Biological Forum – An International Journal **14(1): 771-776(2022)**

ISSN No. (Print): 0975-1130
ISSN No. (Online): 2249-3239

Fitting Inflated Modified Power Series Distribution to Analysis Insect Count Data

Chetan Kumar Saini*, Neelash Patel and Rishabh
Department of Mathematics and Statistics,
College of Agriculture, JNKVV, Jabalpur, (Madhya Pradesh), India.
Department of Agriculture Rabindranath Tagore University, Raisen, (Madhya Pradesh), India.

(Corresponding author: Chetan Kumar Saini)*
(Received 22 October 2021, Accepted 01 January, 2022)
(Published by Research Trend, Website: www.researchtrend.net)

ABSTRACT: The present paper describes the fitting of inflated modified power series distribution (IMPSD) on insect count data. When modelling data consisting of counts, it is not uncommon to observe more zeroes than would be expected. The extra zeros may be a result of over-dispersion in the data. if ignoring extra zero result can be biased parameter estimates and standard errors. To overcome this difficulty, we use zero inflated power series distribution. In a zero-inflated count distribution, an additional term is added to account for these extra structural zeroes. This help to get unbiased parameter of distribution. The parameter of IMPSD estimated by method of moment (MM), Method of proportion of zeroth cell (MPZC) and method of maximum likelihood (MLE). The result show that distribution is adequate to describe inherent variation of insect.

Keywords: Modified power series distribution, Insect count data, Method of moment, MPZC, Method of maximum likelihood, χ^2 test of goodness of fit.

INTRODUCTION

Modified power series distribution (MPSD) describes a mixed population consisting of two groups of individuals. The individuals of the first group always contribute to the r^{th} cell, whereas the individuals belonging to the second group may contribute to any cell according to the MPSD. In such circumstances, the probability distributions based on certain realistic assumptions can be test. Gupta (1974) studied the application of MPSDs in genetics and derived a general expression.

Grassia and Hardy (1970) suggested application of the inflated Poisson distribution to eggs per floret of clover seed moth. Sharma (1988) suggested Polya-Aeppli distribution with zero. Saini and Sharma (2021) suggested size biased generalized negative binomial and Poisson Distributions on Crop Pests Data analysis. Patel and Saini (2020) studies of zero inflated power series distributions for modelling agricultural data. Yau *et al.* (2003) determined the zero - inflated negative binomial mixed regression modeling of over - dispersed count data with extra zeros. Wang (2001) discussed a class of Markov zero- inflated Poisson regression models for a time series of counts with the excess zero relative to Poisson distribution.

A large number of discrete distributions, besides binomial, Poisson and negative binomial are available in recent book of Johnson, Kotz and Kemp (1992). They had advocated some modified forms of the univariate discrete distributions.

MATERIAL AND METHOD

The present investigation advocates the suitable Inflated modified power series distribution which are applied to the insect count data on whiteflies population of moong crop. The data taken from M.Sc. (Agril. Stat) Thesis of Samjksha Bhodriya, Jawaharlal Nehru Krishi Vishwavidyalaya, Jabalpur, Madhya Pradesh.

Modified Power Series Distributions. R.C. Gupta (1974) first originally defined and studies modified power series distributions given below:

Let X be a discrete random variable with probability distribution function:

$$p|X = k| = \frac{\omega(k).\ g(\theta)^K}{f(\theta)}\ k \,\varepsilon\, T$$

Where, T is a subset of the set of non-negative integers, a(k)>0, g(θ) and f(θ)are positive, finite and differentiable. In case g(θ) is invertible, it reduces to Patil's (1962) generalized power series distribution and if in addition of T the entire set of non-negative integers, reduces to power series distribution first given by Noack (1950)

P [X= k] = p_k (k = 0,1 2…. …,.)

Then the inflated modified power series distribution (denoted by primes) is defined as.

$$P_0' = 1 - \omega + \omega P_0 \quad (1)$$
$$P_k' = \omega P_k (k \geq 1) \quad (2)$$
$$with\ 0 < \omega < 1$$

Sometimes ω may be greater than 1 decreasing the proportion of zeros.

Were ω is the proportion of exposure of the risk of happing insect on moong crop.

