Non-central Pólya-Aeppli distribution

Meglena Lazarova, and Leda Minkova

Citation: AIP Conference Proceedings **2048**, 020022 (2018); doi: 10.1063/1.5082040
View online: https://doi.org/10.1063/1.5082040
View Table of Contents: http://aip.scitation.org/toc/apc/2048/1
Published by the American Institute of Physics

Non-central Pólya-Aeppli distribution

Meglena Lazarova[1,a)] and Leda Minkova[2,b)]

[1]*Faculty of Applied Mathematics and Informatics,*
Technical University of Sofia, 1000 Sofia, Bulgaria
[2]*Faculty of Mathematics and Informatics,*
Sofia University "Sv. Kliment Ohridski", 1164 Sofia, Bulgaria

[a)]meglena.laz@tu-sofia.bg
[b)]leda@fmi.uni-sofia.bg

Abstract. In this paper the Noncentral Pólya-Aeppli distribution is introduced. It is a sum of two independent random variables, one that is a Poisson and another one, a Pólya-Aeppli distributed. The probability mass function, moments, recursion formulas and some properties of the defined distribution are derived.

INTRODUCTION

The Non-central negative binomial distribution arises as a model in photon and neural counting, birth and death processes and mixture models, see Ong and Lee [1]. Ong and Lee gave a formulation of the Non-central negative binomial distribution as a Poisson and negative binomial mixture. In fact it is a convolution of independent negative binomial and Pólya-Aeppli distribution. In 1983 Gurland et. al, see [2] considered the Non-central negative binomial distribution and referred it to as a Laguerre series distribution. Lately it had been developed in the work of Ong et. al [3] and Ong and Shimizu [4]. The probability mass function (PMF) and the probability generating function (PMF) of the Noncentral negative binomial distribution are given by

$$P(N = i) = e^{-\lambda p} p^i q^\nu L_i^{\nu-1}(-\lambda q) \tag{1}$$

and

$$\psi_N(s) = \left(\frac{q}{1-ps}\right)^\nu e^{-\left[1-\frac{q}{1-ps}\right]}, \tag{2}$$

where $\nu > 0$ and $\lambda > 0$ are parameters, $0 < p = 1 - q < 1$ and $L_i^\alpha(x)$ are the Laguerre polinomials orthogonal over $(0, \infty)$ with respect to $x^{\alpha-1}e^{-x}$.

The Pólya-Aeppli distribution was derived by Anscombe in 1950, see Anscombe [5] from a model of randomly distributed colonies. In 1953 it was also studied by Evans, see Evans [6]. Anscombe states that in 1930 the distribution was given by A. Aeppli in a thesis and in 1930 developed by G. Pólya. For this reason he called it a Pólya-Aeppli distribution. It is a compound Poisson with geometric compounding distribution. The probability mass function (PMF) and the probability generating function (PGF) of the compounding distribution are given by

$$P(X = i) = (1-\rho)\rho^{i-1}, \qquad i = 1, 2, \ldots \tag{3}$$

and

$$\psi_1(s) = Es^X = \frac{(1-\rho)s}{1-\rho s} \tag{4}$$

Another generalization of the Non-central negative binomial distribution is the I-Delaporte distribution, defined in [7]. It is a convolution of I-Negative binomial distribution (*Inflated - parameter negative binomial distribution*), introduced in [8] as a compound negative binomial distribution with geometric compounding distribution and a Pólya-Aeppli distribution.

Proceedings of the 44th International Conference on Applications of Mathematics in Engineering and Economics
AIP Conf. Proc. 2048, 020022-1–020022-5; https://doi.org/10.1063/1.5082040
Published by AIP Publishing. 978-0-7354-1774-8/$30.00

In the present paper we introduce and analyze a Non-central Pólya-Aeppli distribution as a sum of independent Poisson and Pólya-Aeppli distribution. In [9] we introduced the Non-central Pólya-Aeppli process and applied it as a counting process in risk model. The corresponding distribution is analyzed in this paper. The paper is organized as follows. In the next sections we consider the probability generating function, the probability mass function and the moments of Non-central Pólya-Aeppli distribution. For specific values of the parameters we give graphics and make some conclusions.

