

# On some aspects of intervened generalized Hermite distribution

C. Satheesh Kumar · D. S. Shibu

Received: 27 January 2012 / Accepted: 16 January 2013 / Published online: 10 April 2013  
© Sapienza Università di Roma 2013

**Abstract** Summary The intervened Poisson distribution (IPD) of Shanmugam (Biometrics 41:1025–1029, 1985) has been found suitable for some rare event situations where some intervention arises. The main drawback of IPD is that it is under-dispersed and appropriate for single intervention situation. Through this paper, we introduce a general class of distributions, which includes IPD as its special case, suitable for situations of under-dispersed, over-dispersed and multiple intervention cases. Several properties of this new class of intervened distribution are investigated and illustrated its usefulness with the help of real life data sets.

**Keywords** Factorial moments · Intervened Poisson distribution · Intervened generalized Poisson distribution · Maximum likelihood estimation · Probability generating function

## 1 Introduction

Shanmugam [19] considered a new class of discrete distributions, namely intervened Poisson distribution (IPD) as a replacement for the zero-truncated Poisson distribution (ZTPD) and an independent Poisson random variable. The IPD is the distribution of the sum of a zero-truncated Poisson random variable and an independent Poisson random variable. Shanmugam [19] defined the IPD as follows:

A positive integer valued random variable  $X$  is said to follow IPD if its probability mass function (pmf) has the following form, for  $x = 1, 2, \dots$

$$f(x) = C[(1 + \rho)^x - \rho^x] \frac{\lambda^x}{x!}, \quad (1.1)$$

---

C. S. Kumar (✉) · D. S. Shibu  
Department of Statistics, University of Kerala, Trivandrum 695 581, India  
e-mail: drcsatheeshkumar@gmail.com

D. S. Shibu  
e-mail: dsshibu70@ymail.com

in which  $C = qe^{-\lambda(1+\rho)}$ , with  $q = e^\lambda(e^\lambda - 1)^{-1}$ ,  $\lambda > 0$  and  $\rho \geq 0$ . An advantage of the IPD is that it provides information on how effective various preventive actions taken by health service agents, where ZTPD fails. The IPD is applicable in several areas such as reliability analysis, queuing problems, epidemiological studies, etc. For example, see Shanmugam [19,20], Huang and Fung [7]. The mean and variance of IPD with pmf (1.1) are

$$E(X) = \lambda[\rho + q] = \mu, \tag{1.2}$$

and

$$Var(X) = \mu - \lambda^2q(q - 1). \tag{1.3}$$

From (1.2) and (1.3), it can be observed that the IPD  $(\lambda, \rho)$  is under-dispersed (that is, its variance is less than its mean), for all values of the parameters. The IPD has been further studied by several authors such as Shanmugam [20], Scollnik [16–18], Dhanavanathan [3,4], Patel and Gajjar [13], Kumar and Shibu [11,12].

Consul [1] defined generalized Poisson distribution (GPD) in the following way. A random variable  $Y$  is said to follow the GPD with parameters  $\alpha$  and  $\beta$  [written as  $GPD(\alpha, \beta)$ ], if its pmf is given by

$$g(y) = \frac{\beta(\beta + y\alpha)^{y-1}e^{-(\beta+y\alpha)}}{y!}, \tag{1.4}$$

for  $y = 0, 1, \dots$  with  $0 \leq \alpha < 1$  and  $\beta \geq 0$ , and zero elsewhere. A truncated version of the GPD has been studied by Consul and Famoye [2].

Scollnik [18] obtained the intervened generalized Poisson distribution (IGPD) as in the following. For  $0 \leq \alpha < 1$ ,  $\beta \geq 0$  and  $\lambda > 0$ , let  $Z_1$  and  $Z_2$  be independent random variables such that  $Z_1$  follows zero-truncated  $GPD(\alpha, \beta)$  and  $Z_2$  follows the  $GPD(\alpha, \beta\lambda)$ . Then  $Z = Z_1 + Z_2$  follows IGPD with parameters  $\alpha, \beta$  and  $\lambda$  if its pmf given by

$$g_2(z) = \frac{\lambda [(1 + \beta) \{ (1 + \beta)\lambda + z\alpha \}^{z-1} - \beta \{ \beta\lambda + z\alpha \}^{z-1} ]}{e^{\beta\lambda+z\alpha}(e^\lambda - 1)z!}, \tag{1.5}$$

for  $z = 1, 2, \dots$ . The mean and variance of the IGPD with pmf (1.5) are the following, in which  $q = \frac{e^\lambda}{e^\lambda - 1}$

$$E(Z) = (1 - \alpha)^{-1} [\beta + q] = \nu, \tag{1.6}$$

$$Var(Z) = (1 - \alpha)^{-2} [\nu - \lambda^2q(q - 1)]. \tag{1.7}$$

From the expressions (1.6) and (1.7), it is seen that IGPD may have variance less than or greater than mean depending on the values of  $\alpha, \beta$  and  $\lambda$ .

