 1 + \frac{(2\theta - \phi)^1}{1!} + \frac{(2\theta - \phi)^2}{2!} + \dots \right\} - (\theta - \phi)^2 \left\{ 1 + \frac{(\theta - \phi)^1}{1!} + \frac{(\theta - \phi)^2}{2!} + \dots \right\} \right]$$

$$I = \frac{1}{(e^\theta - 1)e^{(\theta - \phi)}} \left[(2\theta - \phi)^2 e^{(2\theta - \phi)} - (\theta - \phi)^2 e^{(\theta - \phi)} \right]$$

$$I = \frac{1}{(e^\theta - 1)} \left[(2\theta - \phi)^2 e^\theta - (\theta - \phi)^2 \right]$$

$$\therefore E(x^2) = \frac{1}{(e^\theta - 1)} \left[(2\theta - \phi)^2 e^\theta - (\theta - \phi)^2 \right] + E(x)$$

$$\therefore E(x^2) - E(x) = \frac{1}{(e^\theta - 1)} \left[(2\theta - \phi)^2 e^\theta - (\theta - \phi)^2 \right]$$

$$E(x^2) - E(x) = \frac{1}{(e^\theta - 1)} \left[e^\theta \{ \theta + (\theta - \phi) \}^2 - (\theta - \phi)^2 \right]$$

$$E(x^2) - E(x) = \frac{1}{(e^\theta - 1)} \left[e^\theta \{ \theta^2 + (\theta - \phi)^2 + 2\theta(\theta - \phi) \} - (\theta - \phi)^2 \right]$$

$$E(x^2) - E(x) = \frac{1}{(e^\theta - 1)} \left[e^\theta \theta^2 + e^\theta (\theta - \phi)^2 + 2e^\theta \theta (\theta - \phi) - (\theta - \phi)^2 \right]$$

$$E(x^2) - E(x) = \frac{1}{(e^\theta - 1)} \left[e^\theta \theta^2 + 2e^\theta \theta (\theta - \phi) + (\theta - \phi)^2 (e^\theta - 1) \right]$$

$$E(x^2) - E(x) = \frac{1}{(e^\theta - 1)} \left[e^\theta \theta^2 + 2e^\theta \theta^2 - 2e^\theta \theta \phi + (\theta - \phi)^2 (e^\theta - 1) \right]$$

$$E(x^2) - E(x) = \frac{1}{(e^\theta - 1)} \left[\theta e^\theta (3\theta - 2\phi) + (\theta - \phi)^2 (e^\theta - 1) \right]$$

$$E(x^2) - E(x) = \frac{\theta e^\theta (3\theta - 2\phi)}{(e^\theta - 1)} + (\theta - \phi)^2$$

We know that

$$E(x) = \frac{\theta e^\theta}{(e^\theta - 1)} + (\theta - \phi) \quad \Rightarrow (\theta - \phi) = E(x) - \frac{\theta e^\theta}{(e^\theta - 1)}$$

Putting the value of $(\theta - \phi)$ in equation, we get

$$E(x^2) - E(x) = \frac{\theta e^\theta (3\theta - 2\phi)}{(e^\theta - 1)} + \left[E(x) - \frac{\theta e^\theta}{(e^\theta - 1)} \right]^2$$

$$E(x^2) - E(x) = \frac{\theta e^\theta (3\theta - 2\phi)}{(e^\theta - 1)} + \{E(x)\}^2 + \left\{ \frac{\theta e^\theta}{(e^\theta - 1)} \right\}^2 - \frac{2E(x)\theta e^\theta}{(e^\theta - 1)}$$

$$E(x^2) - \{E(x)\}^2 - E(x) = \frac{\theta e^\theta (3\theta - 2\phi)}{(e^\theta - 1)} + \left\{ \frac{\theta e^\theta}{(e^\theta - 1)} \right\}^2 - \frac{2E(x)\theta e^\theta}{(e^\theta - 1)}$$

$$Var(x) - E(x) = \frac{\theta e^\theta}{(e^\theta - 1)} \left[(3\theta - 2\phi) + \frac{\theta e^\theta}{(e^\theta - 1)} - 2E(x) \right]$$

$$[Var(x) - E(x)] \left[\frac{(e^\theta - 1)}{\theta e^\theta} \right] = \left[(3\theta - 2\phi) + \frac{\theta e^\theta}{(e^\theta - 1)} - 2E(x) \right]$$

$$[Var(x) - E(x)] \left[\frac{(e^\theta - 1)}{\theta e^\theta} \right] + 2E(x) - 3\theta - \frac{\theta e^\theta}{(e^\theta - 1)} = -2\phi$$

Now putting the value of ϕ in equation of mean, we get

$$\begin{aligned}
 2E(x) &= \frac{2\theta e^\theta}{(e^\theta - 1)} + 2\theta + [Var(x) - E(x)] \left[\frac{(e^\theta - 1)}{\theta e^\theta} \right] + 2E(x) - 3\theta - \frac{\theta e^\theta}{(e^\theta - 1)} \\
 \Rightarrow \frac{2\theta e^\theta}{(e^\theta - 1)} - \frac{\theta e^\theta}{(e^\theta - 1)} - 3\theta + 2\theta + [Var(x) - E(x)] \left[\frac{(e^\theta - 1)}{\theta e^\theta} \right] &= 0 \\
 \Rightarrow \frac{\theta e^\theta}{(e^\theta - 1)} - \theta + [Var(x) - E(x)] \left[\frac{(e^\theta - 1)}{\theta e^\theta} \right] &= 0 \\
 \Rightarrow \theta \left\{ \frac{e^\theta}{(e^\theta - 1)} - 1 \right\} + [Var(x) - E(x)] \left[\frac{(e^\theta - 1)}{\theta e^\theta} \right] &= 0 \\
 \Rightarrow \theta \left\{ \frac{e^\theta - e^\theta + 1}{(e^\theta - 1)} \right\} + [Var(x) - E(x)] \left[\frac{(e^\theta - 1)}{\theta e^\theta} \right] &= 0 \\
 \Rightarrow \frac{\theta}{(e^\theta - 1)} + \frac{(e^\theta - 1)}{\theta e^\theta} [Var(x) - E(x)] &= 0 \\
 \Rightarrow \frac{\theta}{(e^\theta - 1)} = -\frac{(e^\theta - 1)}{\theta e^\theta} [Var(x) - E(x)] \Rightarrow \frac{\theta}{(e^\theta - 1)} = \frac{(e^\theta - 1)}{\theta e^\theta} [E(x) - Var(x)] \\
 \Rightarrow [E(x) - Var(x)] = \frac{\theta^2 e^\theta}{(e^\theta - 1)^2} \Rightarrow E(x) - Var(x) - \frac{\theta^2 e^\theta}{(e^\theta - 1)^2} &= 0
 \end{aligned}$$