PROBABILITY GENERATING FUNCTION

The Non-central Pólya-Aeppli distribution is defined in [9] as a counting distribution in risk models. It is a sum of independent Poisson and Pólya-Aeppli distribution. Suppose that the first random variable N_1 with a given parameter $\lambda_1 > 0$ has a Poisson distribution i.e

$$P(N_1 = i) = \frac{(\lambda_1)^i}{i!} e^{-\lambda_1}, \ i = 0, 1, \dots \tag{5}$$

We use the notation $N_1 \sim P_o(\lambda_1)$. The PGF of the Poisson distribution with given parameter λ_1 is given by

$$\psi_N(s) = e^{-\lambda_1(1-s)} \tag{6}$$

The mean and the variance of the Poisson distribution are given by

$$E(N_1) = Var(N_1) = \lambda_1$$

The related Fisher index of dispersion is

$$FI(N) = \frac{Var(N)}{E(N)} = 1$$

The second random variable N_2 with given parameters $\lambda_2 > 0$ and $\rho \in [0, 1)$ has a Pólya-Aeppli distribution with PMF

$$P(N_2 = i) = \begin{cases} e^{-\lambda_2}, & i = 0 \\ e^{-\lambda_2} \sum_{j=1}^{i} \binom{i-1}{j-1} \frac{[\lambda_2(1-\rho)]^j}{j!} \rho^{i-j}, & i = 1, 2, \dots \end{cases} \tag{7}$$

We use the notation $N_2 \sim PA(\lambda_2, \rho)$. The distribution given in formula (7) is also called an Inflated-parameter Poisson distribution, see [8]. The mean and the variance of the Pólya-Aeppli distribution are given by

$$E(N_2) = \frac{\lambda_2}{1-\rho}$$

and

$$Var(N_2) = \frac{\lambda_2(1+\rho)}{(1-\rho)^2}.$$

The related Fisher index of dispersion is

$$FI(N_2) = \frac{1+\rho}{1-\rho}$$

i.e for $\rho \neq 0$ the Pólya-Aeppli distribution is over-dispersed related to the Poisson distribution.

The PGF of the Pólya-Aeppli distribution with parameters $\lambda_2 > 0$ and $\rho \in [0, 1)$ is given by

$$\psi_N(s \mid \lambda_2) = e^{-\lambda_2[1-\psi_1(s)]}, \tag{8}$$

where $\psi_1(s)$ is the PGF of the shifted geometric distribution with success probability $1-\rho$, $Ge_1(1-\rho)$, given in (4).

Considering that $N = N_1 + N_2$, where N_1 and N_2 are independent random variables we obtain that the PGF of the Non-central distributed random variable N is given by

$$\psi_N(s) = e^{-\lambda_1(1-s)} e^{-\lambda_2[1-\psi_1(s)]}, \tag{9}$$

where $\psi_1(s)$ is given in (4). This leads to the next definition.

Definition 1. *The random variable N with PGF given in (9) is referred to a Non-central Pólya-Aeppli distribution. We use the notation* $N \sim NPAD(\lambda_1, \lambda_2, \rho)$.

PROBABILITY MASS FUNCTION

The paper of Ong and Lee, see [1] motivated us to call the new distribution Non-central Pólya-Aeppli distribution. It is well known that the Poisson distribution is a limiting case of the Negative binomial distribution. If we take $\nu(1-q) \to \lambda > 0$ in the first term of (2) then we obtain the PGF, given in (9). Thus, if the both terms in the equation (2) have different parameters, then the Non-central Pólya-Aeppli distribution could be a limiting case of the Non-central Negative binomial distribution.

In the next proposition we obtain recursion formulas for the PMF of the defined distribution.

Proposition 1. *The PMF of the Non-central Pólya-Aeppli distribution satisfies the following recursions*

$$\begin{aligned} p_1 &= [\lambda_1 + \lambda_2(1-\rho)]p_0 \\ p_i &= \left[2\rho + \tfrac{[\lambda_1+\lambda_2(1-\rho)-2\rho]}{i}\right] p_{i-1} - \rho\left[\rho + 2\tfrac{\lambda_1-\rho}{i}\right] p_{i-2} + \tfrac{\lambda_1\rho^2}{i} p_{i-3}, \quad i = 2, 3, \ldots, \end{aligned} \tag{10}$$

with $p_0 = e^{-(\lambda_1+\lambda_2)}$ and $p_{-1} = 0$.

Proof. Upon substituting $s = 0$ in the PGF given in (9), we obtain the initial value p_0. The differentiation of the PGF in (9) leads to

$$(1-\rho s)^2 \psi_N'(s) = [\lambda_1(1-\rho s)^2 + \lambda_2(1-\rho)]\psi_N(s), \tag{11}$$

where $\psi_N(s) = \sum_{i=0}^{\infty} p_i s^i$ and $\psi_N'(s) = \sum_{i=0}^{\infty}(i+1)p_{i+1}s^i$. The recursions are obtained by equating the coefficients of s^i for fixed values of $i = 0, 1, 2, \ldots$ on the both sides of the equality (11).