Gupta and Jain [6] introduced the generalized Hermite distribution through the following probability generating function (pgf), in which  $\mu_1 > 0, \mu_2 \geq 0$  and  $m$  is an integer greater than one.

$$G(s) = e^{\mu_1(s-1)+\mu_2(s^m-1)}. \tag{1.8}$$

When  $m = 2$ , (1.8) reduces to the pgf of the well-known Hermite distribution of [8]. Also, when  $m = 1$ , (1.8) reduces to the pgf of Poisson distribution with parameter  $\mu_1 + \mu_2$ . The Hermite distribution has found extensive applications in several areas of research such as biological sciences, physical sciences and operations research. Both Hermite and generalized Hermite distribution belongs to Kumar [10] family of distributions. The generalized Hermite distribution with pgf (1.8) hereafter we will denote as  $H_m(\mu_1, \mu_2)$ .

But in practice, there are situations where more than one intervention arises. The ZTPD, the IPD and the IGPD are all found to be suitable for dealing situations of only one intervention. For tackling such situations of more than one intervention, through this paper, we introduce a wide class of intervened type distributions, namely intervened generalized Hermite distribution (IGHD), as a generalization of IPD and discuss some of its important properties. The paper is organized as follows. In Sect. 2, we present the definition of IGHD and shown that the ZTPD, the IPD and positive Hermite distribution and the intervened Hermite distribution are special cases of IGHD. Some important properties of IGHD such as its pgf, pmf, mean, variance and recurrence relation for its probabilities and factorial moments are also derived in this section. In Sect. 3, we discuss the estimation of parameters of IGHD by method of maximum likelihood. Further, two real life data sets are considered for illustrating the suitability of IGHD compared to the existing models such as the ZTPD, the IPD and the IGPD.

### 2 The intervened generalized Hermite distribution

In this section we present the definition of the intervened generalized Hermite distribution and derive some of its important properties

**Definition 2.1** For any fixed positive integer  $m$ , let  $U_1$  be a non-negative integer valued random variable having zero-truncated generalized Hermite distribution with pgf

$$B(s) = (e^{\mu_1 + \mu_2} - 1)^{-1} (e^{\mu_1 s + \mu_2 s^m} - 1)^{-1} \tag{2.1}$$

in which  $\mu_1 > 0$  and  $\mu_2 \geq 0$ . Let  $U_2$  be a discrete random variable following generalized Hermite distribution  $H_m(\rho_1 \mu_1, \rho_2 \mu_2)$  with  $\rho_1 > 0, \rho_2 \geq 0$ . Assume that  $U_2$  is independent of  $U_1$ . Then the distribution of the random variable  $U = U_1 + U_2$  is called the intervened generalized Hermite distribution with parameters  $\mu_1, \mu_2, \rho_1, \rho_2$  and  $m$  or in short  $IGHD_m(\mu_1, \mu_2, \rho_1, \rho_2)$ .

The  $IGHD_m(\mu_1, \mu_2, \rho_1, \rho_2)$  has the following special cases:

- (1) When  $\mu_2 = 0$ , the  $IGHD_m(\mu_1, \mu_2, \rho_1, \rho_2)$  reduces to  $IPD(\mu_1, \rho_1)$ .
- (2) When  $\rho_2 = 0$ , the  $IGHD_m(\mu_1, \mu_2, \rho_1, \rho_2)$  reduces to an extended form of the IPD with pgf

$$Q_1(s) = c_1 (e^{\mu_1 s + \mu_2 s^m} - 1) e^{\mu_1 \rho_1 s},$$

in which  $c_1 = [e^{\mu_1(1+\rho_1)+\mu_2} - e^{\mu_1 \rho_1}]^{-1}$ .

- (3) When  $\rho_1 = 0$  and  $\mu_2 = 0$ , the  $IGHD_m(\mu_1, \mu_2, \rho_1, \rho_2)$  reduces to  $ZTPD$ .
- (4) When  $\rho_1 = 0$  and  $\rho_2 = 0$ , the  $IGHD_m(\mu_1, \mu_2, \rho_1, \rho_2)$  reduces to zero truncated Hermite distribution.

