AN ALTERNATIVE HYPER-POISSON DISTRIBUTION

C. Satheesh Kumar, B. Unnikrishnan Nair

1. INTRODUCTION

Bardwell and Crow (1964) introduced a two parameter family of discrete distributions namely the hyper-Poisson distribution through the following probability generating function (p.g.f.).

$$G(t) = \phi(1;\gamma;\theta t) / \phi(1;\gamma;\theta), \qquad (1)$$

in which $\gamma > 0$, $\theta > 0$ and

$$\phi(a;b;z) = 1 + \sum_{k=1}^{\infty} (a)_k z^k / [(b)_k k!]$$

is the confluent hypergeometric series (also called the Kummer M function), in which $(a)_k$ is the rising factorial:

$$(a)_k = a(a+1) \cdot \dots \cdot (a+k-1) = \Gamma(a+k)/\Gamma(a),$$

for k = 1, 2, ... and $(a)_0 = 1$. For details on confluent hypergeometric series see Mathai and Haubold (2008) or Abramowitz and Stegun (1965, chapter 13). When $\gamma = 1$, the hyper-Poisson distribution reduces to the Poisson distribution and when γ is a positive integer, the distribution reduces to the displaced Poisson distribution of Staff (1964). Bardwell and Crow (1964) termed the distribution as sub-Poisson when $\gamma < 1$ and super–Poisson when $\gamma > 1$. The hyper-Poisson distribution is also a member of the Kemp family of distributions studied by Kumar (2009). 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 associated with hyper-Poisson arrivals has been developed by Nisida (1962). The estimation of the parameters of the hyper-Poisson distribution using negative moments were attempted by Roohi and Ahmad (2003a). Roohi and Ahmad (2003a) derived expressions for ascending factorial moments of the hyper-Poisson distribution and obtained certain recurrence relations for its negative moments and ascending factorial moments. Kemp (2002) considered a qanalogue of the distribution and Ahmad (2007) introduced the Conway-Maxwell hyper-Poisson distribution. Kumar and Nair (2011, 2012) developed extended versions of the hyper-Poisson distribution and discussed some of their applications.

In this paper, we consider an alternative form of hyper-Poisson distribution which we named as "the alternative hyper-Poisson distribution (AHP distribution)" and study its important properties. In section 2 we give the definition of AHP distribution and derive its p.g.f., expression for factorial moments, raw moments, mean, variance, and recursion formulae for its probabilities, raw moments and factorial moments. Further the estimation of the parameters of AHP distribution by method of factorial moments, method of mixed moments and method of maximum likelihood are discussed in section 3 and illustrated using real life data sets.

We need the following series representations in the sequel.

$$\sum_{x=0}^{\infty} \sum_{r=0}^{\infty} \mathcal{A}(r,x) = \sum_{x=0}^{\infty} \sum_{r=0}^{x} \mathcal{A}(r,x-r)$$
(2)

$$\sum_{r=0}^{\infty} \sum_{n=r}^{\infty} B(r,n) = \sum_{n=0}^{\infty} \sum_{r=0}^{n} B(r,n)$$
(3)

2. THE AHP DISTRIBUTION

In this section we present the definition of the AHP distribution and obtain some of its important properties.

Definition 2.1. A non-negative integer valued random variable X is said to follow the alternative hyper-Poisson distribution (or in short the AHP distribution) if its probability mass function (p.m.f.) has the following form, in which $\gamma > 0$, $\theta > 0$ and x = 0, 1, 2, ...

$$f(x) = P[X = x]$$

= $\frac{\theta^{x}}{(\gamma)_{x}} \phi(1 + x; \gamma + x; -\theta)$ (4)

Clearly, when $\gamma = 1$ the AHP distribution reduces to the Poisson distribution. An important characteristic of the AHP distribution is that it is under-dispersed when $\gamma < 1$ and over-dispersed when $\gamma > 1$, in the light of Remark 2.1. Now we have the following results.

