PRINT ISSN 1119-8362
Electronic ISSN 2659-1499

Full-text Available Online at
https://www.ajol.info/index.php/jasem
https://www.bioline.org.br/ja

J. Appl. Sci. Environ. Manage.
Vol. 29 (2) 401-405 Feb. 2025

Sum of Poisson-Distributed Random Variables: A Convolution Method Approach

***AYENIGBA, AA; AJAO, OM; OKOLIE, FA**

Department of Mathematical Sciences, Ajayi Crowther University Oyo, Nigeria

**Corresponding Author Email: aa.ayenigba@acu.edu.ng*
**ORCID: https://orcid.org/0009-0000-0657-7570*
**Tel: +2348138379326*

Co-Author Email: om.ajao.@acu.edu.ng; fa.okolie@acu.edu.ng

ABSTRACT: This paper presents a two-parameter extension of the classical Poisson distribution, specifically tailored for rare event modeling. The proposed model is constructed as the sum of two independent Poisson random variables, using a convolution method. Some properties of the distribution, including the probability mass function (PMF), moment-generating function (MGF), mean, variance, higher-order moments, Skewness, and kurtosis, are derived.

DOI: https://dx.doi.org/10.4314/jasem.v29i2.8

Cite this Article as: AYENIGBA, A. A; AJAO, O. M; OKOLIE, F. A (2025). Sum of Poisson-Distributed Random Variables: A Convolution Method Approach. *J. Appl. Sci. Environ. Manage.* 29 (2) 401-405

Dates: Received: 23 December 2024; Revised: 27 January 2025; Accepted: 09 February 2025; Published: 28 February 2025

Keywords: Rare event; convolution method; Poisson distribution; Skewness; Kurtosis

The Poisson distribution is a significant probabilistic model utilized to characterize the occurrence of sporadic, random phenomena within a defined time interval or spatial region. Initially formulated by Siméon-Denis Poisson in 1837, it has since found extensive applications across diverse domains, such as telecommunications, public health, and traffic studies (Cox and Isham, 1980). This model is determined by a single parameter, λ, representing the frequency at which events take place over the specified period. A fundamental attribute of the Poisson distribution is its independence assumption, indicating that the occurrence of one event does not impact the likelihood of subsequent events. This feature significantly enhances its relevance in numerous practical applications (Chavez-Demoulin *et al*., 2021).

The Poisson distribution possesses noteworthy statistical characteristics. For example, both the mean and variance of a Poisson random variable are equivalent to λ, emphasizing its property of equidispersion, wherein the variability in event counts is directly proportional to the average rate. Additionally, it exhibits a lack of memory, signifying that the probability of future events remains unaffected by past occurrences. This feature is particularly advantageous for examining systems governed by independent events over time (Johnson *et al*., 1994).

In this study, we expand upon the Poisson distribution by introducing a novel two-parameter framework, which represents the aggregation of two independent Poisson-distributed random variables through the convolution technique. This generalization seeks to address more sophisticated scenarios where event frequencies may vary. By enhancing the analytical scope, we aim to deepen insights into aggregated event behaviors in practical contexts. The paper further explores the statistical

**Corresponding Author Email: aa.ayenigba@acu.edu.ng*
**ORCID: https://orcid.org/0009-0000-0657-7570*
**Tel: +2348138379326*

features of this new framework, shedding light on its potential utility across multiple disciplines.

The convolution methodology, a well-recognized approach for deriving probabilistic models, serves as the foundation for this research. This technique is broadly employed in constructing both discrete and continuous probability distributions. For instance, the Negative Binomial Distribution, derived from the convolution of geometric random variables (Johnson *et al*., 2005), and the Compound Poisson Distribution, obtained from the convolution of Poisson random variables (Klugman *et al*., 2012), exemplify its application in discrete scenarios. Similarly, in continuous contexts, convolution has been crucial in generating distributions such as the sum of exponential random variables (Oguntunde *et al*., 2013), the Beta-Exponential convolution (Shitu *et al*., 2012; Mdziniso, 2012), and the Beta-Weibull convolution (Nadarajah and Kotz, 2006; Sun, 2011). The objective of this paper is to evaluate a twoparameter extension of the classical Poisson distribution.