Note that, since  $U_2$  follows  $H_m(\rho_1 \mu_1, \rho_2 \mu_2)$ ,  $U_2$  can be written as  $U_2 = X_1 + mX_2$ , where  $X_1$  and  $X_2$  are independent Poisson variates with means  $\rho_1 \mu_1$  and  $\rho_2 \mu_2$  respectively. Now the  $IGHD_m(\mu_1, \mu_2, \rho_1, \rho_2)$  variate  $U$  has the following form.

$$\begin{aligned} U &= U_1 + X_1 + mX_2 \\ &= V_1 + mX_2, \end{aligned}$$

where  $V_1$  follows a modified form of  $IPD$ , since when  $\mu_2 = 0$ ,  $U_1$  reduces to  $ZTPD$ . Thus  $V_1$  is enough for tackling single intervention situations and consequently, the  $IGHD_m(\mu_1, \mu_2, \rho_1, \rho_2)$  is suitable for situations where more than one intervention arises. We obtain the pgf of  $IGHD_m(\mu_1, \mu_2, \rho_1, \rho_2)$  through the following result.

**Result 2.1** Let  $X$  follows  $IGHD_m(\mu_1, \mu_2, \rho_1, \rho_2)$ . Then the pgf of  $X$  is the following.

$$Q(s) = c(e^{\mu_1 s + \mu_2 s^m} - 1)e^{\mu_1 \rho_1 s + \mu_2 \rho_2 s^m} \tag{2.2}$$

where

$$c = (e^{\mu_1 + \mu_2} - 1)^{-1} e^{-(\mu_1 \rho_1 + \mu_2 \rho_2)} \tag{2.3}$$

Proof follows from the fact that the pgf of sum of independent random variables is the product of their pgfs.

Now to obtain expressions for pmf and factorial moments of the  $IGHD_m(\mu_1, \mu_2, \rho_1, \rho_2)$ , we need the following lemma.

**Lemma 2.1** For any  $a, b, v \in R$  and for any positive integer  $m$  greater than one,

$$e^{av + bv^m} = \sum_{x=0}^{\infty} \sum_{j=0}^{\lfloor \frac{x}{m} \rfloor} \frac{a^{x-mj} b^j}{(x-mj)! j!} v^x \tag{2.4}$$

and

$$e^{av + b(1+v)^m} = \sum_{r=0}^{\infty} \sum_{j=0}^{\infty} \sum_{i=0}^r \frac{a^{r-i} b^j}{(r-i)! j!} \binom{mj}{i} v^r. \tag{2.5}$$

Proof of (2.4) and (2.5) follow by expanding the exponential and binomial term in the light of the following series representation ([15], pp.134).

$$\sum_{x=0}^{\infty} \sum_{j=0}^{\infty} B(j, x) = \sum_{x=0}^{\infty} \sum_{j=0}^{\lfloor \frac{x}{m} \rfloor} B(j, x - mj).$$

**Result 2.2** The probability mass function  $q_x = P(X = x)$  of  $IGHD_m(\mu_1, \mu_2, \rho_1, \rho_2)$  is the following for  $x = 1, 2, 3, \dots$  in which  $\mu_1 > 0, \mu_2 \geq 0, \rho_1 \geq 0, \rho_2 \geq 0$  and  $[k]$  denotes the integer part of  $k$ .

$$q_x = c \sum_{r=0}^{\lfloor \frac{x}{m} \rfloor} \Lambda_{x-mr, r}(\rho_1, \rho_2) \frac{\mu_1^{x-mr} \mu_2^r}{(x-mr)! r!} \tag{2.6}$$

where for any non-negative integers  $i$  and  $j$ ,

$$\Lambda_{i, j}(\rho_1, \rho_2) = [(1 + \rho_1)^i (1 + \rho_2)^j - \rho_1^i \rho_2^j] \tag{2.7}$$

and  $c$  is as given in (2.3).

*Proof* From (2.1), we have the following.

$$\begin{aligned} Q(s) &= \sum_{x=1}^{\infty} s^x q_x \\ &= c(e^{\mu_1 s + \mu_2 s^m} - 1)e^{\mu_1 \rho_1 s + \mu_2 \rho_2 s^m} \end{aligned} \tag{2.8}$$

$$Q(s) = c \sum_{x=0}^{\infty} \sum_{r=0}^{\lfloor \frac{x}{m} \rfloor} \frac{\Lambda_{x-mr, r}(\rho_1, \rho_2) \mu_1^{x-mr} \mu_2^r s^x}{(x-mr)! r!}, \tag{2.9}$$

since by Lemma 2.1. Now on equating coefficient of  $s^x$ , on right side expressions of (2.8) and (2.9), we get (2.6).