(i) inflated binomial distribution

Jain and Consul (1971) had defined a generalized negative binomial $\frac{a_0}{f(\Theta)}$ distribution given by

$$P[x=k] = \frac{n\Gamma(n+\beta k)}{k!\,\Gamma(n+\beta k-k+1)} \frac{[\theta(1-\theta)^{\beta-1}]^k}{(1-\theta)^{-n}}$$
$$K = 0,1,2….$$

were β is parameter.

In which $g(\theta) = (\theta)(1-\theta)^{\beta-1}$; $f(\theta) = (1-\theta)^{-n}$

Then the mean and variance of inflated binomial distribution becomes when β =0

$\mu_1' = \omega n\theta$

$\mu_2' = \omega n\theta(1-\theta) + \omega(1-\omega) n^2 \theta^2$

Similarly taking p = 1, it reduces to inflated negative binomial distribution, thus mean and variance of this distribution

$$Mean = \frac{\omega g(\theta) f'(\theta)}{f(\theta) g'(\theta)} = \frac{\omega n\theta}{(1-\theta)}$$

$$Variance = \frac{\omega n\theta(1-\theta)}{(1-\theta)^2} + \frac{\omega(1-\omega) n2\,\theta 2}{(1-\theta)^2}$$

(ii) Inflated Poisson distribution

Consul and Jain (1973) had given a generalized Poisson distribution given by

$$P[X=k] = \frac{\lambda_1(\lambda_1 + k\lambda_2)^{k-1}}{k!} e^{-}(\lambda_1+\lambda_2 k)^{\theta}\,\theta^k$$

In which $g(\theta) = \theta_e\text{-}\lambda_2\theta$

$f(\theta) = {}_e\lambda_1\theta$

If $\lambda = 0$, and θ = 1

We have the mean and variance of inflated Poisson distribution and thus for consequently inflated Poisson distribution.

$m_1' = \omega \lambda_1$

$m_2' = \omega\lambda_1 + \omega(1-\omega)\lambda_1^2$

Estimation of Parameters. This distribution consists of two parameters ω and θ and these are estimated by three methods given below:

(i) Method of proportion of zeroth cell (MPZC):

In this method, the observed proportion of zeros (no/N) and sample mean (m_1') are equated to their corresponding theoretical values. It is given below:

$$1 - \omega + \omega \frac{a_0}{f(\Theta)}$$

$$\text{and} \quad m_1' = \frac{\omega g(\Theta) f'(\Theta)}{f'(\Theta) g'(\Theta)}$$

The two parameters were estimated from above.

In particular, inflated binomial distribution

$n_0/N = 1 - \omega + \omega q^n$

$m_1' = \omega\, n\, p$

Assuming n is known, ω and p can be estimated. Similarly in other inflated, distribution, the estimation of parameters was determined.

(ii) Method of moments (MM):

The, r[th] moment about zero of the inflated distribution is given by ω (r[th] moment about zero of the original distribution}

Thus, the mean and variance of inflated modified power series distribution become.

$$mean = \frac{\omega g(\Theta) f'(\Theta)}{f(\Theta) g'(\Theta)}$$

$$variance = \omega \left[\frac{g(\Theta)}{g'(\Theta)} \cdot \frac{d}{d(\Theta)}\ \mu_1' + \mu_2'^2(1-\omega)\right]$$

In particular, inflated binomial distribution

$m_1' = \omega\, n\, p$

$m_2' = \omega npq + \omega(1-\omega) n^2 p^2$

Assuming n is known ω and p can be estimated. Similarly, in other inflated distribution, the estimation of parameters was determined.

(iii) Method of maximum likelihood (MLE)

N function given by (1) and (2), the likelihood function can be written as

$$L = \left[1 - \omega + \omega \frac{a_0}{f(\Theta)}\right] \prod_{K=1}^{R} \left\{\omega\left(\frac{a(k) g(\Theta)^k}{f(\Theta)}\right)\right\} nk$$

where n_k is the sample frequency of k, n is the number of non – zero sample observation (n = N – n_0) and Π the product over n non – zero observations. R is the largest number of insects.

Taking logarithm of L, differentiating with respect to ω and θ and setting the derivative equal to zero gives the estimating equations.

