POLYA-AEPPLI-LINDLEY DISTRIBUTION

MEGLENA D. LAZAROVA, LEDA D. MINKOVA

ABSTRACT: *In this study is introduced the Polya-Aeppli-Lindley distribution as a mixed Polya-Aeppli distribution with Lindley mixing distribution. The probability mass function, recursion formulas and some properties are derived.*

KEYWORDS: *Polya-Aeppli distribution, Inflated – parameter Negative binomial distribution, mixed distribution, Polya-Aeppli-Lindley distribution*

1. Introduction

In 1970 Sankaran introduced a mixed Poisson distribution with Lindley mixing distribution and called it a Poisson-Lindley distribution, see [8]. Minkova (2002) in [6] had defined a mixed Polya-Aeppli distribution with gamma mixing distribution and called it an Inflated-parameter negative binomial (INB) distribution.

We say that the random variable N has a Polya-Aeppli distribution with parameters $\lambda > 0$ and $\rho \in [0,1)$ if the probability mass function is given by:

$$(1) \quad P(N = k \mid \lambda) = \begin{cases} e^{-\lambda}, & k = 0 \\ e^{-\lambda} \sum_{i=1}^{k} \binom{k-1}{i-1} \frac{[\lambda(1-\rho)]^i}{i!} \rho^{k-i}, & k = 1, 2, \ldots \end{cases}$$

see [4]. We use the notation $N \in PA(\lambda, \rho)$.

The mean and the variance of the random variable N are given by: $E(N) = \frac{\lambda}{1-\rho}$ and $Var(N) = \frac{\lambda(1+\rho)}{(1-\rho)^2}$.

The probability generating function (PGF) of the Polya-Aeppli distribution with given parameter λ is:

$$(2) \quad \psi(s \mid \lambda) = e^{-\lambda(1-\psi_1(s))},$$

where $\psi_1(s) = \frac{(1-\rho)s}{1-\rho s}$ is the PGF of the geometric distribution.

In the case of $\rho = 0$, the PMF presented above coincides with the PMF of the classical Poisson distribution.

The random variable Λ with density function $g(\lambda) = \frac{\beta^2}{1+\beta}.e^{-\beta\lambda}(1+\lambda), \ \lambda > 0$ is said to be Lindley distributed with parameter $\beta > 0$. The Lindley distribution was introduced by Lindley in 1958, see [5] and it is a mixture of Gamma (1, β) and Gamma (2, β) distributions i.e.

(3) $$g(\lambda) = \frac{\beta}{1+\beta}\beta e^{-\beta\lambda} + \frac{1}{1+\beta}\beta^2\lambda e^{-\beta\lambda}, \quad \lambda > 0$$

Sankaran (1970), see [8] had introduced the Poisson-Lindley distribution as a mixture of Poisson distribution with Lindley mixing distribution with PMF:

(4) $$P(N=k) = \frac{\beta^2}{(1+\beta)^{k+3}}.[2+\beta+k], \quad k = 0,1,\dots$$

In this paper we introduce the Polya-Aeppli-Lindley distribution.

2. Polya-Aeppli-Lindley distribution

It is known that the properties of the Polya-Aeppli distribution are very close to these of the Poisson distribution, see [1]. On the other hand for $r \geq 1$ the mixture of the Polya-Aeppli distribution with $Gamma\,(r,\beta)$ mixing distribution has the following PMF:

(5) $$P(N=k) = \begin{cases} \left(\dfrac{\beta}{1+\beta}\right)^r, & k=0 \\ \left(\dfrac{\beta}{1+\beta}\right)^r \displaystyle\sum_{i=1}^{k} \binom{k-1}{i-1}\binom{r+i-1}{i}\left(\dfrac{1-\rho}{1+\beta}\right)^i \rho^{k-i}, & k=1,2,\dots \end{cases}$$

The distribution in (5) is called Inflated-parameter Negative Binomial distribution with parameters r,β and ρ i.e. $INB(r,\beta,\rho)$, see [6].

Let us suppose that the parameter λ in the Polya-Aeppli distribution (1) has a Lindley distribution in the form (3). Then from the PGF given in (2) we obtain that the unconditional PGF of N has the form:

$$\psi(\mathrm{s}) = \frac{\beta}{1+\beta}.\frac{\beta}{\beta+(1-\psi_1(s))} + \frac{1}{1+\beta}\left[\frac{\beta}{\beta+(1-\psi_1(s))}\right]^2$$

Denote by $\theta = \dfrac{\beta}{1+\beta}$ the new parameter of the distribution. This parametrization was used for the Poisson-Lindley distribution.