**Result 2.3** The mean and variance of the  $IGHD_m(\mu_1, \mu_2, \rho_1, \rho_2)$  are the following, in which  $\eta = e^{\mu_1 + \mu_2}(e^{\mu_1 + \mu_2} - 1)^{-1}$ .

$$E(X) = \eta(\mu_1 + m\mu_2) + \mu_1\rho_1 + m\mu_2\rho_2 \tag{2.10}$$

and

$$Var(X) = \mu_1\rho_1 + m^2\mu_2\rho_2 + \eta(1 - \eta)(\mu_1 + m\mu_2)^2 + \eta(\mu_1 + m^2\mu_2) \tag{2.11}$$

Proof is simple and hence omitted. Based on the Result 2.3, we have the following practical remark.

*Remark 2.1* In general  $IGHD_m(\mu_1, \mu_2, \rho_1, \rho_2)$  is over-dispersed. A necessary condition for  $IGHD_m(\mu_1, \mu_2, \rho_1, \rho_2)$  to be under-dispersed is  $\text{Max}\{\mu_1, \rho_1, \rho_2\} < 1$  and  $\mu_2$  sufficiently small.

**Result 2.4** The  $r$ -th factorial moment  $\mu_{[r]}$  of  $IGHD_m(\mu_1, \mu_2, \rho_1, \rho_2)$  with pgf (2.1) is the following, for  $r \geq 1$ .

$$\mu_{[r]} = ce^{\mu_1\rho_1}r! \sum_{j=0}^{\infty} \sum_{i=0}^r \frac{\mu_1^{r-i}\mu_2^j}{(r-i)!j!} \binom{mj}{i} [e^{\mu_1}\gamma_1^{r-i}\gamma_2^j - \rho_1^{r-i}\rho_2^j], \tag{2.12}$$

in which  $\gamma_i = 1 + \rho_i$  for  $i = 1, 2$  and  $c$  is as defined in (2.3).

*Proof* The factorial moment generating function (fmgf)  $F(t)$  of  $IGHD_m(\mu_1, \mu_2, \rho_1, \rho_2)$  with pgf (2.1) is

$$\begin{aligned} F(t) &= \sum_{r=0}^{\infty} \frac{t^r}{r!} \mu_{[r]} \\ &= Q(1 + t) \end{aligned} \tag{2.13}$$

$$F(t) = c(e^{\mu_1(1+t)+\mu_2(1+t)^m} - 1)e^{\mu_1\rho_1(1+t)+\mu_2(1+t)^m} \tag{2.14}$$

By applying the Lemma 2.1 in (2.14), we get

$$F(t) = ce^{\mu_1\rho_1}r! \sum_{j=0}^{\infty} \sum_{i=0}^r \frac{\mu_1^{r-i}\mu_2^j}{(r-i)!j!} \binom{mj}{i} [e^{\mu_1}\gamma_1^{r-i}\gamma_2^j - \rho_1^{r-i}\rho_2^j] \frac{t^r}{r!} \tag{2.15}$$

where  $\gamma_i = 1 + \rho_i$  for  $i = 1, 2$ . Equating coefficient of  $(r!)^{-1}t^r$ , we obtain (2.12).

*Remark 2.2* If we put  $\mu_2 = 0$  in (2.12), we get

$$\mu_{[r]} = \frac{\mu_1^r}{(e^{\mu_1} - 1)} [e^{\mu_1}(1 + \rho_1)^r - \rho_1^r],$$

which is the  $r$ -th factorial moment of  $IPD(\mu_1, \rho_1)$ .

**Result 2.5** For  $x \geq m - 1$ , a recurrence relation for the probabilities of  $IGHD_m(\mu_1, \mu_2, \rho_1, \rho_2)$  is the following, in which  $\gamma_i = 1 + \rho_i$ , for  $i = 1, 2$ .

$$(x + 1)q_{x+1} = \mu_1\gamma_1 + m\mu_2\gamma_2q_{x-m+1} + c[\mu_1\xi_x + m\mu_2\xi_{x-m+1}] \tag{2.16}$$

where

$$\xi_r = \sum_{j=0}^{\lfloor \frac{r}{m} \rfloor} \frac{(\mu_1 \rho_1)^{r-mj} (\mu_2 \rho_2)^j}{(r - mj)! j!}, \tag{2.17}$$

for any  $r \geq 1$ , integer.