Result 2.1. The p.g.f G(t) of the AHP distribution with p.m.f. (4) is the following.

$$G(t) = \phi \left[1; \gamma; \theta(t-1)\right] \tag{5}$$

Proof. By definition, the p.g.f. of the AHP distribution with p.m.f. (4) is

$$G(t) = \sum_{x=0}^{\infty} f(x)t^{x}$$
$$= \sum_{x=0}^{\infty} \frac{\theta^{x}}{(\gamma)_{x}} \phi(1+x;\gamma+x;-\theta)t^{x}$$
(6)

On expanding the confluent hypergeometric series in (6), we get

$$G(t) = \sum_{x=0}^{\infty} \sum_{r=0}^{\infty} \frac{\theta^{x} (1+x)_{r}}{(\gamma)_{x} (\gamma+x)_{r}} \frac{(-\theta)^{r} t^{x}}{r!}$$
$$= \sum_{x=0}^{\infty} \sum_{r=0}^{\infty} \frac{(-\theta)^{r}}{r!} \frac{(\theta t)^{x}}{x!} \frac{(x+r)!}{(\gamma)_{x+r}},$$
(7)

since $(\gamma)_x (\gamma + x)_r = (\gamma)_{x+r}$ and $(1+x)_r = (x!)^{-1} (x+r)!$. Now applying (2) in (7) to obtain

$$G(t) = \sum_{x=0}^{\infty} \sum_{r=0}^{x} \frac{(-\theta)^{r}}{r!} \frac{(\theta t)^{x-r}}{(x-r)!} \frac{x!}{(\gamma)_{x}}$$
$$= \sum_{x=0}^{\infty} \frac{1}{(\gamma)_{x}} [\theta(t-1)]^{x},$$
(8)

in the light of binomial expansion of $[\theta(t-1)]^x$. Since $(1)_x = x!$ from (8) we have

$$G(t) = \sum_{x=0}^{\infty} \frac{(1)_{x}}{(\gamma)_{x}} \frac{[\theta(t-1)]^{x}}{x!}$$

which is (5).

Result 2.2. An expression for factorial moments $\mu_{[r]}$ of the AHP distribution is the following, for $r \ge 1$.

$$\mu_{[r]} = \frac{r!\theta^r}{(\gamma)_r} \tag{9}$$

Proof. The factorial moment generating function F(t) of the AHP distribution with p.g.f. (5) is

$$F(t) = G(1+t)$$
$$= \phi(1;\gamma;\theta t).$$

On expanding $\phi(1;\gamma;\theta t)$ and equating the coefficients of $(r!)^{-1} t^r$, we get (9).

Result 2.3. Mean and variance of AHP distribution are

Mean =
$$\frac{\theta}{\gamma}$$

and

Variance =
$$\frac{\theta}{\gamma} \left[1 + \frac{\theta}{\gamma} \frac{(\gamma - 1)}{(\gamma + 1)} \right]$$

Remark 2.1. From Result (2.3) it is obvious that the AHP distribution is underdispersed (that is, mean greater than variance) when $\gamma < 1$ and over dispersed when $\gamma > 1$

Result 2.4. An expression for raw moments μ_n of the AHP distribution is the following, for $n \ge 0$.

$$\mu_n = \sum_{r=0}^n \frac{\theta^r r!}{(\gamma)_r} S(n,r), \qquad (10)$$

where S(n,r) is the Stirling numbers of the second kind. For details see, (Riordan, 1968).

Proof. The characteristic function $\psi(t)$ of the AHP distribution with p.g.f. (5) is the following, for any $t \in \mathbb{R} = (-\infty, \infty)$ and $i = \sqrt{-1}$.

$$\psi(t) = \sum_{n=0}^{\infty} \mu_n \frac{(it)^n}{n!}$$

$$= G(e^{it})$$

$$= \phi[1;\gamma;\theta(e^{it}-1)]$$
(11)

On expanding the confluent hypergeometric function and using the fact $(1)_r = r!$, we get

$$\psi(t) = \sum_{r=0}^{\infty} \frac{\left[\theta(e^{it} - 1)\right]^r}{(\gamma)_r}$$
$$= \sum_{r=0}^{\infty} \sum_{n=r}^{\infty} \frac{\theta^r r!}{(\gamma)_r} S(n, r) \frac{(it)^n}{n!},$$
(12)

since by the equation (1.57) of (Johnson et.al., 2005). By applying (3) in (12) we obtain

$$\psi(t) = \sum_{n=0}^{\infty} \sum_{r=0}^{n} \frac{\theta^{r} r!}{(\gamma)_{r}} S(n, r) \frac{(it)^{n}}{n!}$$
(13)

On equating the coefficient of $(n!)^{-1}(it)^n$ on right hand side expressions of (11) and (13), we get (10).

