Aligarh Journal of Statistics
Vol. 33 (2013), 119-128

ON EXTENDED ALTERNATIVE HYPER-POISSON DISTRIBUTION

C. Satheesh Kumar and B. Unnikrishnan Nair

ABSTRACT

Here we introduce an extended version of the alternative hyper-Poisson distribution of Kumar and Nair (Statistica, 2012a) and study some of its important properties. The maximum likelihood estimation of the parameters of this extended version is also discussed and demonstrated its usefulness with the help of some real life data sets.

1. INTRODUCTION

The hyper-Poisson distribution (HPD) of Bardwell and Crow (1964) has the following probability mass function $(p.m.f.)$, in which $x = 0, 1, 2, \ldots$.

$$f(x) = \frac{\Gamma(\lambda)}{\phi(1;\lambda;\theta)} \frac{\theta^x}{\Gamma(\lambda + x)}, \tag{1.1}$$

where $\lambda > 0$, $\theta > 0$ and $\phi(a;b;x)$ is the confluent hypergeometric series (Mathai and Haubold, (2008) or Slater, (1960)). When $\lambda = 1$, the hyper-Poisson distribution reduces to Poisson distribution and when λ is a positive integer, the distribution is known as the displaced Poisson distribution considered by Staff (1964). Bardwell and Crow (1964) termed the distribution as sub-Poisson when $\lambda < 1$ and super–Poisson when $\lambda > 1$. Various methods of estimation of the parameters of the distribution were discussed in Bardwell and Crow (1964) and Crow and Bardwell (1965). Some queuing theory with hyper-Poisson arrivals has been worked out by Nisida (1962). Roohi and Ahmad (2003a) attempted estimation of the parameters of the hyper-Poisson distribution using negative moments. Roohi and Ahmad (2003b) derived expressions for ascending factorial moments and further obtained certain recurrence relations for negative moments and ascending factorial moments of the hyper-Poisson distribution. Kemp (2002) developed q-analogue of the distribution and Ahmad (2007) introduced and studied Conway-Maxwell hyper-Poisson distribution. Kumar and Nair

(2011, 2012b) developed and studied certain modified form of the HPD. Kumar and Nair (2012a) considered a new version of hyper-Poisson distribution namely `the alternative hyper-Poisson distribution $(AHPD)'$ with $(p.m.f.)$

$$g(y) = \frac{¯^{y}}{(¥)_y} ¤(1+y; ¥ + y; -¯) \tag{1.2}$$

for $y = 0,1,2,...,$ $¥ > 0$, $¯ > 0$

and $(a)_k = a(a+1)...(a+k-1) = \Gamma(a+k)/\Gamma(a),$

for $k = 1,2,...$ and $(a)_0 = 1$. An interesting property of the $AHPD$ is that it is under- dispersed when $¥ < 1$ and over dispersed when $¥ > 1$.

Through this paper, we obtain an extended version of the alternative hyper-Poisson distribution which we call "the extended alternative hyper-Poisson distribution" or in short "the $EAHPD$". In section 2 we establish that the $EAHPD$ possess a random sum structure and it is shown that both the Hermite and generalized Hermite distributions are special cases of the $AHPD$. In section 3, we derive the explicit expressions for its probability mass function, mean and variance. Certain recurrence relations for probabilities, raw moments and factorial moments are also obtained in the section. The maximum likelihood estimation of the parameters of the $EAHPD$ has been discussed in section 4 and section 5 contains some numerical illustrations for emphasizing the usefulness of the model.

