ON SOME GENERALIZED COMPOUND DISCRETE DISTRIBUTIONS*

S. U. Badahdah†, M. Ahmad, and M. N. El-Derini‡

Department of Mathematical Sciences, University of Petroleum and Minerals, Dhahran, Saudi Arabia

الخلاصـة :

لقد أعيدت كتابة دالة توليد الإحتمالات التي أشتقها بها ليراو وقرلند [٣] بواسطة المتسلسلة فوق الهندسية المدمجة وعممت إلى العوائل ذات الأربعة معالم وذات الخمسة معالم .

وعوائل التوزيع المعممة تختلف عن عوائل كاتز [٨] وكمب [٩] . وقد أستخدمت طريقة العزم للحصول على تقدير للمعالم . وقد شرح مثال لتوضيح الطريقة .

ABSTRACT

The probability generating function derived by Bhalerao and Gurland [3] has been rewritten in terms of confluent hypergeometric series functions and generalized to four- and five-parameter families. The generalized family of distributions is different from the Katz [8] and Kemp [9] families. The moment method has been employed to obtain an estimation of the parameters. An example is given to illustrate the method.

---

*This paper was presented at the American Statistical Association Meeting, 1980 at Houston, Texas, USA.

† Present address: *Department of Statistics, Colorado State University, Fort Collins, Colorado, USA.*

‡ Present address: *Computer Engineering Department, University of Alexanaria, Alexandria, Egypt.*

0377–9211/83/010021–05$01.00

S. U. Badahdah, M. Ahmad and M. N. El-Derini

ON SOME GENERALIZED COMPOUND DISCRETE DISTRIBUTIONS

1. INTRODUCTION

Bhalerao and Gurland [3] introduced the probability generating function (p.g.f.)

$$g(z)=\exp\left\{\lambda\left[\left\{1-\frac{\beta}{1-\beta}(z-1)\right\}^{-\alpha/\beta}-1\right]\right\}, \quad (1)$$

where $\lambda>0$, $\alpha>0$, and $\beta<1$. When $\beta<0$, $-\alpha/\beta$ is a positive number. The function (1) is the p.g.f. of a three-parameter family of generalized Poisson distributions and was named Poisson V POLPAB, as it was a mixture of Poisson, Logarithmic, Pascal, and Binomial distributions. This family will be referred to as the B–G family.

In this paper, a generalization of the B–G family is given. In addition, explicit formulae for the density function, moment generating function, and moments are presented, and parameters are estimated using the first two moments and the frequency at zero.

2. GENERALIZATION OF B–G FAMILY

The three-parameter p.g.f. (1) can be rewritten in terms of confluent hypergeometric function as

$$g(z)=\frac{{}_1F_1[1;1;\lambda b(z)]}{{}_1F_1[1;1;\lambda]}, \quad (2)$$

where

$$b(z)=\left[1-\frac{\beta}{1-\beta}(z-1)\right]^{-\alpha/\beta},$$

$${}_1F_1[a;b;t]=\sum_{n=0}^{\infty}\frac{(a)_n}{(b)_n}\frac{t^n}{n!},$$

and

$$(r)_n=r(r+1)\ldots(r+n-1),\ r>0.$$

This has motivated the generalization to a four- or five-parameter family of discrete distributions. This is done by introducing two more parameters in the arguments of the confluent hypergeometric function. The new p.g.f. will, therefore, be given by

$$g_*(z)=k\,{}_1F_1[a;\theta;\lambda b(z)], \quad (3)$$

where $k=({}_1F_1[a;\theta;\lambda])^{-1}$ and $a>0$, $\lambda>0$, and $\theta>0$. The function $g_*(z)$ is evidently a p.g.f., for $g_*(z)$ converges absolutely at least for $|\frac{a}{\theta}\cdot\frac{\lambda}{n}|\leq 1$, since $f^{(n)}(z)$ constitutes a bounded sequence of real numbers and $g_*(1)=1$.

The probability function as the coefficient of z^x in the expansion of $g_*(z)$ is

$$f(x)=k\sum_{n=0}^{\infty}\frac{(a)_n}{(\theta)_n}\frac{\lambda^n}{n!}(1-\beta)^{n\alpha/\beta}\left[\binom{-n\alpha/\beta}{x}(-\beta)^x\right],\quad x=0,1,2,3,\ldots \quad (4)$$