Then the PGF has the form:

(6) $$\psi(s) = \frac{\theta}{1-(1-\theta)\psi_1(s)}.\left[\theta + (1-\theta).\frac{\theta}{1-(1-\theta)\psi_1(s)}\right] = \frac{\theta(1-\rho s)}{1-(1-\theta(1-\rho))s}.\left[\theta+(1-\theta).\frac{\theta(1-\rho s)}{1-(1-\theta(1-\rho))s}\right]$$

Definition: The distribution of the random variable N with PGF given in (6) is called Polya-Aeppli-Lindley distribution with parameters θ and ρ. We use the notation $N \sim PAL(\theta,\rho)$.

From the PGF in (6) it follows that the random variable N can be represented as a sum of two independent variables $N = N_1 + N_2$, where N_1 has a compound geometric distribution and

N_2 has a compound modified geometric distribution with the same geometric compounding distribution.

The corresponding PGFs are given by:

$$\psi_{N_1}(s) = \frac{\theta}{1-(1-\theta)\psi_1(\mathrm{s})}$$

and

$$\psi_{N_2}(s) = \theta + (1-\theta)\frac{\theta}{1-(1-\theta)\psi_1(s)}$$

2.1 Probability mass function

Using the form (3) of the mixing distribution we obtain the following theorem:

Theorem1: The PMF of the Polya-Aeppli-Lindley distribution is given by:

$$(7)\quad P(N=k) = \begin{cases} \dfrac{\beta^2}{(1+\beta)^3}.(2+\beta), & k=0 \\ \dfrac{(1-\rho)\beta^2}{(1+\beta)^4}.(3+\beta), & k=1 \\ \dfrac{(1-\rho)\beta^2(1+\beta\rho)^{k-2}}{(1+\beta)^{k+3}}[(1+\beta\rho)(3+\beta)+(\mathrm{k}-1)(1-\rho)], & k=2,3,... \end{cases}$$

Proof: The PMF in (7) is obtained from the PMF of Polya-Aeppli distribution given in (1), where the parameter λ has a Lindley distribution, given in (3).

The graphics below show the PMF of the Polya-Aeppli-Lindley distribution obtained by giving the different values of the parameter ρ. The first parameter β is fixed and the second one is increasing.

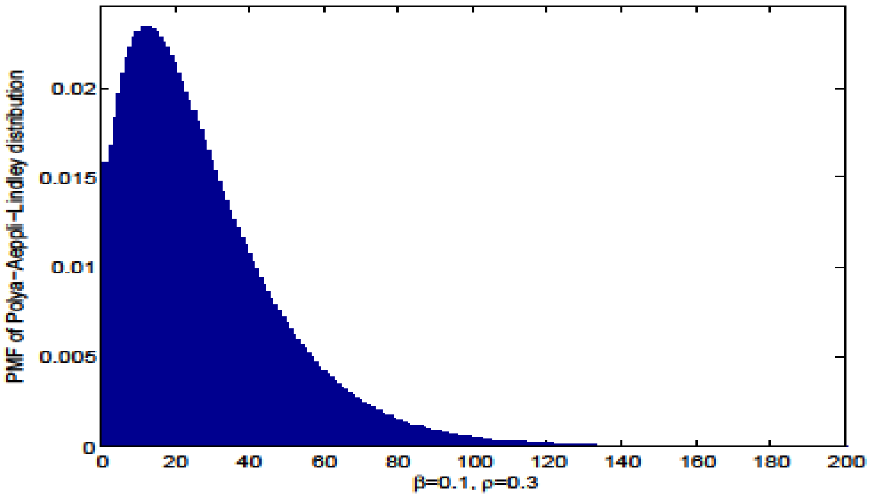

Figure1: Graphic of PMF of Polya-Aeppli-Lindley distribution with given parameters $\beta = 0,1$ and $\rho = 0,3$