MATERIALS AND METHODS

The convolution method is a fundamental technique in probability theory used to determine the distribution of the sum of two independent random variables. Given two independent and identically distributed (i.i.d.) random variables Y_1 and Y_2, their sum $Z = Y_1 + Y_2$ has a probability distribution obtained via convolution. This method is particularly useful for discrete distributions, such as the Poisson and Binomial distributions (Ross, 2014).

If Y_1 and Y_2 are independently and identically distributed Poisson random variables with parameter λ_1 and λ_2 respectively,

Then, the probability of the sum $Z = Y_1 + Y_2$ is given by the convolution of their individual probability mass functions:

$$F_z(Z) = P(Z = z) = \sum_{m=0}^{z} P(Y_1 = m)P(Y_2 = z - m) \quad (1)$$

This formula represents the summation over all possible values of Y_1 and Y_2, such that their sum equals Z. Intuitively, this method accounts for all ways the individual values of Y_1 and Y_2, can combine to form Z (Feller, 1968).

The Probability Mass Functions of X_1 and X_2 are given below:

$$P(Y_1 = m) = \frac{\lambda_1^{m}\, e^{-\lambda_1}}{m!}, \quad m = 0, 1, 2, \ldots \quad (2)$$

$$P(Y_2 = z - m) = \frac{\lambda_2^{z-m}\, e^{-\lambda_2}}{(z-m)!}, \quad m = 0, 1, 2, \ldots \quad (3)$$

Using (1), we have:

$$F_z(Z) = P(Z) = \sum_{m=0}^{z} \frac{\lambda_1^{m}\, e^{-\lambda_1}}{m!} . \frac{\lambda_2^{z-m}\, e^{-\lambda_2}}{(z-m)!}$$

$$= e^{-(\lambda_1+\lambda_2)} \sum_{m=0}^{z} \frac{\lambda_1^{m}}{m!} . \frac{\lambda_2^{z-m}}{(z-m)!}$$

$$= e^{-(\lambda_1+\lambda_2)} \frac{1}{z!} \sum_{m=0}^{z} \frac{z!}{m!\,(z-m)!} \lambda_1^{m} \lambda_2^{z-m}$$

$$= \frac{e^{-(\lambda_1+\lambda_2)}}{z!} \sum_{m=0}^{z} \binom{z}{m} \lambda_1^{m} \lambda_2^{z-m}$$

$$P(Z) = \frac{e^{-(\lambda_1+\lambda_2)}((\lambda_1+\lambda_2))^{z}}{z!}, z = 0, 1, 2, \ldots \quad (4)$$

The model in equation (4) above represents the probability model for the sum of two independently and identically distributed Poisson random variables

RESULTS AND DISCUSSION

Validity of the Model $P(Z)$: The model is valid if and only if

$$\sum_{m=0}^{z} P(Z) = 1$$

$$\sum_{m=0}^{z} \frac{e^{-(\lambda_1+\lambda_2)}\big((\lambda_1+\lambda_2)\big)^{z}}{z!}$$

$$= e^{-(\lambda_1+\lambda_2)} \sum_{m=0}^{z} \frac{((\lambda_1+\lambda_2))^{z}}{z!}$$

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

$$= 1$$

The Cumulative Density Function (CDF) of the Model P(Z)

By definition, CDF is derived by:

$$F_z(Z) = P(Z \le z)$$

$$= \sum_{m=0}^{z} P_z(Z = m)$$

$$= \sum_{m=0}^{z} \frac{e^{-(\lambda_1+\lambda_2)}\big((\lambda_1+\lambda_2)\big)^{m}}{m!}$$

$$= e^{-(\lambda_1+\lambda_2)} \sum_{m=0}^{z} \frac{((\lambda_1+\lambda_2))^z}{z!}$$

It can be deduced that as $z \to \infty$, *the CDF as approach* 1

The Shape of the Model: The shape is distribution is determined by obtaining the mode of the distribution. The mode can be obtained by comparing consecutive probabilities $P_z(Z = z) = P_z(Z = z + 1)$.