2. GENESIS AND SPECIAL CASES

Consider a non-negative integer valued random variable y following the $AHPD$ with $p.m.f.$ (1.2) and probability generating function $(p.g.f.)$

$$G(t) = ¤[1; ¥; ¯(t-1)]. \tag{2.1}$$

Let $\{Z_n, \ n \geq 1\}$ be a sequence of independent and identically distributed random variables, where Y_n has the following $p.g.f.$, in which m is a positive integer.

$$Q(t) = \sim t + (1-\sim)t^m. \tag{2.2}$$

Let $¯_1 > 0$ and $¯_2 \geq 0$ such that $¯ = ¯_1 + ¯_2$ and $\sim = ¯_1 ¯^{-1}$. Suppose that $Y, \ Z_1, Z_2, ...$ are statistically independent and let $S_0 = 0$. Set $S_Y = \sum_{n=0}^{Y} Z_n$. Then the $p.g.f.$ of S_Y is

$$H(t) = E(t^{S_Y})$$

$$= E_Y\{E[t^{S_Y} / Y]\}$$

$$= G(Q(t))$$

$$= ¤[1;¥;¯_1(t-1)+¯_2(t^m-1)]. \qquad (2.3)$$

We define a distribution with $p.g.f.$ (2.3) as ‘the extended alternative hyper-Poisson distribution’ or in short ‘the $EAHPD$’. Clearly, $EAHPD$ with $m=1$ or $¯_2 = 0$ is the $AHPD$. When $¥ = 1$ the EAHPD with p.g.f. (2.3) reduces to the $p.g.f.$ of generalized Hermite distribution of Gupta and Jain (1974) and when $¥ = 1$ and $m = 2$, the $EAHPD$ reduces to the $p.g.f.$ of the Hermite distribution of Kemp and Kemp (1965).

3. SOME IMPORTANT PROPERTIES

In this section we obtain some important properties of the $EAHPD$. Let U be a random variable following the $EAHPD$ with $p.g.f.$

$$H(t) = ¤[1;¥;¯_1(t-1)+¯_2(t^m-1)].$$

$$= \sum_{n=0}^{\infty} h_n(1,¥)t^n, \qquad (3.1)$$

in which $h_n(1,¥) = P(U=n),\ n = 0,1,2,...$. On expanding (3.1) and equating the coefficients of t^n, we get the following result.

Result 3.1 The probability mass function $(p.m.f.)$ $h_n(1,¥)$ of the $EAHPD$ with $p.g.f.$ (3.1) is the following, in which $¢ = n-(m-1)k$.

$$h_n(1,¥) = \sum_{k=0}^{\left[\frac{n}{m}\right]} \frac{¢!}{(¥)_¢} ¤[1+¢;¥+¢;-(¯_1+¯_2)] \frac{¯_1^{n-mk}¯_2^{k}}{(n-mk)!k!} \quad , \qquad (3.2)$$

for $n = 0,\ 1,\ 2,\ \ldots,\ (a)_r = a(a+1)...(a+r-1),$ for $r \ge 1,$ such that $(a)_0 = 1$ and $[k]$ denote the integer part of k. Further we obtain the following results.

Result 3.2 The mean and variance of the $EAHPD$ with $p.g.f.$ (3.1) are

$$E(U) = \frac{1}{¥}(¯_1 + m¯_2), \qquad (3.3)$$

$$Var(U) = \frac{1}{¥}\left(\frac{2}{¥+1} - \frac{1}{¥}\right)(¯_1 + m¯_2)^2 + \frac{1}{¥}(¯_1 + m^2¯_2), \qquad (3.4)$$

The proof is simple and hence omitted.

Remark 3.1 Result 3.2 shows that the $EAHPD$ is over-dispersed (that is, mean less than variance) when $¥ > 1$ and under dispersed when $¥ < 1$ and satisfying the following inequality, for $m > 1$ and $¯_2 > 0$.

$$\frac{(1-¥)}{¥(¥+1)}(¯_1+m¯_2)^2>m(m-1)¯_2$$

Result 3.3 A recurrence relation for the probabilities $h_r(1,¥)$ of the *EAHPD* with $p.g.f.$ (3.1), for $n\geq(m-1)$ is

$$(n+1)h_{n+1}(¥^*)=d_0\left\{¯_1h_n(¥^*+1)+m¯_2h_{n-m+1}(¥^*+1)\right\}, \tag{3.5}$$

in which $¥^*=(1,¥)$, $¥^*+1=(2,¥+1)$ and $d_0=¥^{-1}$.