*Proof* From (2.2), we have the following.

$$Q(s) = \sum_{x=1}^{\infty} q_x s^x = c(e^{\mu_1 s + \mu_2 s^m} - 1)e^{\mu_1 \rho_1 s + \mu_2 \rho_2 s^m} \tag{2.18}$$

$$= c[e^{\mu_1 \gamma_1 s + \mu_2 \gamma_2 s^m} - e^{\mu_1 \rho_1 s + \mu_2 \rho_2 s^m}] \tag{2.19}$$

On differentiating right hand side expressions of (2.18) and (2.19) with respect to  $s$ , we obtain the following in the light of Lemma 2.1.

$$\sum_{x=1}^{\infty} x q_x s^{x-1} = (\mu_1 \gamma_1 + m \mu_2 s^{m-1})Q(s) + c(\mu_1 + m \mu_2 s^{m-1})e^{\mu_1 \rho_1 s + \mu_2 \rho_2 s^m} \tag{2.20}$$

Equating coefficient of  $s^x$  on both sides of (2.20), we get (2.16).

### 3 Estimation

In this section we discuss the maximum likelihood estimation of the parameters  $\mu_1, \mu_2, \rho_1$  and  $\rho_2$  of the  $IGHD_m(\mu_1, \mu_2, \rho_1, \rho_2)$ . The log likelihood of the sample is

$$\ln L = \sum_{x=1}^z n_x \ln q_x, \tag{3.1}$$

where  $n_x$  is the observed frequency of  $x$  events and  $z$  is the highest value of  $x$  observed. On differentiating (3.1) with respect to  $\mu_1, \mu_2, \rho_1$  and  $\rho_2$  and equating to zero, we obtain the following likelihood equations, in which  $\sum_I$  denotes the 2-tuple sum over the set  $I = \{(x, r) : x = 1, 2, \dots, z; r = 1, 2, \dots, \lfloor \frac{x}{m} \rfloor\}$  and  $\Lambda_{ij} = \Lambda_{ij}(\rho_1, \rho_2)$  is as defined in (2.5).

$$\frac{\partial \ln L}{\partial \mu_1} = 0$$

implies

$$\sum_I \frac{n_x \Lambda_{x-mr,r} \mu_2^r \mu_1^{x-mr}}{q_x (x - mr - 1)! r!} = (\eta + \rho_1) \sum_I \frac{n_x \Lambda_{x-mr,r} \mu_2^r \mu_1^{x-mr}}{q_x (x - mr)! r!} \tag{3.2}$$

$$\frac{\partial \ln L}{\partial \mu_2} = 0$$

implies

$$\sum_I \frac{n_x \Lambda_{x-mr,r} \mu_2^{r-1} \mu_1^{x-mr-1}}{q_x(x-mr)!(r-1)!} = (\eta + \rho_2) \sum_I \frac{n_x \Lambda_{x-mr,r} \mu_2^r \mu_1^{x-mr}}{q_x(x-mr)!r!} \tag{3.3}$$

$$\frac{\partial \ln L}{\partial \rho_1} = 0$$

implies

$$\sum_I \frac{n_x \Lambda_{x-mr,r} \mu_2^r \mu_1^{x-mr}}{q_x(x-mr-1)r!} = \sum_I \frac{n_x \Lambda_{x-mr,r} \mu_2^r \mu_1^{x-mr}}{q_x(x-mr)!r!} \tag{3.4}$$

and

$$\frac{\partial \ln L}{\partial \rho_2} = 0$$

implies

$$\sum_I \frac{n_x \Lambda_{x-mr-1,r} \mu_2^r \mu_1^{x-mr}}{q_x(x-mr)!(r-1)!} = \sum_I \frac{n_x \Lambda_{x-mr,r} \mu_2^r \mu_1^{x-mr}}{q_x(x-mr)!r!} \tag{3.5}$$