Define the following shorter notations, which we need in the sequel.

$$\gamma^* = (1, \gamma)$$

and

$$\gamma^* + j = (1 + j, \gamma + j),$$

for j = 1, 2, ...

Result 2.5. The following is a simple recursion formula for probabilities $f_x(\gamma^*, \theta) = f(x)$ of the AHP distribution with p.g.f. (5), for $x \ge 0$.

$$f_{x+1}(\gamma^*,\theta) = \frac{\theta}{\gamma(x+1)} f_x(\gamma^*+1,\theta)$$
(14)

Proof. From (4) we have

$$G(t) = \phi [1; \gamma; \theta(t-1)]$$
$$= \sum_{x=0}^{\infty} f_x(\lambda^*, \theta) t^x.$$
(15)

On differentiating (15) with respect to t, we obtain the following.

$$\sum_{x=0}^{\infty} (x+1) f_{x+1}(\gamma^*, \theta) t^x = \frac{\theta}{\gamma} \phi [2; \gamma+1; \theta(t-1)].$$
(16)

Also, from (15) we have

$$\phi[2;\gamma+1;\theta(t-1]] = \sum_{x=0}^{\infty} f_x(\gamma^*+1,\theta)t^x$$
(17)

Relations (16) and (17) together lead to the following.

$$\sum_{x=0}^{\infty} (x+1) f_{x+1}(\gamma^*, \theta) t^x = \frac{\theta}{\gamma} \sum_{x=0}^{\infty} f_x(\gamma^*+1, \theta) t^x .$$
(18)

on equating the coefficients of t^x on both sides of (18) we get (14).

Result 2.6. The following is a recursion formula for raw moments $\mu_r(\gamma^*) = \mu_r$ of the AHP distribution, for $r \ge 0$.

$$\mu_{r+1}(\gamma^{*}) = \frac{\theta}{\gamma} \sum_{k=0}^{r} \binom{r}{k} \mu_{r-k}(\gamma^{*}+1)$$
(19)

Proof. On differentiating (11) with respect to t, we get the following.

$$\frac{\theta}{\gamma} e^{it} \phi [2; \gamma + 1; \theta(e^{it} - 1)] = \sum_{r=1}^{\infty} \mu_r (\gamma^*) \frac{(it)^{r-1}}{(r-1)!}$$
(20)

From (11) we have

$$\phi[2;\gamma+1;\theta(e^{it}-1)] = \sum_{r=1}^{\infty} \mu_r(\gamma^*+1) \frac{(it)^r}{r!}.$$
(21)

Equations (20) and (21) lead to the following.

$$\sum_{r=0}^{\infty} \mu_{r+1}(\gamma^{*}) \frac{(it)^{r}}{r!} = \frac{\theta}{\gamma} e^{it} \sum_{r=0}^{\infty} \mu_{r}(\gamma^{*}+1) \frac{(it)^{r}}{r!}$$
$$= \frac{\theta}{\gamma} \sum_{r=0}^{\infty} \sum_{k=0}^{\infty} \frac{(it)^{k}}{k!} \mu_{r}(\gamma^{*}+1) \frac{(it)^{r}}{r!}$$
$$= \frac{\theta}{\gamma} \sum_{r=0}^{\infty} \sum_{k=0}^{r} \mu_{r-k}(\gamma^{*}+1) \frac{(it)^{r}}{k!(r-k)!}, \qquad (22)$$

in the light of (2). Now, on equating the coefficients of $(r!)^{-1} (it)^r$ on both sides of (22) we get (19).