If $a=\theta=1$, $f(x)$ in (4) is the B–G probability function. If $\beta\in(0,1)$, then $f(x)$ in (4) is the Poisson–Pascal distribution, and if $\beta<0$, then $f(x)$ is a Poisson–binomial type. If α/β is an integer, then $f(x)$ reduces to a Poisson–binomial distribution. Other special and limiting cases are discussed by Badahdah [1]. The moment generating function for (4) is

$$M(t)=k\,{}_1F_1[a;\theta;\lambda b(e^t)] \quad (5)$$

The moments from (5) can easily be obtained using the following property of the confluent hypergeometric function [4, p.283],

$${}_1F_1(a;\theta;\lambda x)=\sum_{n=0}^{\infty}\frac{(a)_n(x-1)^n\lambda^n}{(\theta)_n n!}\,{}_1F_1(a+n,\theta+n,\theta).$$

In the derivation of moments, the following identity, proof of which is simple, is also used

$$\sum_{n=j}\frac{(n)_j}{(\theta)_n}\lambda^n=\frac{j!\lambda^j}{(\theta)_j}\,{}_1F_1[j+1;\theta+j;\lambda].$$

Alternatively, the moments of the distribution (4) can be expressed as a linear combination of moments of the negative binomial probability distribution. The moment generating function, $M(t)$ is defined as

$$M(t)=E(e^{tx})=k\sum_{n=0}^{\infty}\frac{(a)_n}{(\theta)_n}\frac{\lambda^n}{n!}(1-P/Q)^N\times\sum_{x=0}^{\infty}\binom{-N}{x}\left(-\frac{P}{Q}\right)^x e^{tx}, \quad (6)$$

where $P=\beta/(1-\beta)$, $Q=1/(1-\beta)$, and $N=n\alpha/\beta$. The rth moment about the origin is the coefficient of $t^r/r!$ in the expansion (6) and is given by

$$\mu_r'=k\sum_{n=0}^{\infty}\frac{(a)_n}{(\theta)_n}\frac{\lambda^n}{n!}\sum_{x=0}^{\infty}x^r(1-P/Q)^N\binom{N+x-1}{N-1}\left(\frac{P}{Q}\right)^x$$

$$=\sum_{n=0}^{\infty}c_n\mu_r'(N,P),\ r=1,2,3,\ldots,$$

where $c_n = [(a)_n/(\theta)_n]\lambda^n/n!$ and $\mu'_r(N,P)$ is the rth moment about the origin of the negative binomial probability function with parameters N and P.

If $a=1$, the first four moments are

$$\mu'_1 = \frac{k\alpha}{1-\beta}\,\frac{\lambda}{\theta}\,{}_1F_1(2;\theta;\lambda), \tag{7}$$

$$\mu'_2 = \frac{\mu'_1}{(1-\beta)}\left[\frac{2\alpha\lambda}{(\theta+1)}F(3)+(1+\alpha)\right], \tag{7a}$$

$$\mu'_3 = \frac{\mu'_1}{(1-\beta)^2}\left[\frac{6\alpha^2\lambda^2}{(\theta+1)_2}F(4)+\frac{6\lambda\alpha(\alpha+1)}{(\theta+1)}F(3) + (\alpha^2+3\alpha+\beta+1)\right], \tag{8}$$

and

$$\mu'_4 = \frac{\mu'_1}{(1-\beta)^3}\left[\frac{24\alpha^3\lambda^3}{(\theta+1)_3}F(5) + 36\alpha^2(\alpha+1)\frac{\lambda^2}{(\theta+1)_2}F(4) + 2(7\alpha^3+18\alpha^2+4\alpha\beta+7\alpha)\frac{\lambda}{\theta+1}F(3) + (\alpha^3+6\alpha^2+7\alpha+4\alpha\beta+\beta^2+4\beta+1)\right], \tag{9}$$

where $F(r) = {}_1F_1[r;\theta+r-1;\lambda] \div {}_1F_1[2;\theta;\lambda]$, $r=3,\ 4,\ 5$.

3. ESTIMATION WHEN $a=1$

Using a sample moment method, we obtain the moment estimators of the parameters. With θ and λ known, the moment estimators of α and β can be obtained as

$$\tilde\alpha = \theta(1-\tilde\beta)m'_1/[k\lambda F(2)], \tag{10}$$

$$\tilde\beta = 1-\frac{\lambda k m'_1}{\lambda k m'_2 - a(\lambda,\theta)}, \tag{11}$$

where $a(\lambda,\theta)=\theta m_1'^2\left(1+\frac{2\lambda F(3)}{\theta+1}\right)$ and m'_i is the ith sample moment about zero.