These likelihood equations do not always have a solution because the  $IGHD_m(\mu_1, \mu_2, \rho_1, \rho_2)$  is not a regular model. Therefore, when likelihood equations do not have a solution the maximum of the likelihood function is attained at the border of the domain of parameters. It also happens for the generalized Hermite distribution and necessary and sufficient conditions to ensure that the maximum likelihood estimator is the solution of likelihood equations are given in [14]. As such, we obtained the second order partial derivatives of  $\ln q_x$  with respect to parameters  $\mu_1, \mu_2, \rho_1$  and  $\rho_2$  as given in Appendix –A. Now, by using MATHCAD software one can observe that these equations are negative for all  $\mu_i \in [0, 1]$  and  $\rho_i \in [0, 1]$ , for  $i = 1, 2$ . Thus the  $IGHD_m(\mu_1, \mu_2, \rho_1, \rho_2)$  is log concave and hence the maximum likelihood estimators (MLE) of the parameters  $\mu_1, \mu_2, \rho_1$  and  $\rho_2$  are unique under these parametric restrictions (cf. [14]). Now the MLE of the parameters  $\mu_1, \mu_2, \rho_1$  and  $\rho_2$  can be obtained by solving the system of likelihood equations (3.2–3.5) with the help of some mathematical softwares such as MATHCAD, MATHEMATICA, R etc.

We have fitted the ZTPD, the IPD, the IGPD and the  $IGHD_m(\mu_1, \mu_2, \rho_1, \rho_2)$  to the following two data sets by the method of maximum likelihood with the help of MATHCAD software and presented in Tables 1 and 2.

The first data set, given in Table 1 indicates the distribution of number of articles on theoretical Statistics and Probability for years 1940–1949 and initial letter N-R of the author’s name. For details, see Kendall [9]. The second data set given in Table 2 represents the distribution of 1,534 biologists according to the number of research papers to their credit in the review of applied entomology, volume 24, 1936.  $IGHD_m(\mu_1, \mu_2, \rho_1, \rho_2)$ . For details see Williams [21]. We have computed the value of estimators and their standard errors (SE) and included in respective tables. For the second dataset, we have pooled the classes corresponding to the observed values 6, 7, 8, 9, 10 and 10+ together so as to satisfy the requirement of Pearson Goodness of fit (See Greenwood [5]) and consequently degrees of freedom of Chi-square statistic becomes  $k - p - 1$ , where  $k$  is the number of categories and  $p$  is the number of estimated parameters. Based on Chi-square values and P values given in the tables, it can be observed that  $IGHD_3(\mu_1, \mu_2, \rho_1, \rho_2)$  gives a better fit to the first data

**Table 1** Comparison of fit of  $IGHD_m$  by method of maximum likelihood for the first dataset

x	Observed frequency	Expected frequency					
		ZTPD	IPD	IGPD	IGHD <sub>2</sub>		
1	198	215	189	185	200		
2	60	50	66	66	54		
3	20	18	27	27	18		
4	21	17	19	19	26		
5	10	12	11	11	12		
6	5	7	10	8	7		
7	32	20	36	30	29		
Total	346	346	346	346	346		
Estimates of parameters with corresponding SE within brackets		$\hat{\lambda} = 1.88 (0.074)$	$\hat{\lambda} = 1.47 (0.067)$	$\hat{\alpha} = 0.50 (0.14)$	$\hat{\mu}_1 = 0.42 (0.0147)$	$\hat{\mu}_1 = 0.26 (0.00289)$	$\hat{\mu}_1 = 0.32 (0.0175)$
			$\hat{\rho} = 0.21 (0.043)$	$\hat{\beta} = 0.86 (0.0476)$	$\hat{\mu}_2 = 0.25 (0.347)$	$\hat{\mu}_2 = 0.15 (0.00247)$	$\hat{\mu}_2 = 0.10 (0.475)$
				$\hat{\lambda} = 0.12 (0.0057)$	$\hat{\rho}_1 = 0.48 (0.027)$	$\hat{\rho}_1 = 0.12 (0.00122)$	$\hat{\rho}_1 = 0.15 (0.00345)$
					$\hat{\rho}_2 = 0.26 (0.0075)$	$\hat{\rho}_2 = 0.16 (0.00312)$	$\hat{\rho}_2 = 0.17 (0.011)$
Degrees of freedom		5	4	3	2	2	2
Chi square value		6.34	4.41	4.83	3.09	2.81	5.15
P value		0.39	0.62	0.57	0.80	0.83	0.70





set whereas  $IGHD_2(\mu_1, \mu_2, \rho_1, \rho_2)$  gives a better fit to the second data set compared to the existing models such as ZTPD, IPD and IGPD.