□

An alternative recursion formula is given in the next corollary.

Corollary 1. *The PMF of the Non-central Pólya-Aeppli distribution satisfies the following recursions*

$$\begin{aligned} p_1 &= [\lambda_1 + \lambda_2(1-\rho)]p_0 \\ p_i &= \tfrac{[\lambda_1+\lambda_2(1-\rho)]}{i} p_{i-1} + \lambda_2(1-\rho)\textstyle\sum_{j=0}^{i-2}\left(1 - \tfrac{j}{i}\right)\rho^{i-j-1} p_j, \quad i = 2, 3, \ldots, \end{aligned} \tag{12}$$

with $p_0 = e^{-(\lambda_1+\lambda_2)}$.

Proof. The required recursions are obtained from the first derivative of PGF given in

$$\psi_N'(s) = \left[\lambda_1 + \frac{\lambda_2(1-\rho)}{(1-\rho s)^2}\right]\psi_N(s)$$

□

Recursively from (10) we obtain the PMF of the Non-central Pólya-Aeppli distribution in the next theorem.

Theorem 1. *The PMF of a Non-central Pólya-Aeppli distributed random variable is given by*

$$P(N = i) = \begin{cases} e^{-(\lambda_1+\lambda_2)}, & i = 0 \\ e^{-(\lambda_1+\lambda_2)}\left[\frac{(\lambda_1)^i}{i!} + \sum_{j=1}^{i}\frac{(\lambda_1 t)^{i-j}}{(i-j)!}\sum_{k=1}^{j}\binom{j-1}{k-1}\frac{[\lambda_2(1-\rho)t]^k}{k!}\rho^{j-k}\right], & i = 1, 2, \ldots \end{cases} \tag{13}$$

For specific values of the parameters ρ, λ_1 and λ_2 we construct some useful graphics for the probability mass function of the Non-central Pólya-Aeppli distribution. Initially we take the parameter ρ to be a constant and we change the values of the parameters λ_1 and λ_2. We are interested in the sensitivity of the distribution on the parameters λ_1 and λ_2.

In Figure 1 we introduce three graphics for the probability mass function of the Non-central Pólya-Aeppli distribution with fixed parameter $\rho = 0.5$. We take $\lambda_1 < \lambda_2$. It is seen that on the left graphic we have more events in the beginning than the events in the other two graphics. We can conclude that in this case the larger parameter λ_2 leads to a heavier tail of the distribution.

In Figure 2 we fix the parameters λ_1, λ_2 and we change the values of the parameter ρ. On the left graphic we take $\rho = 0$ and thus we obtain the PMF of the sum of two independent Poisson distributions. On the right graphic we take the value $\rho = 0.9$. It is seen that for a larger value of the parameter ρ the distribution has a longer right tail.