As the use of the third moment reduces the efficiency of an estimator, we may use zero frequency and the first two sample moments to estimate α, β, and λ, when $\theta=1$. Hinz and Gurland [7] observed that the estimators based on zero frequency and low-order sample moments attain high asymptotic efficiency. The estimators of α, β, and λ using the frequency at zero and the first two sample moments are

$$\tilde\lambda = \frac{1+\tilde\alpha}{\tilde\alpha}\,\frac{m_1'^2}{m_2}, \tag{12}$$

$$\tilde\beta = 1-\frac{\tilde\alpha\tilde\lambda}{m'_1}, \tag{13}$$

and

$$M = \frac{1+\tilde\alpha}{\tilde\alpha}\left[1-\{c(1+\tilde\alpha)\}^{\frac{\tilde\alpha}{1-c(1+\tilde\alpha)}}\right], \tag{14}$$

where $\quad M=\frac{m_2}{m_1'^2}\ln f_0^{-1}\quad$ and $\quad c=\frac{m'_1}{m_2}$.

The M-function in (14) involves α only. The value of α is estimated when c and M are known. If $c<1$, then $M>1$, and the M-function is a monotonically decreasing function of α with $M=1$ as an asymptote. If $c=1$ then $M=1$ for all α. The M-function is a monotonically increasing function of α for $c>1$ with $M=1$ as the asymptote. It can easily be shown that

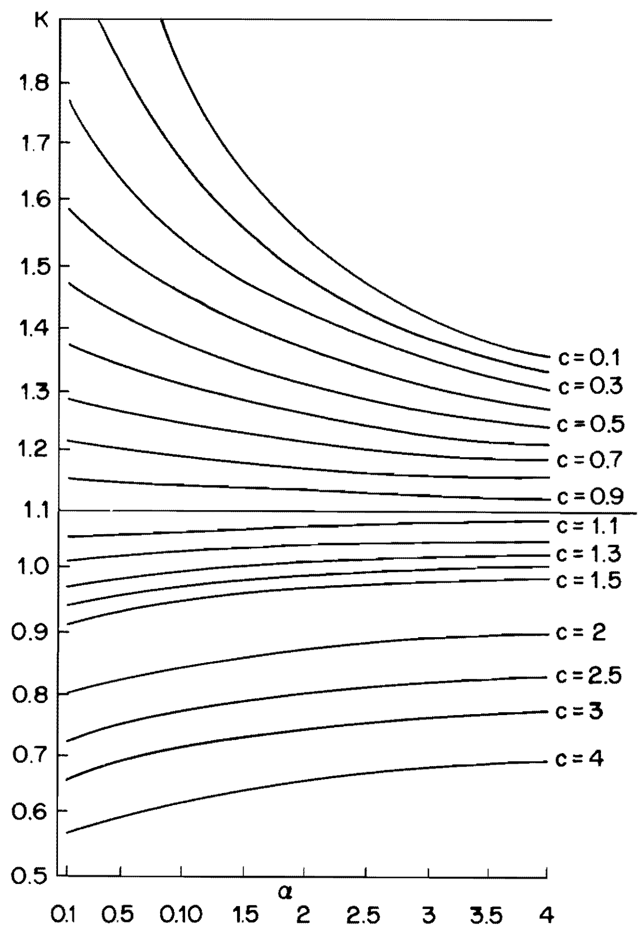

Figure 1. Graphs of the M-function

$\lim_{\alpha \to \infty} [c(1+\alpha)]^{\alpha/\{1-c(1+\alpha)\}} = 0$. When $c \to \infty$, $M \to 0$ for all α. Graphs of the M-function are drawn (see Figure 1) for different values of α and c. For given values of M and c, a value of α can be interpolated. A table of M-functions for various values of α and c has also been constructed for computational purposes (see Table 2).

4. EXAMPLE

Williford and Price [10] examined three distinct categories of physical situations. They have fitted modified compound distributions to various types of data using the method of modified minimum chi-square estimation. They found that some of the modified distributions would provide a better fit than either the Poisson, binomial, or negative binomial distributions.

We consider the data on the frequency of days with X thunderstorm outcomes. The three-parameter generalized compound distribution is fitted to the data using frequency at zero and first two sample moments. Table 1 contains estimates of the three parameters, and observed and expected frequencies. For comparative purposes, the negative binomial distribution is included. The moment estimation method (MM) is less efficient than the maximum likelihood (ML) estimation method, though some of their asymptotic properties are roughly the same. The fit seems better than the negative binomial distribution using the maximum likelihood method. In the example, the negative value of β suggests that the data follow Poisson–binomial type distribution.