Proof On differentiating (3.1) with respect to t, we get the following.

$$\sum_{n=0}^{\infty}(n+1)h_{n+1}(¥^*)t^n=\frac{1}{¥}(¯_1+m¯_2t^{m-1})¤[2;¥+1;¯_1(t-1)+¯_2(t^m-1)] \tag{3.6}$$

Replacing $¥$ by $¥+1$ in (3.1) to obtain the following.

$$¤[2;¥+1;¯_1(t-1)+¯_2(t^m-1)]=\sum_{n=0}^{\infty}h_n(¥^*+1)t^n, \tag{3.7}$$

Relations (3.6) and (3.7) together lead to the following relationship:

$$\sum_{r=0}^{\infty}(n+1)h_{n+1}(¥^*)t^n=\frac{1}{¥}\left\{¯_1\sum_{n=0}^{\infty}h_n(¥^*+1)t^n+m¯_2\sum_{n=0}^{\infty}h_n(¥^*+1)t^{m+n-1}\right\} \tag{3.8}$$

Now, on equating the coefficients of t^n on both sides of (3.8) we get (3.5).

Result 3.4 For $n\geq 1$, a recurrence relation for factorial moments $«_{[r]}(¥^*)$ of the *EAHPD* is

$$«_{[r+1]}(¥^*)=d_0¯_1«_{[r]}(¥^*+1)+m\,d_0¯_2\sum_{k=0}^{m-1}\binom{m-1}{k}\frac{r!}{(r-k)!}«_{[r-k]}(¥^*+1), \tag{3.9}$$

in which $«_{[0]}(¥^*)=1$

Proof The factorial moment generating function $F_U(t)$ of the *EAHPD* with $p.g.f.$ (3.1) is the following

$$\begin{aligned}F_U(t)&=H(1+t)\\&=¤[1;¥;¯_1t+¯_2(1+t)^m-1]\\&=\sum_{r=0}^{\infty}«_{[r]}(¥^*)\frac{t^n}{n!}\end{aligned} \tag{3.10}$$

On differentiating (3.10) with respect to 't' to obtain

$$\frac{1}{¥}[¯_1+m¯_2(1+t)^{m-1}]¤[2;¥+1;¯_1t+¯_2(1+t)^m-1]=\sum_{r=0}^{\infty}«_{[r]}(¥^*)\frac{t^{n-1}}{(n-1)!} \tag{3.11}$$

By using (3.10) with $¥$ replaced by $¥+1$ we get the following from (3.11).

$$\sum_{r=0}^{\infty} «_{[r+1]}(¥^*)\frac{t^n}{n!} = \frac{1}{¥}[¯_1 + m¯_2(1+t)^{m-1}]\sum_{r=0}^{\infty} «_{[r]}(¥^*)\frac{t^n}{n!}$$

$$= \frac{1}{¥}¯_1\sum_{r=0}^{\infty} «_{[r]}(¥^*+1)\frac{t^n}{n!} + m¯_2\sum_{r=0}^{\infty}\sum_{k=0}^{m-1}\binom{m-1}{k}«_{[r]}(¥^*+1)\frac{t^{r+k}}{r!} \tag{3.12}$$

On equating coefficients of $(r!)^{-1}t^r$ on both sides of (3.12) we get (3.9).

Result 3.5 For $r \geq 0$, a recurrence relation for the raw moments $«_r(¥^*)$ of the *EAHPD* is

$$«_{r+1}(¥^*) = \frac{1}{¥}\sum_{j=0}^{r}\binom{r}{j}(¯_1 + m^{j+1}¯_2)«_{r-j}(¥^*+1) \tag{3.13}$$