We find the point where $P_z(Z = z + 1) \leq P_z(Z = z)$ (6)

$$\frac{P_z(Z = z+1)}{P_z(Z = z)} \leq 1$$

$$\frac{\frac{e^{-(\lambda_1+\lambda_2)}((\lambda_1+\lambda_2))^z}{z!}, (5)}{\frac{e^{-(\lambda_1+\lambda_2)}((\lambda_1+\lambda_2))^{z+1}}{(z+1)!}} \leq 1$$

$$\frac{\lambda_1+\lambda_2}{z} \leq 1$$

$z \geq (\lambda_1+\lambda_2) - 1$ (7)

Since $(\lambda_1+\lambda_2)$ must be an integer, the mode of the model is the largest integer less than or equal to parameter.

$z = \lfloor \lambda_1 + \lambda_2 \rfloor$

It is Unimodal; it has a single peak at the mode.

The graph in Figure 1 shows the shape of the distribution for $(\lambda_1+\lambda_2) = 2$ and $(\lambda_1+\lambda_2) = 3$

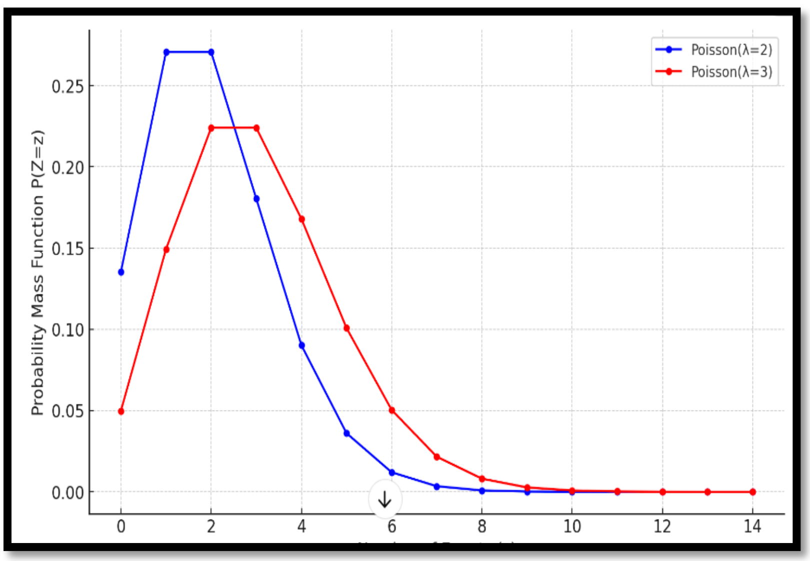

Fig. 1: Graph for pdf of Z (where $(\lambda_1+\lambda_2) = 2$ and $(\lambda_1+\lambda_2) = 3$)

The graph illustrates that as the parameter for $(\lambda_1+\lambda_2)$ increases from 2 to 3, the distribution becomes more concentrated, with a higher peak. This suggests an increased probability of observing values closer to the mean. The increased Peakness is accompanied by a wider spread and a longer right tail, indicating that the distribution remains Unimodal.

Parameter Estimation: Let $Z_1, Z_2, Z_3, \ldots, Z_n$ denote a random sample of n independent and identically distributed random variables, each with the probability density function (pdf) as derived in Equation (2) above. Applying the method of maximum likelihood estimation, the likelihood function is expressed as:

$$L(Z/(\lambda_1+\lambda_2)) = e^{-n(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^{\sum z} \prod_{i=1}^{n} \frac{1}{z_i} \quad (8)$$

Let $L = Log_e(Z/(\lambda_1+\lambda_2)$

$$L = Log_e\left[e^{-n(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^{\sum z} \prod_{i=1}^{n} \frac{1}{z_i}\right]$$

$$L = -n(\lambda_1+\lambda_2) + \sum_{i=1}^{n} z_i \, Log_e(\lambda_1+\lambda_2) - Log_e\left(\prod_{i=1}^{n} \frac{1}{z_i}\right)$$

$$\frac{\partial L}{\partial(\lambda_1+\lambda_2)} = -n + \frac{\sum z}{(\lambda_1+\lambda_2)} = 0 \quad (9)$$

Solving equation (3) above we have:
$\widehat{(\lambda_1 + \lambda_2)} = \bar{Z}$