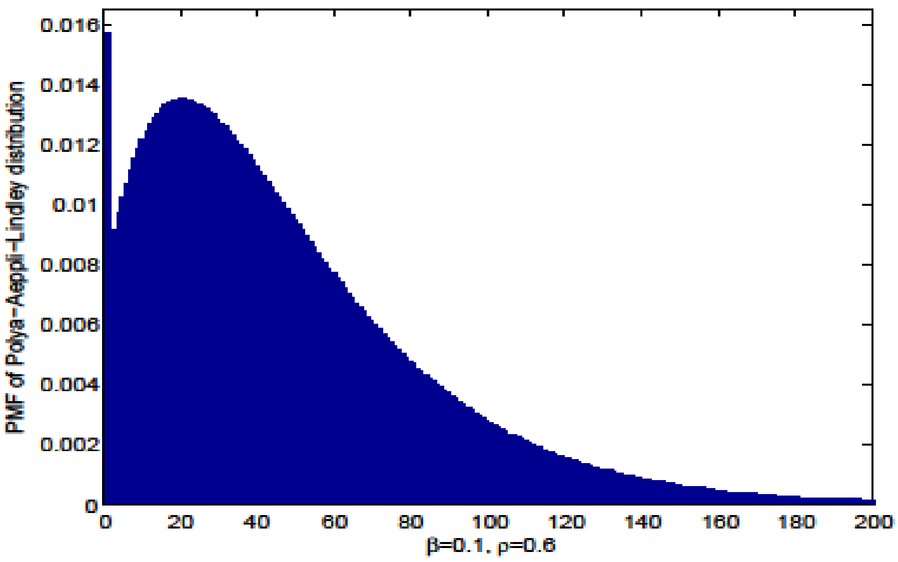

Figure2: Graphic of PMF of Polya-Aeppli-Lindley distribution with given parameters $\beta = 0{,}1$ and $\rho = 0{,}6$

Remark: In the case of $\rho = 0$ the PMF in (7) coincides with the PMF of the Poisson-Lindley distribution in (4), see Figure3.

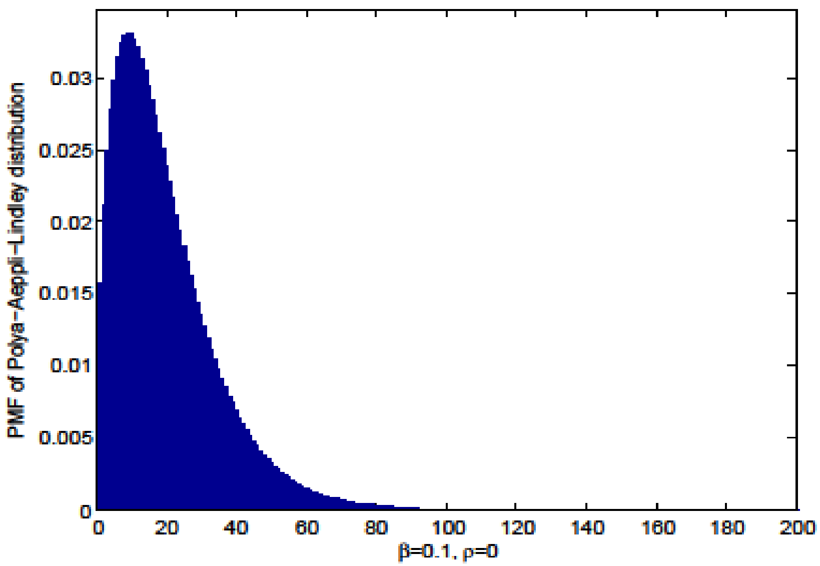

Figure3: Graphic of PMF of Polya-Aeppli-Lindley distribution with given parameters $\beta = 0{,}1$ and $\rho = 0$

On the graphics above it is seen that for fixed parameter $\beta = 0{,}1$ and for increasing values of the parameter ρ the Polya-Aeppli-Lindley distribution has a heavy tail.

From the PGFs of N_1 and N_2 we obtain the following PMFs:

$$P(N_1 = i) = \begin{cases} \theta, & i = 0 \\ \theta(1-\theta)(1-\rho)[1-\theta(1-\rho)]^{i-1}, & i = 1,2,\ldots \end{cases}$$

and

$$P(N_2 = i) = \begin{cases} \theta(2-\theta), & i = 0 \\ \theta(1-\theta)^2(1-\rho)[1-\theta(1-\rho)]^{i-1}, & i = 1,2,\ldots \end{cases}$$