Result 2.7. The following is a simple recursion formula for factorial moments $\mu_{[r]}(\gamma^*) = \mu_{[r]}$ of the AHP distribution, for $r \ge 1$, in which $\mu_{[0]}(\gamma^*) = 1$.

$$\mu_{[r+1]}(\gamma^*) = \frac{\theta}{\gamma} \mu_{[r]}(\gamma^* + 1), \qquad (23)$$

Proof. The factorial moment generating function F(t) of the AHP distribution with p.g.f. given in (5) has the following series representation.

$$F(t) = G(1+t)$$

= $\phi(1;\gamma;\theta t)$
= $\sum_{r=0}^{\infty} \mu_{[r]}(\gamma^*) \frac{t^r}{r!}$ (24)

On differentiating (24) with respect to t to get

$$\frac{\theta}{\gamma}\phi(2;\gamma+1;\theta\,t) = \sum_{r=1}^{\infty}\mu_{[r]}(\gamma^*)\frac{t^{r-1}}{(r-1)!}$$
(25)

By using (24) we get the following from (25).

$$\sum_{r=0}^{\infty} \mu_{[r+1]}(\gamma^*) \frac{t^r}{r!} = \frac{\theta}{\gamma} \sum_{r=0}^{\infty} \mu_{[r]}(\gamma^* + 1) \frac{t^r}{r!}$$
(26)

Now, on equating coefficients of $(r!)^{-1} t^r$ on both sides of (26) we get (23). \Box

3. ESTIMATION

Here we consider the estimation of the parameters γ and θ of the AHP distribution by method of factorial moments, method of mixed moments and the method of maximum likelihood.

3.1 Method of factorial moments

In method of factorial moments, the first two factorial moments $\mu_{[1]}$, $\mu_{[2]}$ of the AHP distribution are equated to the corresponding sample factorial moments, say $m_{[1]}$, $m_{[2]}$. Thus we obtain the following system of equations.

$$\frac{\theta}{\gamma} = m_{[1]} \tag{27}$$

$$\frac{2.\theta^2}{\gamma(\gamma+1)} = m_{[2]} \tag{28}$$

On solving (27) and (28) we obtain the factorial moment estimators $\overline{\gamma}$ and $\overline{\theta}$ of γ and θ of the AHP distribution as

$$\overline{\gamma} = \frac{m_{[2]}}{m_{[1]}^2 - m_{[2]}}$$

and

$$\overline{\theta} = \frac{m_{[1]} m_{[2]}}{m_{[1]}^2 - m_{[2]}}$$

3.2 Method of mixed moments

In method of mixed moments, the parameters are estimated by using the first sample factorial moment and the first observed frequency of the distribution. That is, the estimators are obtained by solving the following equation together with (29).

$$N\phi^{-1}(1;\gamma;-\theta) = p_0, \qquad (29)$$

where p_0 is the observed frequency of the distribution corresponding to the observation zero and N is the total observed frequency.

3.3 Method of maximum likelihood

Let a(x) be the observed frequency of x events based on the observations from a sample with independent components and let y be the highest value of x observed. Then the likelihood function of the sample is

$$L = \prod_{x=0}^{\mathcal{Y}} [f(x)]^{a(x)}$$

which implies

$$\log L = \sum_{x=0}^{y} a(x) \log f(x)$$

Assume that γ is known. Let $\hat{\theta}$ denote the maximum likelihood estimate of θ . Now $\hat{\theta}$ is obtained by solving the normal equation (30) given below.

$$\frac{\partial \log L}{\partial \theta} = 0$$

Equivalently,

$$\sum_{x=0}^{y} a(x) \left[\frac{x}{\theta} - \frac{\frac{1+x}{\gamma+x}\phi(2+x;\gamma+x+1;-\theta)}{\phi(1+x;\gamma+x;-\theta)} \right] = 0$$
(30)

Here the estimate of γ is used for obtaining the maximum likelihood estimator $\hat{\theta}$ of θ . Let $\overline{\gamma}$ denote the factorial moment estimator of γ and $\tilde{\gamma}$ denote the mixed moment estimator of γ . All these procedures discussed in this section are illustrated using two real life data sets, obtained from (Albert, 1991) [or see page 133 of (Hand *et al.*, 1994)] and (Stirrett et.al., 1937) [or Bliss, 1953)] with the help of the mathematical software – MATHCAD and presented in Table 1 and Table 2. From these tables it can be viewed that the AHP distribution gives better fit compared to the hyper-Poisson distribution. Further study on the properties and the comparisons of the estimators of the parameters of the AHP distribution obtained in the paper will be published in the sequel.