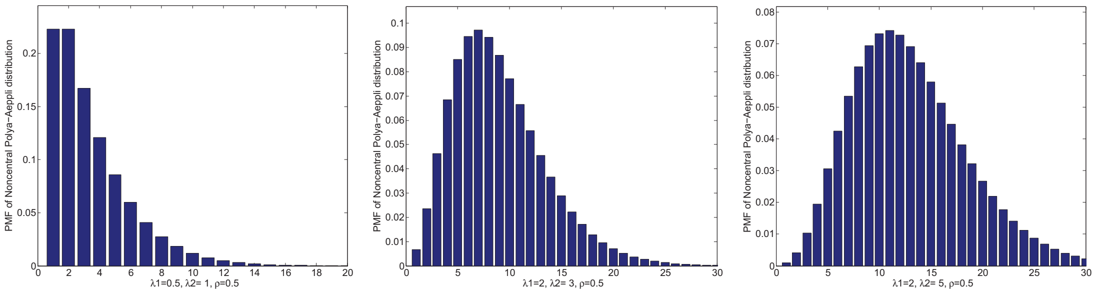

FIGURE 1. Probability mass function of NPAD with fixed parameter $\rho = 0.5$

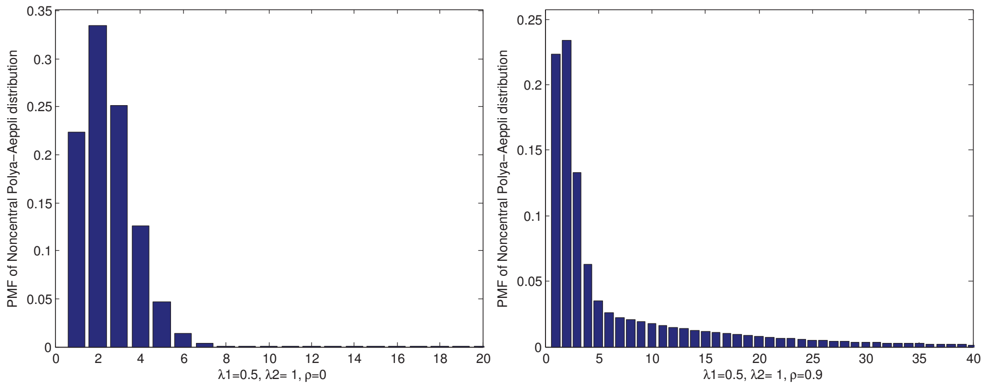

FIGURE 2. Probability mass function of NPAD with fixed parameters λ_1 and λ_2

MOMENTS

The mean and the variance of the Noncentral Pólya-Aeppli distribution are given by

$$E(N) = \left(\lambda_1 + \frac{\lambda_2}{1-\rho}\right) \quad \text{and} \quad Var(N) = \left[\lambda_1 + \lambda_2 \frac{1+\rho}{(1-\rho)^2}\right]$$

For the Fisher index we obtain

$$FI(N) = \frac{\lambda_1(1-\rho)^2 + \lambda_2(1+\rho)}{(1-\rho)[\lambda_1(1-\rho) + \lambda_2]}$$

It is easy to check that

$$FI(N) = 1 + \frac{2\lambda_2\rho}{(1-\rho)[\lambda_1(1-\rho) + \lambda_2]}$$

i.e. NPAD is over-dispersed related to Poisson distribution and

$$FI(N) = \frac{1+\rho}{1-\rho} - \frac{2\lambda_1\rho}{\lambda_1(1-\rho) + \lambda_2} < \frac{1+\rho}{1-\rho},$$

i.e. NPAD is under-dispersed related to Pólya-Aeppli distribution

CONCLUDING REMARKS

In the present paper we have analyzed the Noncentral Pólya-Aeppli distribution with its probability generating function, probability mass function, recursion formulas and moments. We also gave the connection to some well-known distributions like Poisson distribution, Pólya-Aeppli distribution and Noncentral negative binomial distribution.

ACKNOWLEDGMENTS

This research is partially supported by project ”Stochastic project and simulation models in the field of medicine, social sciences and dynamic systems” founded by the National Science Fund of Ministry of Education and Science of Bulgaria (Contract No. DN 12/11/20 Dec. 2017).

REFERENCES

[1] S. H. Ong and P. Lee, The Non-central Negative Binomial Distribution. Biometric J. **21**, 611–628 (1979).

[2] J. Gurland, E. Chen, and F. Hernandez, A new discrete distribution involving Laguerre polinomials. Commun. Statist. Theory and Methods **12**, 1987–2004 (1983).

[3] S. H. Ong, S. Chakraborty, T. Imoto, and K. Shimizu, Generalizations of Non-central Negative Binomial, Charlier Series Distributions and their extensions by Lagrange expansion, Commun. Statist. - Theory and Methods **41**, 571–590 (2012).

[4] S. H. Ong and K. Shimizu, A Discrete Distribution Arising as a Solution of a Linear Difference Equation: Extension of the Non-central Negative Binomial Distribution,Commun. Statist. - Theory and Methods **38**, 927–938 (2009).

[5] F. Anscombe, Sampling theory of the negative binomial and logarithmic series distributions. Biometrika **37**, 358–382 (1950).

[6] D. A. Evans, Experimental evidence concerning contagious distributions in ecology. Biometrika **40**, 186–211 (1953).

[7] M. D. Lazarova and L. D. Minkova, On the I-Delaporte distribution. National Conference An application of the Mathematics, Statistics and Information technology for modeling of economic and business processes, UNSS, Sofia, 16–22 (2016).

[8] L. D. Minkova, A family of compound discrete distributions. Compt. Randue Bulg. Acad. Sci. **54**, 11–14 (2001).

[9] M. D. Lazarova and L. D. Minkova, Noncentral Pólya-Aeppli process and Applications. Doctoral Conference in Mathematics and Informatics (MIDOC 2015) Proceedings, Sofia, 80–87 (2016).