Further, we have verified that mean values of the  $IGHD_m(\mu_1, \mu_2, \rho_1, \rho_2)$  for each value of  $m$  agrees with the corresponding sample mean which shows the consistency of estimators of parameters obtained through maximum likelihood estimation discussed in this section. Thus, it can be seen that  $IGHD_m(\mu_1, \mu_2, \rho_1, \rho_2)$  model gives better fit to both datasets compared to the classical models for particular value of  $m$ . This may be due to the fact that  $IGHD_m(\mu_1, \mu_2, \rho_1, \rho_2)$  considers the effect of further interventions such as academic promotion criteria, research incentives etc. Also, it can be observed from tables that the number of articles published in different categories of datasets are not very small for higher observed  $x$  values, which indicates some interventions such as measures taken for academic promotion criteria, research incentives etc. that favours the number of publications in each case.

**Acknowledgements** The authors would like to express their gratitude to the Editor-in-Chief and both anonymous referees for their valuable comments on an earlier version of the paper, which greatly improved the quality of the paper.

### Appendix

The second order partial derivatives of  $\ln q_x$  with respect to  $\mu_1, \mu_2, \rho_1$  and  $\rho_2$  are obtained as given below:

$$\begin{aligned} \frac{\partial^2}{\partial \mu_1^2} \ln q_x &= \frac{e^{\mu_1 + \mu_2}}{(e^{\mu_1 + \mu_2} - 1)^2} + \sum_{r=0}^{\lfloor \frac{x}{m} \rfloor} \sum_{r=0}^{\lfloor \frac{x-2}{m} \rfloor} \frac{\Lambda_{x-mr,r}^2}{((x-mr)!)^2 r!(r-2)!} (\mu_1^{x-mr-1} \mu_2^r) \\ &\quad \times \left( \sum_{r=0}^{\lfloor \frac{x}{m} \rfloor} \frac{\Lambda_{x-mr,r} \mu_1^{x-mr} \mu_2^r}{(x-mr)! r!} \right)^{-2} - \left[ \sum_{r=0}^{\lfloor \frac{x-1}{m} \rfloor} \frac{\Lambda_{x-mr,r}}{(x-mr-1)! r!} \mu_1^{x-mr-1} \mu_2^r \right]^2 \\ &\quad \times \left( \sum_{r=0}^{\lfloor \frac{x}{m} \rfloor} \frac{\Lambda_{x-mr,r} \mu_1^{x-mr} \mu_2^r}{(x-mr)! r!} \right)^{-2} \end{aligned} \tag{4.1}$$

$$\begin{aligned} \frac{\partial^2}{\partial \mu_2^2} \ln q_x &= \frac{e^{\mu_1 + \mu_2}}{(e^{\mu_1 + \mu_2} - 1)^2} + \sum_{r=0}^{\lfloor \frac{x}{m} \rfloor} \sum_{r=2}^{\lfloor \frac{x}{m} \rfloor} \frac{\Lambda_{x-mr,r}^2}{((x-mr)!)^2 r!(r-2)!} (\mu_1^{x-mr} \mu_2^{r-1}) \\ &\quad \times \left( \sum_{r=0}^{\lfloor \frac{x}{m} \rfloor} \frac{\Lambda_{x-mr,r} \mu_1^{x-mr} \mu_2^r}{(x-mr)! r!} \right)^{-2} - \left[ \sum_{r=1}^{\lfloor \frac{x}{m} \rfloor} \frac{\Lambda_{x-mr,r}}{(x-mr)!(r-1)!} \mu_1^{x-mr} \mu_2^{r-1} \right]^2 \\ &\quad \times \left( \sum_{r=0}^{\lfloor \frac{x}{m} \rfloor} \frac{\Lambda_{x-mr,r} \mu_1^{x-mr} \mu_2^r}{(x-mr)! r!} \right)^{-2} \end{aligned} \tag{4.2}$$

$$\frac{\partial^2}{\partial \rho_1^2} \ln q_x = \sum_{r=0}^{\lfloor \frac{x}{m} \rfloor} \sum_{r=0}^{\lfloor \frac{x-2}{m} \rfloor} \frac{\Lambda_{x-mr,r} \Lambda_{x-mr-2,r}}{(x-mr)!(x-mr-2)!(r!)^2} (\mu_1^{x-mr} \mu_2^r) \left( \sum_{r=0}^{\lfloor \frac{x}{m} \rfloor} \frac{\Lambda_{x-mr,r} \mu_1^{x-mr} \mu_2^r}{(x-mr)!r!} \right)^{-2} - \left[ \sum_{r=0}^{\lfloor \frac{x-1}{m} \rfloor} \frac{\Lambda_{x-mr-1,r} \mu_1^{x-mr} \mu_2^r}{(x-mr-1)!r!} \right] \left( \sum_{r=0}^{\lfloor \frac{x}{m} \rfloor} \frac{\Lambda_{x-mr,r} \mu_1^{x-mr} \mu_2^r}{(x-mr)!r!} \right)^{-2} \tag{4.3}$$