	function and and	Expecte	ected frequency by	Expected	Expected frequency by	Exp	sected frequency by	Expected frequency by method of maximum likelihood estimation	ikelihood estimation
	rrequercy	method of	of Factorial moments	method of	method of mixed moments		using $\overline{\gamma}$		using $\tilde{\gamma}$
		Hyper-Poisson distribution	Alternative hyper-Poisson distribution	Hyper-Poisson distribution	Alternative hyper-Poisson distribution	Hyper-Poisson distribution	Alternative hyper-Poisson distribution	Hyper-Poisson distribution	Alternative hyper-Poisson distribution
0	126	121.20	122.19	126.01	126.00	120	122.16	126.01	124.97
_	80	88.37	87.28	87.83	87.06	88.07	87.27	87.83	86.72
~1	59	59.46	58.86	57.52	57.35	59.62	58.86	57.52	57.43
~	42	37.14	37.23	35.54	35.84	37.48	37.24	35.54	36.11
4	24	21.64	21.99	20.77	21.18	21.98	22.00	20.77	21.49
10	8	11.82	12.10	11.52	11.81	12.08	12.11	11.52	12.07
2	5	6.08	6.20	6.08	6.20	6.25	6.20	6.08	6.39
2	4	2.95	2.96	3.06	3.07	3.05	2.96	3.06	3.19
~	3	1.36	1.32	1.47	1.43	1.41	1.32	1.47	1.50
[otal	351	351	351	351	351	351	351	351	351
Estimated value of	ralue of	$\overline{\gamma} = 11.940$	$\overline{y} = 3.528$	$\tilde{y} = 15.602$	$\tilde{y} = 4.115$	$\hat{\theta} = 8.682$	$\hat{\theta} = 5.450$	$\hat{\theta} = 10.874$	$\hat{\theta} = 6.328$
parameters		$\overline{\theta} = 8.710$	$\overline{\theta} = 5.448$	$\tilde{ heta}=10.874$	$\tilde{ heta} = 6.253$				
χ^2 - value		3.541	3.285	4.473	3.795	3.380	3.285	4.473	4.12
P - value		0.738	0.772	0.724	0.704	0.760	0.772	0.724	0.765

TABLE 1

2	
ц	
B	
7	
Ц	

Count Observed		Expected frequency by method	Expected fre	Expected frequency by method	Щ	xpected frequency by	Expected frequency by method of maximum likelihood estimation	likelihood estimation
frequency		of Factorial moments	ofmix	of mixed moments	-	using $\overline{\gamma}$		using \tilde{p}
	Hyper-Poisson distribution	Alternative hyper-Poisson distribution	Hyper-Poisson distribution	Alternative hyper-Poisson distribution	Hyper-Poisson distribution	Alternative hyper-Poisson distribution	Hyper-Poisson distribution	Alternative hyper-Poisson Distribution
19	12.65	17.38	18.96	18.96	10.45	17.35	14.62	18.87
12	16.90	16.62	19.62	17.48	14.73	16.59	16.31	17.41
18	19.20	15.70	18.73	15.91	17.65	15.67	16.80	15.86
18	18.97	14.53	16.61	14.23	18.40	14.51	16.07	14.19
11	16.58	13.06	13.75	12.43	16.97	13.06	14.35	12.42
12	12.99	11.30	10.67	10.55	14.04	11.30	12.01	10.55
7	9.23	9.31	7.79	8.63	10.52	9.32	9.47	8.65
8	5.99	7.28	5.38	6.78	7.21	7.29	7.05	6.82
4	3.59	5.37	3.51	5.09	4.55	5.38	4.97	5.13
4	1.99	3.72	2.18	3.64	2.66	3.74	3.33	3.68
1	1.03	2.43	1.29	2.48	1.45	2.45	2.12	2.52
9	0.50	1.50	0.73	1.60	0.74	1.51	1.29	1.63
otal 120	120	120	120	120	120	120	120	120
Estimated value of	$\overline{\gamma} = 5.677$	$\overline{\gamma} = 2.278$	$\tilde{y} = 12.015$	$\tilde{y} = 2.633$	$\hat{\theta} = 8.003$	$\hat{\theta} = 8.519$	$\hat{\theta} = 13.410$	$\hat{\theta} = 9.674$
parameters	$\overline{\theta} = 3.776$	$\overline{\theta} = 8.502$	$\tilde{ heta}=12.428$	$\tilde{ heta} = 9.624$				
² - value	7.958	4.67	12.095	4.799	14.786	4.186	8.428	4.269
P - value	0.088	0.846	0.147	0.851	0.097	0.899	0.494	0.893

