Bivariate Noncentral Pólya-Aeppli Distribution. An Application to Insurance Risk Theory

Meglena Lazarova

Faculty of Applied Mathematics and Informatics,
Technical University of Sofia, Bulgaria

in collaboration with Leda Minkova and Michail Todorov

Workshop on Discrete Distributions
in memory of Adrienne Freda Kemp
Harokopio University of Athens, Greece, 13 April 2024

Outline

- 1. Introduction
- 2. Univariate Noncentral Pólya-Aeppli distribution
  2.1. Probability generating function
  2.2.Probability mass function and moments
- 3. Bivariate Noncentral Pólya-Aeppli distribution
  3.1. Joint probability generating function
  3.2. Joint probability mass function and moments

In this talk a convolution of Poisson distribution and Pólya-Aeppli distribution is analyzed. The resulting distribution is called a Noncentral Pólya-Aeppli distribution.

Suppose that the nonnegative random variable Z_3 with a given parameter λ_3 has a Poisson distribution, i.e.,

$$P(Z_3 = i \mid \lambda_3) = \frac{\lambda_3^i}{i!} e^{-\lambda_3}, \ \ i = 0, 1, \dots . \qquad (1)$$

where $\lambda_3 > 0$ and we use the notation $Z_3 \sim Po(\lambda_3)$. It is known that $\mathrm{E}(Z_3) = \mathrm{Var}(Z_3) = \lambda_3$.

The mean and the variance of Z_1 are given by

$$\mathrm{E}(Z_1) = \frac{\lambda_1}{1-\rho}$$

and

$$\mathrm{Var}(Z_1) = \frac{\lambda_1(1+\rho)}{(1-\rho)^2}$$

The Fisher index is

$$FI(Z_1) = \frac{1+\rho}{1-\rho}$$

The Pólya-Aeppli distribution is a compound Poisson with geometric compounding distribution and it is analyzed in Minkova (2002) [4].

2.1. Probability generating function

The Noncentral Pólya-Aeppli distribution is a sum of independent Pólya-Aeppli distribution and Poisson distribution, i.e., $M = Z_1 + Z_3$, where Z_1 and Z_3 are independent random variables. Then the probability generating function (PGF) of the random variable M is given by

$$\Psi_N(s) = e^{-\lambda_1(1-\psi_1(s))} e^{-\lambda_3(1-s)}, \tag{2}$$

where

$$\psi_1(s) = \frac{(1-\rho)s}{1-\rho s} \tag{3}$$

is the PGF of the geometric distribution, $Ge_1(1-\rho)$.

Definition

The random variable M, defined by the PGF (2), has a Noncentral Pólya-Aeppli distribution (NPAD).

In 1979 Ong and Lee defined the Noncentral negative binomial distribution, [8]. It is the distribution of a sum of two independent variables. The first variable has a Negative binomial distribution and the second one is Pólya-Aeppli distributed. This was our motivation to name the resulting distribution Non-central Pólya-Aeppli distribution.

The probability mass function (PMF) of the Noncentral Pólya-Aeppli distribution is given by

$$P(M=i)$$

$$= \begin{cases} e^{-(\lambda_1+\lambda_3)}, \quad i=0 \\ e^{-(\lambda_1+\lambda_3)} \left[\sum_{j=1}^{i} \frac{(\lambda_3)^{i-j}}{(i-j)!} \sum_{k=1}^{j} \binom{j-1}{k-1} \frac{[\lambda_1(1-\rho)]^k}{k!} \rho^{j-k} + \frac{\lambda_3^i}{i!}\right], \\ \quad i=1,2,\dots, \end{cases} \tag{4}$$

Proposition

Denote by $p_i = P(M = i),\ i = 0, 1, \ldots.$
Then the PMF of the Univariate Noncentral Pólya-Aeppli distribution satisfies the following recursions:
$p_1 = [\lambda_1(1-\rho) + \lambda_3]p_0,$ *where* $p_0 = e^{-(\lambda_1+\lambda_3)}$ *and for* $i = 1, 2, \ldots$

$$p_{i+1} = \frac{[\lambda_1(1-\rho)+\lambda_3]}{i+1} p_i + \lambda_1(1-\rho)\sum_{j=0}^{i-1}\left[1 - \frac{j}{i+1}\right]\rho^{i-j} p_j$$