Hazard Function: According to its definition, the hazard function for a random variable Z is given by:

$$h(z) = \frac{f(z)}{1 - F(z)}$$

$$h(z) = \frac{\frac{e^{-(\lambda_1+\lambda_2)}\left((\lambda_1 + \lambda_2)\right)^z}{z!}}{1 - e^{-(\lambda_1+\lambda_2)} \sum_{m=0}^{z} \frac{\left((\lambda_1 + \lambda_2)\right)^m}{m!}} \quad (10)$$

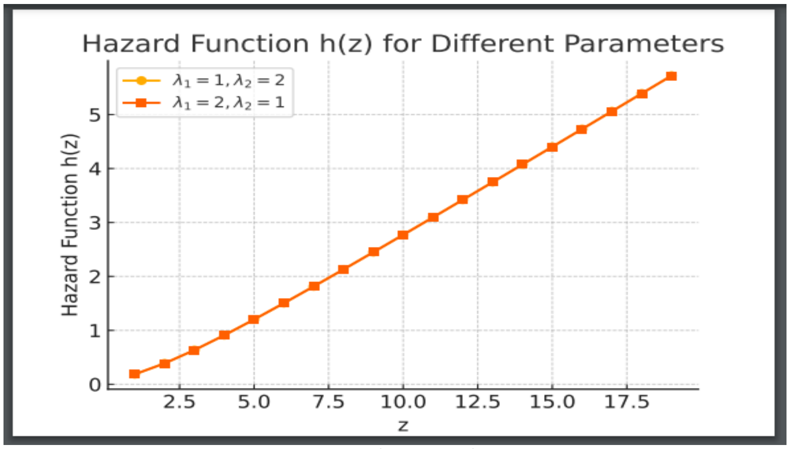

Fig. 2: Graph for hazard function

Figure 2 demonstrates that the hazard rate increases as variable Z increases, indicating a higher initial risk or probability of occurrence that diminishes over time. This suggests that the model in Equation (2) is well-suited for events characterized by a high initial risk that gradually decreases as time progresses.

Asymptotic Behaviour of the Model: We aim to analyze the behaviour of the model in equation (4) as Z →0 and as Z → ∞. This includes evaluating $\lim_{z\to 0} f(z)$ and $\lim_{z\to 0} F(z)$

$$\lim_{z\to 0} f(z) = \lim_{z\to 0}\left[\frac{e^{-(\lambda_1+\lambda_2)}((\lambda_1 + \lambda_2))^z}{z!}\right]$$
$$= e^{-(\lambda_1+\lambda_2)} \quad (11)$$

$$\lim_{z\to \infty} f(z) = \lim_{z\to \infty}\left[\frac{e^{-(\lambda_1+\lambda_2)}((\lambda_1 + \lambda_2))^z}{z!}\right] = 0 \quad (12)$$

These results further confirm that the model presented in Equation (4) exhibits a single mode (uni-modal).

Moment Generating Function: The moment generating function (m.g.f.) of a random variable Z is represented by:

$M_z(t) = E(e^{tz}),\quad (13)$

Where $Z = Y_1 + Y_2$

$M_{Y_1+Y_2}(t) = E(e^{ty_1}).E(e^{ty_2})$

Where:

$M_{Y_1}(t) = e^{\lambda_1(e^t-1)} \quad (14)$

$M_{Y_2}(t) = e^{\lambda_2(e^t-1)} \quad (15)$

Where $E(e^{ty_1})$ *and* $E(e^{ty_2})$ are moments generating functions for convoluted Poisson distribution.

$M_z(t) = e^{(\lambda_1+\lambda_2)(e^t-1)} \quad (16)$

The Characteristic function of the model is given as:

$\emptyset_z(t) = E(e^{tz}) \quad (17)$

$\emptyset_z(t) = e^{(\lambda_1+\lambda_2)(e^{it}-1)} \quad (18)$

From the result in Equation (18), we can confidently generalize that if $Y_1, Y_2, Y_3, \dots Y_n$ are independently and identically distributed random variables, each having Exponential distribution with parameter, the moment generating function of the sum can be expressed as $Z = Y_1 + Y_2 + Y_3 + \cdots + Y_n$