We have

$$\log L = n_0 \log\left[1 - \omega + \omega \frac{a_0}{f(\Theta)}\right] + \sum_{k=1}^{R} n_k \log\omega \frac{a(k) g(\Theta)^k}{f(\Theta)}$$

$$\frac{\partial(\log L)}{\partial\omega} = \frac{-n}{1-\omega+\omega\left(\frac{a_0}{f(\Theta)}\right)} \cdot \left(1 - \frac{a_0}{f(\Theta)}\right) + \frac{n}{\omega} = 0$$

$$\frac{\partial(\log L)}{\partial\theta} = \frac{-n\theta \cdot \omega \cdot a_0 f(\theta)^{-2} \cdot f'(\theta)}{1-\omega+\omega\left(\frac{a_0}{f(\Theta)}\right)} + \sum_{k=1}^{R} \frac{n_k}{p_k} \frac{k\frac{a(k)}{g(\theta) f(\theta)} - a(k) f(\theta)}{f(\theta))^2}$$

In above mentioned' parameters in each, they were estimated by equating the observed proportion of zeroes, mean and variance to their theoretical values and by maximum likelihood method.

RESULTS AND DISCUSSION

Inflated modified power series distribution. Table 1-4 describes the distribution of observed and expected number of moong plant according to number of whiteflies observation recorded 20 days interval before and after insecticide spray. The inflated binomial, and Inflated Poisson distribution was fitted by three methods described as above. In these distributions, the values of $\hat{\omega}$, $\hat{p}$ and $\hat{\lambda}$ in three methods are found to be 1.0085, 1.9568, 1.00 and 0.3165, 0.3336, 0.3192 respectively before first insecticide spray. 0.9979, 0.9779, 0.9800 and 0.5047, 0.5150, 0.5139 after first insecticide spray, whereas the value of $\hat{\omega}$ and $\hat{\lambda}$ are 1.00, 1.00, 1.02 and 0.8871, 0.8155, 0.8768 respectively before second insecticide spray. 0.8477, 0.9044, 0.86 and 0.865, 0.81, 0.8459 respectively after second insecticide spray. It shows that all the moong plants are exposed to the incidence of whitefly. The MPZC and M.L.E. provide good fitting of inflated binomial. For applying a x^2 test, some last cells are grouped. The values of x^2 in Table 1 are 14.73, 16.60, 15.09 in Table 2 are 3.97, 9.25, 4.63, for Table 3 are 11.72, 7.55, 7.22, and Table 4 are 0.27, 0.76, 0.30 respectively under method I, II and III. Since the calculated value of χ^2 is found to be less than tabulated value at 5% level of significance therefore data is well fitted. For visual displayed, the graphical representation of inflated binomial and inflated Poisson distribution using three methods of fitting are exhibited in Fig. 1-4.

Table 1: Distribution of observed and expected number of moong plants according to number of whiteflies on the observation recorded before first Insecticide spray.

No. of whiteflies	Observed frequency	Inflated binomial distribution		
		MPZC	MM	MLE
0	8	8.0	22.4	10.4
1	46	45.2	36.9	43.7
2	95	83.6	73.8	82.1
3·	86	90.4	86.2	89.8
4	52	62.8	64.8	63.1
5	22	29.1	32.4	29.6
6	15	10.9	13.5	11.3
7	3			
8	2			
9	1			
Total	330	330.0	330.0	330.0
Estimates of parameters		$\hat{\omega}$ =1.008 $\hat{p}$ =0.3165	$\hat{\omega}$ =0.9568 $\hat{p}$ =0.3336	$\hat{\omega}$ = 1.00 $\hat{p}$ =0.3192
x^2 d.f:		14.73 5	16.60 5	15.09 5