$$\frac{\partial^2}{\partial \rho_2^2} \ln q_x = \sum_{r=0}^{\lfloor \frac{x}{m} \rfloor} \sum_{r=0}^{\lfloor \frac{x}{m} \rfloor} \frac{\Lambda_{x-mr,r} \Lambda_{x-mr,r-2}}{((x-mr)!)^2 r!(r-2)!} (\mu_1^{x-mr} \mu_2^{r-1})^2 \left( \sum_{r=0}^{\lfloor \frac{x}{m} \rfloor} \frac{\Lambda_{x-mr,r} \mu_1^{x-mr} \mu_2^r}{(x-mr)!r!} \right)^{-2} - \left[ \sum_{r=1}^{\lfloor \frac{x}{m} \rfloor} \frac{\Lambda_{x-mr,r-1} \mu_1^{x-mr} \mu_2^r}{(x-mr)!(r-1)!} \right] \left( \sum_{r=0}^{\lfloor \frac{x}{m} \rfloor} \frac{\Lambda_{x-mr,r} \mu_1^{x-mr} \mu_2^r}{(x-mr)!r!} \right)^{-2} \tag{4.4}$$

**References**

1. Consul, P.C.: Generalized Poisson distribution. In: Properties and Applications. Marcel Dekker, New York (1989)
2. Consul, P.C., Famoye, F.: The truncated generalized Poisson distribution and its estimation. Commun. Stat. Theory Methods **18**, 3635–3648 (1989)
3. Dhanavananthan, P.: Compound intervened Poisson distribution. Biometr. J. **40**, 641–646 (1998)
4. Dhanavananthan, P.: Estimation of the parameters of compound intervened Poisson distribution. Biometr. J. **40**, 315–320 (2000)
5. Greenwood, P.E., Nikulin, M.S.: A Guide to Chi-Squared Testing. Wiley, New York (1996)
6. Gupta, R.P., Jain, G.C.: A generalized Hermite distribution and its properties. Siam J. Appl. Math. **27**, 359–363 (1974)
7. Huang, M., Fung, K.Y.: Intervened truncated Poisson distribution. Sankhya Ser. B **51**, 302–310 (1989)
8. Kemp, A.W., Kemp, C.D.: Some properties of the Hermite distribution. Biometrika **52**, 381–394 (1965)
9. Kendall, M.G.: Natural law in science. J. R. Stat. Soc. Ser. A **124**, 1–18 (1961)
10. Kumar, C.S.: Extended generalized hyper geometric probability distributions. Stat. Probab. Lett. **59**, 1–7 (2002)
11. Kumar, C.S., Shibu, D.S.: Modified intervened Poisson distribution. Statistica **71**, 489–499 (2011)
12. Kumar, C.S., Shibu, D.S.: An alternative to truncated intervened Poisson distribution. J. Stat. Appl. **5**, 131–141 (2012)
13. Patel, M.N., Gajjar, A.V.: Estimation from power series distribution. Metron **LVIII**, 171–184 (2000)
14. Puig, P.: Characterizing additively closed discrete models by a property of their MLEs, with an application to generalized Hermite distributions. J. Am. Stat. Assoc. **98**, 687–692 (2003)
15. Saksena, K.M., Kazim, M.A., Pathan, M.A.: Elements of Special Functions. P.C. Dwadash Shreni & Co. (Pvt.) Ltd., Aligarh (1972)
16. Scollnik, D.P.M.: Bayesian analysis of an intervened Poisson distribution. Commun. Stat. Theory Methods **24**, 735–754 (1995)
17. Scollnik, D.P.M.: On the analysis of the truncated generalized Poisson distribution using a Bayesian method. Aust. Bull. **28**, 135–152 (1998)
18. Scollnik, D.P.M.: On the intervened generalized Poisson distribution. Commun. Stat. Theory Methods **35**, 953–963 (2006)
19. Shanmugam, R.: An intervened Poisson distribution and its medical application. Biometrics **41**, 1025–1029 (1985)
20. Shanmugam, R.: An inferential procedure for the Poisson intervention parameter. Biometrics **48**, 559–565 (1992)
21. Williams, C.B.: Number of publications written by biologists. Ann. Eugenics **12**, 143–146 (1944)