Proof The characteristic function $¤_U(t)$ of the *EAHPD* with $p.g.f.$ (3.1) has the following series representation. For $t \in R$,

$$¤_U(t) = H(e^{it})$$

$$= ¤[1;¥;¯_1(e^{it}-1) + ¯_2(e^{mit}-1)]$$

$$= \sum_{r=0}^{\infty} «_r(¥^*)\frac{(it)^r}{r!} \tag{3.14}$$

On differentiating (3.14) with respect to t, we obtain

$$\frac{1}{¥}(¯_1e^{it} + m¯_2e^{mit})¤[2;¥+1;¯_1(e^{it}-1) + ¯_2(e^{mit}-1)] = \sum_{r=1}^{\infty} «_r(¥^*)\frac{(it)^{r-1}}{(r-1)!} \tag{3.15}$$

By using (3.14) with $¥$ replaced by $¥+1$, we get the following from (3.15).

$$\sum_{r=0}^{\infty} «_{r+1}(¥^*)\frac{(it)^r}{r!} = \frac{1}{¥}(¯_1e^{it} + m¯_2e^{mit})\sum_{r=0}^{\infty} «_r(¥^*+1)\frac{(it)^r}{r!}$$

$$= \frac{1}{¥}\sum_{r=0}^{\infty}\left[¯_1\sum_{j=0}^{\infty}\frac{(it)^j}{j!} + m¯_2\sum_{j=0}^{\infty}\frac{(mit)^j}{j!}\right]«_r(¥^*+1)\frac{(it)^r}{r!} \tag{3.16}$$

On equating the coefficients of $(r!)^{-1}(it)^r$ on both sides we get (3.13).

4. MAXIMUM LIKELIHOOD ESTIMATION

In this section we consider the estimation of the parameters of the *EAHPD* by the method of maximum likelihood. Here we assume that m is a fixed known positive integer. Let a(x) be the observed frequency of *x* events and let y be the highest value of n observed. Then the likelihood function of the sample is

$$L = \prod_{x=0}^{y} [h_x(¥^*)]^{a(x)}. \tag{4.1}$$

Taking logarithm on both sides of (4.1) we get

$$l = \log L = \sum_{x=0}^{k} a(x) \log[h_x(¥^*)]. \tag{4.2}$$

Let $\hat{¯}_1$, $\hat{¯}_2$ and $\hat{¥}$ denote the maximum likelihood estimators of $¯_1$, $¯_2$ and ¥ respectively. Now $\hat{¯}_1$, $\hat{¯}_2$ and $\hat{¥}$ are computed by solving the following equations, obtained from (4.2) on differentiation with respect to $¯_1$, $¯_2$ and ¥ respectively and equating to zero.

$$\frac{\partial l}{\partial ¯_1} = \sum_{x=0}^{y} \sum_{k=0}^{\left[\frac{x}{m}\right]} \Delta(x;k)\{(x-mk)¯_2 ¤[1+¢;¥+¢;-(¯_1+¯_2)] - ¶(¥)\} = 0, \tag{4.3}$$

$$\frac{\partial l}{\partial ¯_2} = \sum_{x=0}^{y} \sum_{k=0}^{\left[\frac{x}{m}\right]} \Delta(x;k)\{k¯_1 ¤[1+¢;¥+¢;-(¯_1+¯_2)] - ¶(¥)\} = 0, \tag{4.4}$$

and

$$\frac{\partial l}{\partial ¥} = \sum_{x=0}^{y} \sum_{k=0}^{\left[\frac{x}{m}\right]} ¯_1 ¯_2 \Delta(x;k)\{¤[1+¢;¥+¢;-(¯_1+¯_2)]¦(¥,¢)$$

$$+ \sum_{j=0}^{\infty} \frac{(1+¢)_j}{(¥+¢)_j} \frac{[-(¯_1+¯_2)]^j}{j!} ¦(¥+j;¢)\} = 0, \tag{4.5}$$

in which ¢ is as given in (3.2), $·(\ ) = [\Gamma(\ )]^{-1}\left[\frac{d}{d}\Gamma(\ )\right]$ for > 0 , $¦(¥;¢) = ·(¥) - ·(¥+¢)$,

$$\Delta(x;k) = a(x)\frac{1}{h_x(¥^*)}\frac{[n-(m-1)k]!¯_1^{n-mk-1}¯_2^{k-1}}{(n-mk)!k!(¥)_¢}, \tag{4.6}$$

for $x = 0,1,...$, $k = 0,1,...$ and $¶(¥) = ¯_1 ¯_2 \frac{1+¢}{¥+¢} ¤[2+¢;¥+¢+1;-(¯_1+¯_2)]$.