Moments

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

$$\mathrm{E}(M) = \left(\frac{\lambda_1}{1-\rho} + \lambda_3\right) \quad \text{and} \quad \mathrm{Var}(M) = \left(\lambda_1 \frac{1+\rho}{(1-\rho)^2} + \lambda_3\right)$$

For the Fisher index we obtain:

$$FI(M) = \frac{\lambda_3(1-\rho)^2 + \lambda_1(1+\rho)}{(1-\rho)[\lambda_3(1-\rho) + \lambda_1]}$$

It is easy to check that:

$$FI(M) = 1 + \frac{2\lambda_1\rho + \lambda_3(1-\rho)}{(1-\rho)[\lambda_3(1-\rho) + \lambda_1]},$$

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

$$FI(M) = \frac{1+\rho}{1-\rho} - \frac{2\lambda_3\rho}{\lambda_3(1-\rho) + \lambda_1} < \frac{1+\rho}{1-\rho},$$

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

3. Bivariate Noncentral Pólya-Aeppli distribution

Let Z_1, Z_2 and Z_3 are independent variables. Suppose that $Z_i,\ i = 1, 2$ are Pólya-Aeppli distributed with parameters λ_i and ρ_i,, i.e, $Z_i \sim PA(\lambda_i, \rho_i)$. The random variable Z_3 has a Poisson distribution with parameters λ_3, i.e $Z_3 \sim Po(\lambda_3)$.
Let us also define the random variables:

$$M = Z_1 + Z_3 \ \text{ and } \ N = Z_2 + Z_3.$$

It is clear that:

$$M \sim NPA(\lambda_1 + \lambda_3, \rho_1) \ \text{ and } \ N \sim NPA(\lambda_2 + \lambda_3, \rho_2).$$

The Joint PGF of (M, N) is given by:

$$\psi(s_1, s_2) = e^{-\lambda_1(1-\frac{(1-\rho_1)s_1}{1-\rho_1 s_1})-\lambda_2(1-\frac{(1-\rho_2)s_2}{1-\rho_2 s_2}-\lambda_3(1-s_1 s_2)} \quad (5)$$

where

$$\psi_1(s_1) = \frac{(1-\rho_1)s_1}{1-\rho_1 s_1} \quad \text{and} \quad \psi_1(s_2) = \frac{(1-\rho_2)s_2}{1-\rho_2 s_2}$$

are the PGF-s of the compounding distribution given in (3).

Definition

The probability of (M, N) corresponding to (5) is referred to as a Bivariate Noncentral Pólya-Aeppli distribution, i.e., $BivNPA(\lambda_1, \lambda_2, \lambda_3, \rho_1, \rho_2)$.

Denote by $f(i,j) = P(M = i, N = j), \ i,j = 0,1,\ldots$ the Joint PMF of (M, N).
Then the following proposition gives an extension of the Panjer recursion formulas; see [9].

Proposition

The PMF of the BivPA$(\lambda_1, \lambda_2, \lambda_3, \rho_1, \rho_2)$ distribution satisfies the following recursions:

$$f(i+1,j) = \left[2\rho_1 + \tfrac{(1-\rho_1)\lambda_1 - 2\rho_1}{i+1}\right] f(i,j) - \rho_1^2(1 - \tfrac{2}{i+1}) f(i-1,j)$$

$$+ \tfrac{\lambda_3}{i+1}\left[f(i,j-1) - 2\rho_1 f(i-1,j-1) + \rho_1^2 f(i-2,j-1)\right],$$

$$f(i,j+1) = \left[2\rho_2 + \tfrac{(1-\rho_2)\lambda_2 - 2\rho_2}{j+1}\right] f(i,j) - \rho_2^2(1 - \tfrac{2}{j+1}) f(i,j-1)$$

$$+ \tfrac{\lambda_3}{j+1}\left[f(i-1,j) - 2\rho_2 f(i-1,j-1) + \rho_2^2 f(i-1,j-2)\right], \tag{6}$$