$M_z(t) = e^{(\sum_{i=1}^{n}\lambda_i)(e^t-1)} \quad (19)$

Moment: The rth raw moment of a random variable Z is given by:

$$E(Z^r) = \frac{d^r M_z(t)}{dt^r}\bigg| t = 0 \quad (20)$$

From Equation (7), the first 4 moments are derived below:

$E(Z) = M'_z(0) = (\lambda_1 + \lambda_2) \quad (21)$

$E(Z^2) = M''_z(0) = (\lambda_1 + \lambda_2)^2 + (\lambda_1 + \lambda_2) \quad (22)$

$E(Z^3) = M'''_z(0)$
$= (\lambda_1 + \lambda_2)^3 + 3(\lambda_1 + \lambda_2)^2 + (\lambda_1 + \lambda_2) \quad (23)$

$E(Z^4) = M'^v_z(0)$
$= (\lambda_1 + \lambda_2)^4 + 6(\lambda_1 + \lambda_2)^3 + 7(\lambda_1 + \lambda_2)^2 + (\lambda_1 + \lambda_2) \quad (24)$

The following results are obtained from moment above:

The mean:
$$E(z) = M'_z(0) = (\lambda_1 + \lambda_2) \quad (25)$$

The variance:
$$Var(z) = M''_z(0) - \left(M'_z(0)\right)^2 = (\lambda_1 + \lambda_2) \quad (26)$$

Skewness:
$$Sk = \frac{E[(Z - E[Z])^3]}{(Var(Z))^{3/2}} = \frac{1}{\sqrt{\lambda_1 + \lambda_2}} \quad (27)$$

Kurtosis:
$$KU = \frac{E[(Z - E(Z))^4]}{(Var(Z))^2} = 3 + \frac{1}{\sqrt{\lambda_1 + \lambda_2}} \quad (28)$$

Conclusion: In this study, a new two-parameter distribution by combining two independent Poisson distributions through convolution has been developed. Our analysis showed that this distribution is positively skewed and has a single peak, making it Unimodal. The shape of its hazard function suggests that it's well-suited for situations where the likelihood of an event initially starts high and then gradually decreases over time. This makes it particularly useful for applications in insurance and reliability fields, such as predicting claim frequencies in insurance policies or failure rates in reliability studies, where risks tend to decrease over time. Overall, this model offers a practical way to capture patterns in data where frequencies play a central role.

Declaration of Conflict of Interest: The authors declare no conflict of interest.

Data Availability Statement: Data is available upon request from the first author or corresponding author or any of the other authors

REFERENCES

Chavez-Demoulin, V; Embrechts, P; Hofert, M (2021). An Introduction to Applied Probability (3rd ed.). Springer, p. 350.

Cox, DR; Isham, V (1980). Point Processes. Chapman and Hall, p. 188.

Feller, W. (1968). An Introduction to Probability Theory and Its Applications, Vol. 1 (3rd ed.). Wiley. p. 160.

Johnson, NL; Kotz, S; Kemp, AW (1994). Univariate Discrete Distributions (2nd ed.). Wiley, p. 565.

Johnson, NL; Kemp, AW; Kotz, S (2005). Univariate Discrete Distributions (3rd ed.). Wiley, p. 648.

Klugman, SA; Panjer, HH; Willmot, GE (2012). Loss Models: From Data to Decisions (4th ed.). Wiley, p. 752.

Montgomery, DC; Runger, GC (2003). Applied Statistics & Probability for Engineers (3rd ed.). Wiley, p. 640.

Mdziniso, N (2012). The Convoluted Beta-Exponential Distribution and Its Applications [Unpublished master's thesis]. University of Swaziland.

Nadarajah, S; Kotz, S (2006). The Beta Weibull Distribution. *Reliab. Eng. Syst. Saf.* 91(6): 689–697. DOI: https://doi.org/10.1016/j.ress.2005.07.003

Oguntunde, PE; Odetunmibi, OA; Adejumo, AO (2013). On the Sum of Two Exponential Random Variables. *Int. J. Probab. Stat. Stud.* 1(4): 25–30.

Ross, S. M. (2014). Introduction to Probability Models (11th ed.). Academic Press p. 66.

Sun, Y (2011). On the Convoluted Beta-Weibull Distribution [Master's thesis, University of Alberta]. *Univ. Alberta Libr.*