5. CONCLUDING REMARKS

In section 2, we have shown that the *EAHPD* possess a random sum structure. Such random sum distributions have found extensive applications in several areas of scientific research. For a detailed account of random sum distributions refer chapter 9 of Johnson *et al*. (2005). Here we consider two real life data sets for demonstrating the estimation procedures discussed in section 4 and for illustrating the usefulness of the model. We have obtained maximum likelihood estimates and variances of the parameters of *EAHPD* for $m = 1, 2, 3, 4$ and using *MATHCAD* software and there by computed expected frequencies, χ^2 - values and p - values. The results obtained are given in Table1 and 2. It can be observed from tables that both *HPD* and *AHPD* are not giving good fit to both data sets, where as the *EAHPD* with $m = 4$ gives the best fit in case of first data set and the *EAHPD* with $m = 3$ gives the best fit in case of second data set compared to the existing model.

Acknowledgement

The authors are grateful to the Chief Editor and the anonymous referee for their valuable suggestions. The second author is particularly thankful to University Grants Commission, New Delhi, India for the financial support.

Table 1: Observed distribution of *Ribes* [Fracker and Brischle 1944] and the expected frequencies computed using *HPD*, *AHPD* and the *EAHPD* for $m = 2, 3, 4$ and 5.

x	Observed frequency	Expected frequency by method of maximum likelihood					
		HPD	*AHPD*	*EAHPD* with $m = 2$	*EAHPD* with $m = 3$	*EAHPD* with $m = 4$	*EAHPD* with $m = 5$
0	43	34.14	33.53	45.18	41.68	40.52	39.48
1	15	22.23	21.68	14.91	19.60	19.17	19.61
2	8	12.61	12.69	13.08	5.76	9.00	9.69
3	6	6.34	6.70	4.02	6.97	4.20	4.76
4	3	2.86	3.19	1.99	3.59	2.91	2.32
5	4	1.17	1.38	0.55	1.10	1.80	1.63
6	0	0.44	0.54	0.20	0.74	1.04	1.04
7	1	0.15	0.19	0.05	0.36	0.58	0.63
Total	80	80	80	80	80	80	80

Estimated value of parameters	$\hat{-}_1 = 4.39$ ¥ $= 6.748$	$\hat{-}_1 = 3.65$ ¥ $= 3.167$	$\hat{-}_1 = 0.436$ $\hat{-}_2 = 0.339$ ¥ $= 1.104$	$\hat{-}_1 = 0.736$ $\hat{-}_2 = 0.216$ ¥ $= 1.374$	$\hat{-}_1 = 33.041$ $\hat{-}_2 = 1.671$ ¥ $= 36.117$	$\hat{-}_1 = 42.938$ $\hat{-}_2 = 1.102$ ¥ $= 43.413$
i^2	8.13	49.694	9.012	207.13	1.869	2.742
P-value	0.004	0.014	0.011	0.097	0.172	0.098

Table 2: Observed distribution of the counts of Red mites on Apple leaves [P.Garman,1951] and the expected frequencies computed using *HPD* and the *AHPD* as well as the *EAHPD* for $m = 2$ and $m = 3$ by method of maximum likelihood.