with $f(-1,0) = 0, f(0,-1) = 0$ *and* $i, j = 0, 1, \ldots$

Theorem

The PMF of the BivPA$(\lambda_1, \lambda_2, \lambda_3, \rho_1, \rho_2)$ distribution is given by:

$$f(i,0) = \sum_{m=1}^{i} \binom{i-1}{i-m} \frac{[\lambda_1(1-\rho_1)]^m}{m!} \rho_1^{i-m} f(0,0), \quad i = 1, 2, \dots,$$

$$f(0,j) = \sum_{l=1}^{j} \binom{j-1}{j-l} \frac{[\lambda_2(1-\rho_2)]^j}{j!} \rho_2^{j-l} f(0,0), \quad j = 1, 2, \dots,$$

$$f(i,j)$$

$$= \Bigg[\left(\sum_{m=1}^{i} \binom{i-1}{i-m} \frac{[\lambda_1(1-\rho_1)]^m}{m!} \rho_1^{i-m} \right) \left(\sum_{l=1}^{j} \binom{j-1}{j-l} \frac{[\lambda_2(1-\rho_2)]^l}{l!} \rho_2^{j-l} \right)$$

$$+ \sum_{k=1}^{i \wedge j} \frac{(\lambda_3)^k}{k!} \sum_{m=0}^{i-k} \binom{i-1}{i-k-m} \frac{[\lambda_1(1-\rho_1)]^m}{m!} \rho_1^{i-k-m}$$

$$\times \sum_{l=0}^{j-k} \binom{j-1}{j-k-l} \frac{[\lambda_2(1-\rho_2)]^l}{l!} \rho_2^{j-k-l} \Bigg] f(0,0), \quad i, j = 1, 2, \dots,$$

with $f(0,0) = e^{-(\lambda_1+\lambda_2+\lambda_3)}$.

Moments of the Bivariate Noncentral Pólya-Aeppli distribution

For the means of M and N we obtain:

$$\mathrm{E}(M) = \frac{\lambda_1}{1-\rho_1} + \lambda_3$$

and

$$\mathrm{E}(N) = \frac{\lambda_2}{1-\rho_2} + \lambda_3$$

The variances of M and N are:

$$\mathrm{Var}(M) = \lambda_3 + \frac{\lambda_1}{1-\rho_1}\left(1 + \frac{2\rho_1}{1-\rho_1}\right)$$

and

$$\mathrm{Var}(N) = \lambda_3 + \frac{\lambda_2}{1-\rho_2}\left(1 + \frac{2\rho_2}{1-\rho_2}\right).$$

The product moment of M and N is given by

$$\mathrm{E}(MN) = \lambda_3 + \left(\lambda_3 + \frac{\lambda_1}{1-\rho_1}\right)\left(\lambda_3 + \frac{\lambda_2}{1-\rho_2}\right)$$

which readily yield to the covariance between M and N to be

$$\mathrm{Cov}(M, N) = \lambda_3$$

The correlation coefficient:

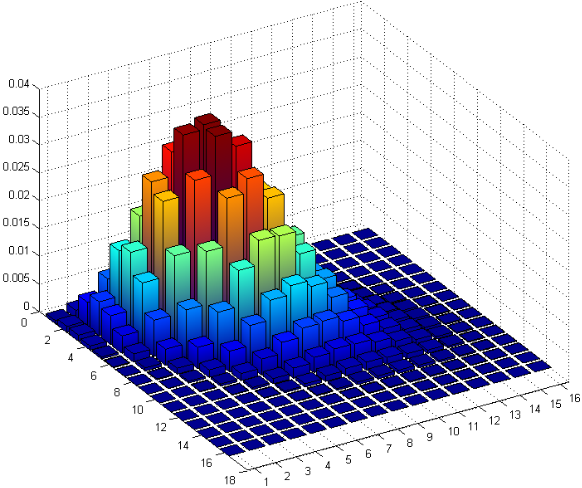
$$\mathrm{Corr}(M, N)$$

$$= \frac{\lambda_3 (1-\rho_1)(1-\rho_2)}{\sqrt{\left[\lambda_3 (1-\rho_1)^2 + \lambda_1 (1+\rho_1)\right] \left[\lambda_3 (1-\rho_2)^2 + \lambda_2 (1+\rho_2)\right]}}$$