ACKNOWLEDGMENTS

The authors would like to express their sincere thanks to the Editor and the anonymous Referee for carefully reading the paper and for valuable comments. The second author is particularly thankful to UGC, Govt. of India for financial support.

Department of Statistics University of Kerala C. SATHEESH KUMAR B. UNNIKRISHNAN NAIR

REFERENCES

M. ABRAMOWITZ, I.A. STEGUN, (1965), Hand book of Mathematical Functions, Dover, New York.

- M. AHMAD, (2007), A short note on Conway-Maxwell-hyper Poisson distribution, "Pakistan Journal of Statistics", 23, pp. 135-137.
- P.S. ALBERT, (1991), A two state Markov mixture model for a time series of epileptic seizure counts, "Biometrics", 47, pp. 1371-1381.
- G.E. BARDWELL, E.L. CROW, (1964), A two parameter family of hyper-Poisson distributions, "Journal of American Statistical Association", 59, pp. 133-141.
- C.I. BLISS, (1953), Fitting the negative binomial distribution to biological data, "Biometrics", 9, pp. 176-200.
- S. CHAKRAVORTHY, (2010), On some distributional properties of the family of weighted generalized Poisson distribution, "Communications in Statistics-Theory and methods", 39, pp. 2767-2788.

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

- D.J. HAND, F. DALY, A.D. LUNN, K.J. MCCONWAY, E. OSTROWSKI, (1994), A Hand Book of Small Data Sets, Chapman and Hall, London.
- N.L. JOHNSON, A.W. KEMP, S. KOTZ, (2005), Univariate Discrete Distributions, Wiley, New York.
- C.D. KEMP, (2002), *q-analogues of the hyper-Poisson distribution*, "Journal of Statistical Planning and Inference", 101, pp. 179-183.
- c.s. KUMAR, (2009), Some properties of Kemp family of distributions, "Statistica", 69, pp. 311-316.
- C.S. KUMAR, B.U. NAIR, (2011), A modified version of hyper Poisson distribution and its applications, "Journal of Statistics and Applications", 6, pp. 25-36.
- C.S. KUMAR, B.U. NAIR, (2012), An extended version of byper Poisson distribution and some of its applications, "Journal of Applied Statistical Sciences", 19 (To appear).
- A.M. MATHAI, H.J. HAUBOLD, (2008), Special functions of applied sciences, Springer, New York.
- T. NISIDA, (1962), On the multiple exponential channel queuing system with hyper-Poisson arrivals, "Journal of the Operations Research Society", 5, pp. 57-66.
- J. RIORDAN, (1968), *Combinatorial identities*, Wiley, NewYork.
- A. ROOHI, M. AHMAD, (2003a), Estimation of the parameter of hyper-Poisson distribution using negative moments, "Pakistan Journal of Statistics", 19, pp. 99-105.
- A. ROOHI, M. AHMAD, (2003a), Inverse ascending factorial moments of the hyper-Poisson probability distribution, "Pakistan Journal of Statistics", 19, pp. 273-280.
- P.J. STAFF, (1964), *The displaced Poisson distribution*, "Australian Journal of Statistics", 6, pp. 12-20.
- G.M. STIRRETT, G.BEALL, M.TIMONIN, (1937), A field experiment on the control of the European corn borer, "Pyrausta nubilalis Hubn". by Beauveria bassiana Vuill. Scient. Agric, 17, pp. 587-591.

SUMMARY

An alternative hyper-Poisson distribution

An alternative form of hyper-Poisson distribution is introduced through its probability mass function and studies some of its important aspects such as mean, variance, expressions for its raw moments, factorial moments, probability generating function and recursion formulae for its probabilities, raw moments and factorial moments. The estimation of the parameters of the distribution by various methods are considered and illustrated using some real life data sets.