Table 1. Frequency of Days with X Thunderstorm Outcomes

No. of days (x)	Observed frequency	Expected frequency: Negative binomial (ML)	Expected frequency: Generalized compound (MM)
0	511	496.33	511.0
1	216	239.21	201.8
2	96	105.90	112.2
3	65	45.50	54.2
4	24	19.25	30.0
5 or more	8	13.81	10.8
	920	920.0	920.0
$\bar{x}=0.79$,	$s^2=1.21$	$\chi^2=12.00$	7.6
$f_o=0.5654$	$\tilde{\alpha}=1.674$	df$=3$	2
$c=0.6529$	$\tilde{\beta}=0.8239$	% 0.993	0.97
$k=1.14003$	$\tilde{\lambda}=-0.7459$		

Table 2. Table of M-functions

α \ c	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	2.416	1.941	1.678	1.500	1.368	1.265	1.182	1.112	1.052
0.5	2.017	1.730	1.548	1.416	1.313	1.229	1.158	1.098	1.046
1.0	1.733	1.566	1.442	1.345	1.264*	1.196	1.138	1.086	1.041
1.5	1.563	1.458	1.370	1.295*	1.230	1.173	1.122	1.077	1.037
2.0	1.452	1.383	1.318	1.258	1.204	1.155	1.111	1.071	1.034
2.5	1.375	1.328	1.278	1.229	1.183	1.141	1.101	1.065	1.031
3.0	1.320	1.286	1.247	1.206	1.167	1.129	1.094	1.060	1.029
3.5	1.278	1.254	1.222	1.187	1.153	1.119	1.087	1.056	1.027
4.0	1.245	1.227*	1.201	1.172	1.141	1.111	1.082	1.053	1.026

*Values are computed on the basis of $\alpha+0.0005$.

α \ c	1.1	1.2	1.3	1.4	1.5	2	2.5	3	4
0.1	0.955	0.914	0.878	0.845	0.816	0.700	0.618	0.556	0.469
0.5	0.959	0.922	0.889	0.857	0.831	0.072	0.641	0.580	0.492
1.0	0.963	0.930	0.899	0.871	0.845	0.740	0.663	0.602	0.514
1.5	0.966	0.936	0.907	0.881	0.856	0.755	0.679	0.620	0.531
2.0	0.969	0.940	0.913	0.888	0.865	0.767	0.693	0.634	0.545
2.5	0.971	0.944	0.918	0.895	0.872	0.778	0.705	0.646	0.557
3.0	0.973	0.947	0.923	0.900	0.878	0.786	0.714	0.656	0.568
3.5	0.974	0.950	0.926	0.904	0.884	0.794	0.723	0.665	0.577
4.0	0.975	0.952	0.930	0.908	0.888	0.801	0.731	0.673	0.585

ACKNOWLEDGMENT

The authors are indebted to the referees for their helpful comments. M. Ahmad is also grateful to the UPM Research Committee for the award of a travel grant in order to attend ASA meetings in Houston, Texas, U.S.A.

REFERENCES

[1] S. U. Badahdah, 'On a New Generalized Discrete Distribution', *Unpublished M.S. Thesis*, University of Petroleum and Minerals, Dhahran, Saudi Arabia, 1980.

[2] S. U. Badahdah, M. Ahmad, and M. N. El-Derini, 'On Some Generalized Mixed Discrete Distributions', *Technical Report No. 34*, Dept. of Math. Sciences, University of Petroleum and Minerals, Dhahran, Saudi Arabia (1981).

[3] N. R. Bhalerao and J. Gurland, 'Statistical Inference Concerning Some Compound and Generalized Discrete Distributions', *Technical Summary Report, 1611*, University of Wisconsin, U.S.A. (1976), pp. 1–72.

[4] A. Erdelyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi, *Higher Transcendental Functions*, Vol. 1. Bateman Manuscript Project, New York: McGraw-Hill, 1953, p. 283.

[5] A. V. Fiacco and G. P. McCormick, *Non-linear Programming Sequential Unconstrained Minimization Techniques*. New York: Wiley 1968.

[6] R. Fletcher and M. J. D. Powell, 'A Rapidly

Convergent Descent Method for Minimization', *Computer Journal*, **6**(1963), pp. 163–168.
[7] P. Hinz and J. Gurland, 'Simplified Techniques for Estimating Parameters for Some Generalized Poisson Distributions', *Biometrika*, **54** (1967), pp. 555–566.
[8] L. Katz, 'Unified Treatment of a Broad Class of Discrete Probability Distributions', in *Classical and Contagious Discrete Distribution*, ed. G. P. Patil, New York: Pergamon Press, 1953, pp. 175–182.
[9] A. W. Kemp, 'A Wide Class of Discrete Distributions and the Associated Differential Equations', *Sankhya*, **30** (1968), pp. 401–410.
[10] W. O. Williford and A. L. Price, 'Estimation in Modified and Compound Modified Discrete Distributions', *Biometrics*, **34** (1978), pp. 631–643.

Paper Received 16 November 1981; Revised 30 May 1982.