Taking into account that $N = N_1 + N_2$ we obtain the following PMF of N in terms of θ:

$$(8)\quad P(N = k) = \begin{cases} \theta^2(2-\theta), & k = 0 \\ \theta^2(1-\theta)(1-\rho)(3-2\theta), & k = 1 \\ \theta^2(1-\theta)(1-\rho)[1-\theta(1-\rho)]^{k-2}[(1-\theta(1-\rho))(3-2\theta+(k-1)(1-\theta)^2(1-\rho)], & k = 2,3,\ldots \end{cases}$$

Denoting by $p_k = P(N = k),\ k = 0,1,\ldots$ the PMF of the random variable N we obtain the following theorem:

Theorem2: The PMF of Polya-Aeppli-Lindley distribution satisfies the following recursion formulas:

$$p_0 = \theta^2(2-\theta), \qquad k = 0$$

$$(2-\theta)\,\mathrm{p}_1 = (1-\rho).(1-\theta).(3-2\theta)\,\mathrm{p}_0, \qquad k = 1$$

$$(2-\theta)\,\mathrm{p}_2 = (2-\theta).[1-\theta(1-\rho)]\,\mathrm{p}_1 + (1-\rho)^2(1-\theta)^3\,p_0 \qquad k = 2$$

and for k= 3,4,…

$$(2-\theta).(k+1).p_{k+1} = [(1-\theta)(1+\rho)+2\rho].k.p_k - \rho(1-\theta+\rho).(k-1).\mathrm{p}_{k-1} +$$
$$+(1-\rho)(1-\theta)(3-2\theta).\mathrm{p}_k + 2(1-\rho)^2(1-\theta)^3.\sum_{j=0}^{k-1}[1-\theta(1-\rho)]^{k-j-1}.p_j$$

Proof: The initial value p_0 is obtained upon substituting $s = 0$ in the PGF, given in formula (6). Differentiation in (6) leads to:

$$(9)\quad (1-\rho s)[1-(1-\theta(1-\rho))s][2-\theta-(1+\rho)\mathrm{s}].\psi'(\mathrm{s}) = (1-\rho)(1-\theta)[3-2\theta-(1-(2-\theta)\rho)\mathrm{s}].\psi(\mathrm{s}),$$

where $\psi(s) = \sum_{k=0}^{\infty} p_k s^k$ and $\psi'(\mathrm{s}) = \sum_{k=0}^{\infty}(k+1)p_{k+1}s^k$.

The recursions are obtained by equating the coefficients of s^k on both sides for fixed k=0,1,2,…

2.2 Moments

The mean and the variance of the Polya-Aeppli-Lindley distribution in the terms of θ are given by:

$$E(N)=\frac{(1-\theta).(2-\theta)}{\theta(1-\rho)} \quad \text{and} \quad Var(N)=\frac{(1-\theta).\left[2(1+\theta\rho)-\theta^2(\theta+\rho)\right]}{\theta^2(1-\rho)^2}$$

For the Fisher index of dispersion we obtain:

$$FI(N)=\frac{Var(N)}{E(N)}>\frac{1+\rho}{1-\rho}$$

It is known that the Fisher index of Polya-Aeppli distribution is equal to $\frac{1+\rho}{1-\rho}$, see [6]. It follows that the Polya-Aeppli-Lindley distribution is over-dispersed related to Polya-Aeppli distribution.

REFERENCES

1. Chukova, S. and Minkova, L.D. Characterization of the Polya-Aeppli Process, Stochastic Analysis and Applications, 31, 590-599.2, 2013
2. Ghitany, M.E. and S.A. Al-Awadhi. A Unified Approach to Some Mixed Poisson Distributions, Tamsui Oxford Journal of Mathematical Sciences, 17, 147-161, 2001
3. Ghitany, M.E. and Al-Mutairi D. Size-based Poisson-Lindley distribution and its application , Metron-International Journal of Statistics, LXVI, 299-311, 2008
4. Johnson, N.L., Kotz, S. and Kemp, A.W. Univariate Discrete Distributions, Wiley Series in Probability and Mathematical Statistics, third edition 1, 2005
5. Lindley, D.V. Fiducial distributions and Bayes’ theorem, Journal of the Royal Statistical Society, 20, 102-107.1, 1958
6. Minkova, L.D. A Generalization of the Classical Discrete Distributions, Commun.Statist –Theory and Methods, 31, 871-888. 1, 2, 2.3, 2002
7. Nadarajah, S. Some gamma distributions, Statistics, 42, 77-94, 2008
8. Sankaran, M. The discrete Poisson-Lindley distribution, Biometrics, 26, 145-149. 1,1, 1970