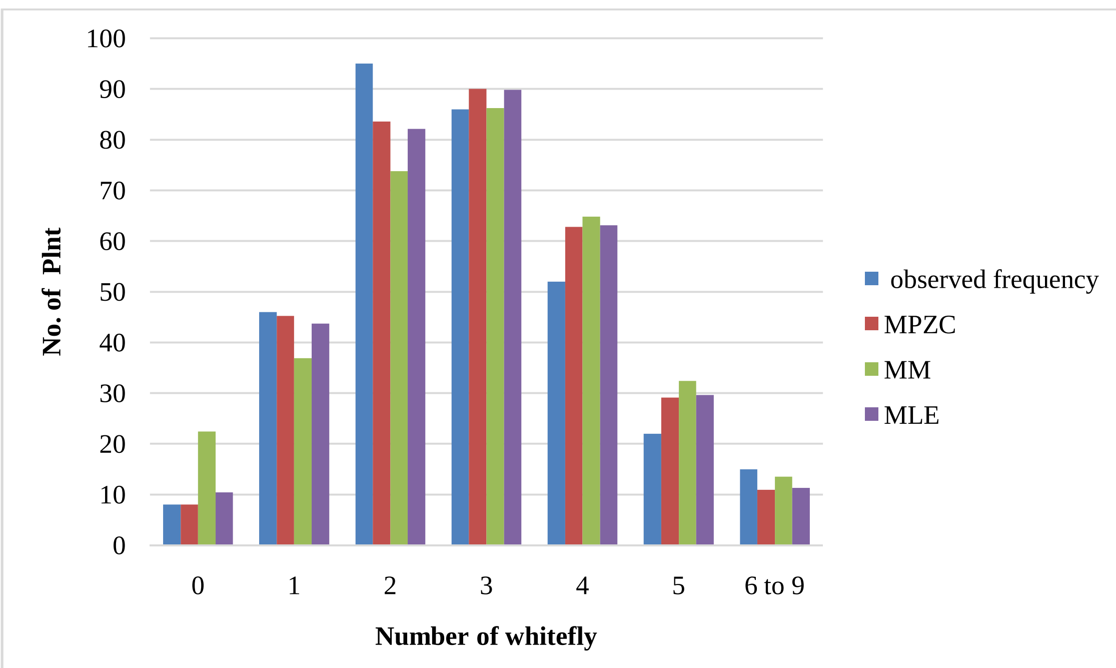

Fig. 1. Distribution of observed and expected number of moong plants according to number of whiteflies on the observation recorded before first insecticide spray.

Table 2: Distribution of observed and expected number of moong plants according to number of whiteflies on the observation recorded after first insecticide spray.

Number of whiteflies	Observed frequency	Inflated Poisson distribution		
		MPZC	MM	MLE
0	137	137.0	147.0	137.0
1	118	106.0	96.9	107.3
2	42	53.0	49.8	53.3
3	19	21.0	22.1	21.5
4	6	13.0	14.2	10.9
5	5			
6	2			
7	1			
Total	330	330.0	330.0	330.0
Estimates of parameters		$\hat{\omega}$ = 1.00 $\hat{\lambda}$ = 08871	$\hat{\omega}$ = 1.00 $\hat{\hat{\lambda}}$ = 0.8155	$\hat{\omega}$ = 1.02 $\hat{\lambda}$=0.8768
x^2 d.f.		3.97 3	9.25 3	4.63 3

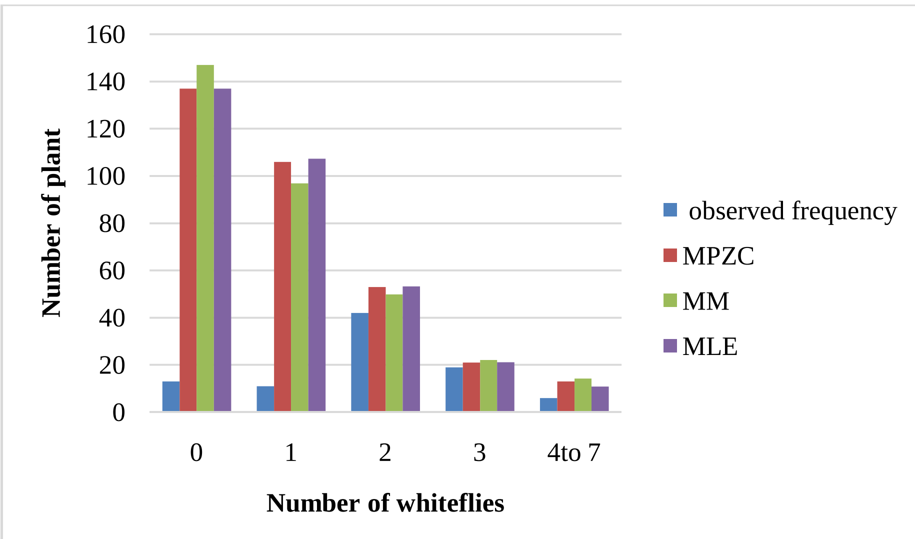

Fig. 2. Distribution of observed and expected number of moong plants according to number of whiteflies on the observation recorded after first insecticide spray.