x	Observed frequency	Expected frequency by method of maximum likelihood			
		HPD	*AHPD*	*EAHPD* with $m = 2$	*EAHPD* with $m = 3$
0	70	60.32	46.12	26.19	69.94
1	38	42.80	58.60	44.89	40.60
2	17	25.24	31.29	43.85	14.64
3	10	12.73	10.67	19.32	12.63
4	9	5.61	2.68	10.37	7.13
5	3	2.19	0.53	3.46	2.60
6	2	0.77	0.09	1.35	1.38
7	1	0.25	0.01	0.39	0.68
8	0	0.07	0.001	0.12	0.24
Total	150	150	150	150	150
Estimated value of parameters		$\hat{-}_1 = 3.485$ ¥ $= 4.911$	$\hat{-}_1 = 0.962$ ¥ $= 0.863$	$\hat{-}_1 = 0.209$ $\hat{-}_2 = 0.153$ ¥ $= 0.519$	$\hat{-}_1 = 0.887$ $\hat{-}_2 = 0.193$ ¥ $= 1.326$
i^2		9.5	32.986	96.68	1.95
P-value		0.009	0.11	0.27	0.52

REFERENCES

Ahmad, M. (2007): A short note on Conway-Maxwell-hyper Poisson distribution. *Pakistan J. Statist.*, **23**, 135-137.

Bardwell, G. E. and Crow, E. L. (1964): A two parameter family of hyper-Poisson distributions. *J. Amer. Statist. Assoc.*. **59**, 133-141.

Beall, G., and Rescia, R. R. (1953): A generalization of Neyman's contagious distributions. *Biometrics*, **9**, 354-386.

Crow, E.L. and Bardwell, G.E. (1965): *Estimation of the parameters of the hyper-Poisson distributions*. *Classical and Contagious Discrete Distributions*. G. P. Patil (editor), 127-140, Pergamon Press, Oxford.

Fracker, S. B. and Brischle, H. A. (1944): Measuring the local distribution of Ribes. *Ecology*, **25**, 283-303.

Gupta, R. P. and Jain, G. C. (1974): A generalized Hermite distribution and its properties. *SIAM J. Appl. Math.*, **27**, 359-363.

Hand, D. J. , Daly, F., Lunn, A.D., Mc. Conway, K. J. and Ostrowski, E. (1994): *A hand book of small data sets*, Chapman and Hall. London.

Kemp, C. D. (2002): q-analogues of the hyper-Poisson distribution. *J. Statist. Plann. Inference*. **101**, 179-183.

Kemp, C.D. and Kemp, A. W. (1965): Some properties of Hermite distribution, *Biometrika*. **52**, 381-394.

Kumar, C. S. and Nair, B. U. (2011): A modified version of hyper-Poisson distribution and its applications. *Journal of Statistics and Applications*. **6**, 23-34.

Kumar, C. S and Nair, B.U. (2012a): An alternative hyper-Poisson distribution, *Statistica*, **72**(**3**), 357-369.

Kumar, C. S. and Nair, B.U. (2012b): An extended version of hyper Poisson distribution and some of its applications, *Journal of Applied Statistical Sciences* **19**(**1**), 81-88.

Mathai, A. M. and Haubold, H.J. (2008): *Special Functions of Applied Sciences*, Springer, NewYork.

Nisida, T. (1962): On the multiple exponential channel queuing system with hyper-Poisson arrivals. *J. Oper. Res. Soc. Japan*, **5**, 57-66.

Ong, S. H. and Lee, P. A. (1979): The noncentral negative binomial distribution. *Biometrics*, **21**, 611-627.

Riordan, J. (1968): *Combinatorial Identities*, Wiley, New York.

Roohi, A. and Ahmad, M. (2003a): Estimation of the parameter of hyper-Poisson distribution using negative moments. *Pakistan J. Statist.*, **19**, 99-105.

Roohi A. and Ahmad, M. (2003b): Inverse ascending factorial moments of the hyper-Poisson probability distribution. *Pakistan J. Statist.*, **19**, 273-280.

Slater, L. J. (1960): *Confluent Hypergeometric Functions*. Cambridge: Cambridge University Press.

Staff, P. J. (1964): The displaced Poisson distribution. *Australian Journal of Statistics*, **6**, 2-20.

Received: 29.07.2012

Received: 20.12.2012

C. Satheesh Kumar and B. Unnikrishnan Nair

Department of Statistics
University of Kerala
Trivandrum-695 581, India

E-mail: drcsatheeshkumar@gmail.com
E-mail: ukswathy@gmail.com