Remark

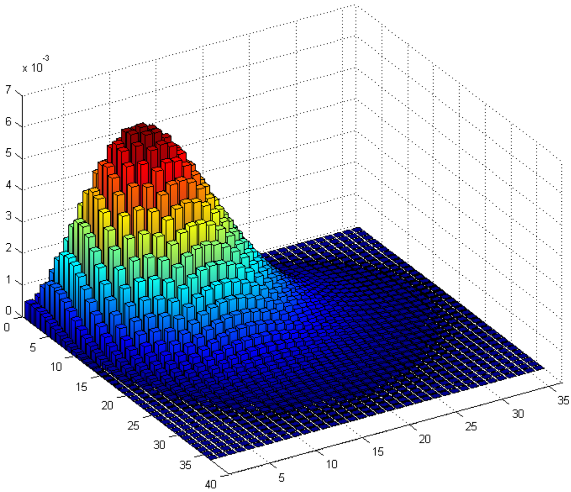
When $\rho_1 = \rho_2 = 0$ the PMF of the Noncentral Pólya-Aeppli distribution coincides to the bivariate Poisson distribution.

$\rho_1 = \rho_2 = 0$

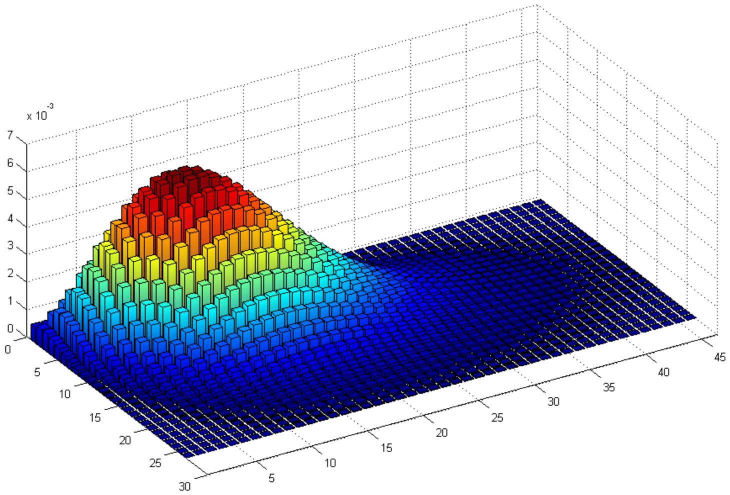


0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
0
2
4
6
8
10
12
14
16
18
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16

$\rho_1 = \rho_2 = 0.5$



$\rho_1 = 0.4, \ \rho_2 = 0.6$



References

1. Chukova S. and Minkova L.D. (2013) Characterization of the Pólya - Aeppli process, *Stochastic Analysis and Applications*, **31**, 590–599.
2. Grandell, J. (1997) *Mixed Poisson Processes*, Chapman & Hall, London.
3. Klugman S.A., Panjer H. and Willmot G. (2004) *Loss Models. From Data to Decisions*, 2nd edition, John Wiley & Sons, Inc.
4. Minkova L. (2002) A generalization of the classical discrete distributions, *Commun. Statist.-Theory and Methods*, **31**, 871–888.
5. Minkova L.D. (2004) The Pólya-Aeppli process and ruin problems, *J. Appl. Math. Stoch. Anal.*, **2004(3)**, 221–234.
6. Minkova L.D. (2011) I-Pólya Process and Applications, *Commun. Statist. - Theory and Methods*, **40**, 2847–2855.
7. Minkova L. D. (2012) *Distributions in Insurance Risk Models*, Doctor of Science Thesis, available in: www.fmi.uni-sofia.bg.

8. Ong S. H. and Lee P. A. (1979) The non-central negative binomial distribution, *Biom. J.*, **21**, 611–628.
9. Panjer H.(1981) Recursive evaluation of a family of compound discrete distributions, *ASTIN Bulletin*,**12**, 22–26

Thanks

Thank you for your kind attention !