Table 3: Distribution of observed and expected number of moong plants according to number of whiteflies on the observation recorded before second insecticide spray.

Number of whiteflies	Observed frequency	Inflated binomialdistribution		
		MPZC	MM	MLE
0	1	1.00	7.5	6.8
1	6	3.0	2.5	2.5
2	18	13.7	11.8	12.0
3	35	37.7	33.4	33.8
4	62	66.2	62.0	62.5
5	78	81.0	79.1	79.3
6	67	68.8	69.9	69.8
7	38	40.1	42.4	42.2
8	17	15.3	16.9	16.7
9	7	3.2	4.5	4.4
10 ·	1			
Total	330	330.0	330.0	330.0
Estimates of parameters		$\hat{\omega}$ = 0.9979 $\hat{p}$ = 0.5047	$\hat{\omega}$ = 0.9779 $\hat{p}$= 0.5150	$\hat{\omega}$ = 0.98 $\hat{p}$=0.5139
x^2 d.f.		11.72 7	7.55 7	7.12 7

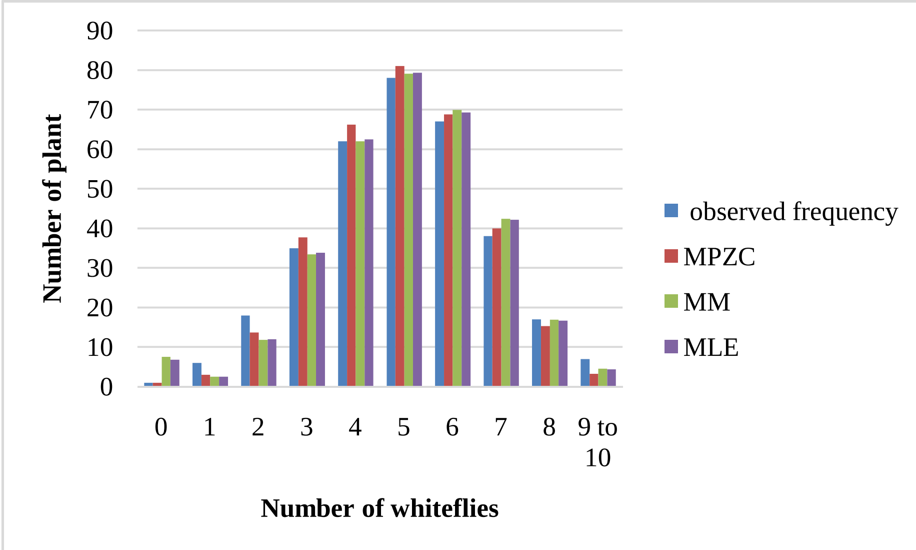

Fig. 3. Distribution of observed arid expected number of moong plants according to number of whiteflies on the observation recorded before second insecticide spray.

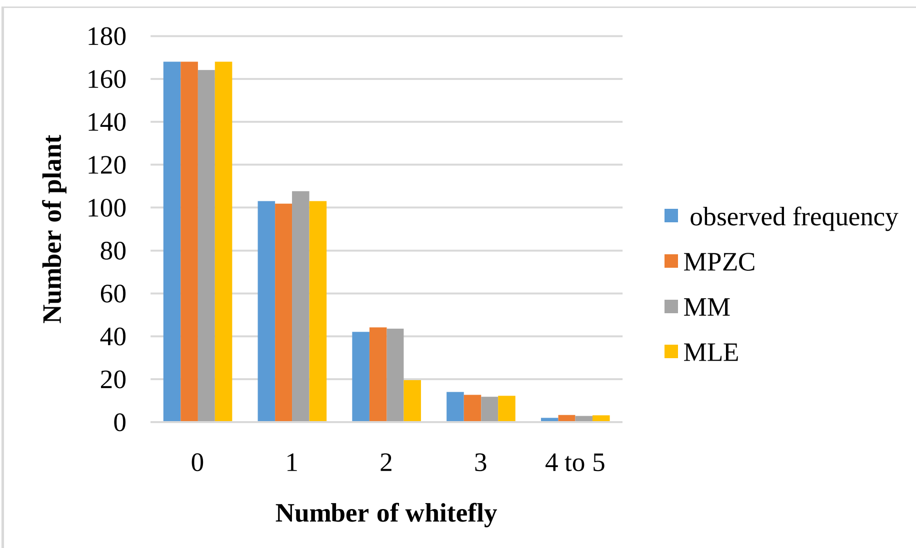

Fig. 4. Distribution of observed and expected number of moong plants according to number of whitefly-on the observation recorded after second insecticide spray.

Table 4: Distribution of observed and expected number of moong plants according to number of whiteflies on the observation recorded after second insecticide spray.

No. of whiteflies	Observed frequency	Inflated Poisson distribution		
		MPZC	MM	MLE
0	168	168.0	164.2	168.0
1	103	101.9	107.6	103.0
2	42	44.1	43.6	43.6
3	14	12.7	11.8	12.3
4	2			
5	1			
Total	330	330.0	330.0	330.0
Estimates of parameters		$\hat{\omega} = 0.8477$ $\hat{\lambda} = 0.865$	$\hat{\omega} = 0.9053$ $\hat{\lambda} = 0.81$	$\hat{\omega} = 0.86$ $\hat{\lambda} = 0.8459$
x^2 d.f.		0.27 2	0.76 2	0.30 2

COCLUSION

1. The present result shows that Inflated, binomial and Poisson distribution was found adequate to explain the inherent variation of insect population on moong crop.
2. For estimation of parameters the method of moments, method of proportion of zeroth cell (MPZC), and maximum likelihood estimation method were found to be the suitable.

FUTURE SCOPE

When the number of occurrences of whitefly were more, it was difficult to make a regular frequency distribution. Therefore, the fitting of adequate distributions help to designing efficient sampling programs for population estimation, development of population models and pest management.

Acknowledgement. The authors are grateful to the editor and referee's for giving their cogent comments on the paper that have really changed the shape of the paper
Conflict of Interest. None

REFERENCES

Bhodriya, S. (2009). 'Inflated Modified Power Series Distribution and Path Analysis Application to Agricultural Data, unpublished M.Sc. thesis, *J.N.K.V.V., Jabalpur.*

Gupta, (1974). Modified power series distribution and some of its applications. *Sankhiya. Series 836*: 288-298.

Grassia, A. and Hardy, R.J. (1970). Application of the inflated Poisson distribution to eggs per floret of clover seed moth coleophora alcyonipennella kollar, *Australian Journal of Entomology, 9*(2): 162-164.

Jain, G. C. and Consul, P. C. (1971). A Generalized negative binomial distribution. *SIAM J. Applied Mathematics, 21*: 501-518.

Johnson, N. L., Kotz, Sand Kemp, A. W. (1992). Univariate Discrete Distribution. 2nd edition. *Publ. John Wiley & Sons, Inc. NY*

Noack, A. (1950). On a class of discrete random variables. *Annals of mathematical statistics, l21*: 127-132

Patil, G. P. (1962). Estimation by two-moments method for generalized power series distribution and certain applications. *Sankhya, 24*: 201- 214.

Patel, N., and Saini, C. K. (2020). Application of zero inflated power series distributions for modelling agricultural data related to insects. *Indian Journal of Entomology 82*(3): 423-425.

Consul, P. C. and Jain, G. C. (1973). A generalization of the Poisson distribution. *Technimetrics, 15*: 791–799.

Jain, G. C., & Consul, P. C. (1971). A generalized negative binomial distribution. *SIAM Journal of Applied Mathematics, 21*(4), 501-513

Saini, C. K., Sharma, H.L., and Patel, P. (2021) Fitting of Size Biased Generalized Negative Binomial and Poisson Distributions on Crop Pests Data. *Journal of the Indian Society of Agricultural Statistics 75*(1): 9–12.

Sharma, H. L. (1988). Estimating the parameters of a Polya - Aeppli distribution. *Rural Demography, 15*(1&2).

Yau, K. K.W. and Wang, K. (2003). Zero-inflated negative binomial mixed regression modeling of over-dispersed count data with extra zeros. *Biometrical Journal 45*(4): 437-452.

Wang, P. M. (2001). Markov zero-inflated Poisson regression models for a time series of counts with excess zeros. *Journal of Applied Statistics, 28*(5): 623-632.

How to cite this article: Chetan Kumar Saini, Neelash Patel and Rishabh (2022). Fitting Inflated Modified *Power Series Distribution to Analysis Insect Count Data*. *Biological Forum – An International Journal, 14*(1): 771-776.