RESEARCH ARTICLE

Modeling Biological Uncertainty Using the Double Fuzzy Poisson Distribution

Adel Abbood Najm1*, Bashar Khalid Ali 2

¹ Department of Statistics, Administration and Economics College, Sumer University-Iraq; ² Department of Statistics, Administration and Economics College, Kerbala University-Iraq.

ABSTRACT

This paper introduces a novel probability distribution, the Double Fuzzy Poisson Distribution, designed to address two levels of uncertainty commonly encountered in biological data analysis. The first layer of fuzziness arises from measurement inaccuracies, where observed data are represented as fuzzy numbers. The second layer accounts for uncertainty in the distribution parameter itself, which is also expressed as a fuzzy number, incorporating prior information from previous studies or expert opinions related to the biological phenomenon under investigation. To evaluate the performance of the proposed distribution, extensive Monte Carlo simulations were conducted alongside applications to real-world data involving HIV infection rates. The results demonstrate that the Double Fuzzy Poisson Distribution offers superior accuracy and greater flexibility compared to the traditional Poisson distribution. Specifically, it provides more reliable estimations of infection rates and enhances risk analysis by effectively capturing the inherent uncertainty in medical data. These findings suggest that the proposed model is a robust and adaptable tool for handling biological data characterized by multiple sources of imprecision, making it a valuable addition to the field of biomathematical modeling.

Introduction

In recent years, fuzzy set theory has found important application in developing models of statistical distributions in a fuzzy environment. The Poisson distribution forms the basis of many probabilistic models, although there are very few fuzzy versions of it. The fuzziness concept related to the theory of logic of fuzzy, as it allows us to deal with data or variables that contain uncertainty or ambiguity. In many phenomena in nature and within the traditional concept of it, this can be considered accurate measurements. But in many of them there is not enough accuracy for measurement and thus falls within the scope of fuzzy logic. Unlike traditional logic, this takes into account the affiliation of each element with a certain degree of affiliation. Such phenomena include the Poisson distribution, which depends on one parameter, which is the rate of occurrence of the event, which we cannot be accurate in measuring, so it can be considered fuzzy and expressed as a fuzzy set. Thus, reaching a fuzzy distribution that is more accurate than the traditional distribution. At the same time, there can be data fuzziness, which refers to the uncertainty or ambiguity in the values or information that are collected and analyzed.

Traditionally, statistical data is precise, where values are precisely defined, such as height, weight, or time, but in many real-world cases, these values may be imprecise or contain a degree of ambiguity. Therefore, this research aims to find a double fuzzy Poisson distribution that takes into account the fuzziness in the distribution parameter as well as in the data in order to increase the accuracy in measuring the probability of occurrence of events. M.A.S. Monfared & J.B. Yang, combined fuzzy logic (Monfared & Yang, 2004), control theory and optimization theory to develop an intelligent manufacturing control system considering that the system parts are dynamically arriving according to the Poisson distribution (Garg et al., 2013). Jamkhaneh Ezzatallah et al. established a double acceptance sampling plan is the percentage of false items is a fuzzy number and modeled it according to Poisson distribution considering fuzziness by observations and parameter (Jamkhaneh & Gildeh, 2011). Moraes & Machado, converted the Naiv Bayes classifier to the fuzzy Naiv Bayes classifier considering that the data is inaccurate and found the fuzzy probability Poisson distribution using certain belonging functions (Moraes & Machado, 2015). Neamah & Ali, estimated the fuzzy reliability of the Frechet distribution by transforming it into a fuzzy distribution with the data (Ali & Neamah, 2022). Sirbiladze et al., expanded the binomial distribution into a fuzzy distribution using data transformations (Sirbiladze et al., 2022). Shebib et al., transformed the Lindley

ARTICLE HISTORY

Received December 25, 2023 Accepted June 23, 2024

KEYWORDS

Poisson Distribution, Double Fuzzy, Maximum Likelihood Estimation, Membership Function, Fuzziness, HIV Infections 52

Najm, A. A: Modeling Biological Uncertainty Using the Double Fuzzy Poisson Distribution

distribution into a fuzzy distribution using rainfall data in Iraq (Shebib et al., 2023).

2. Fuzzy Information System (FIS)

Let $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_n)$ be a fuzzy events represent all subsets in Ω with membership functions that have a Borel measure witch have orthogonally $\sum_{\tilde{x} \in X} \mu_{\tilde{x}}(x) = 1$, then \tilde{A} is a fuzzy subset from Ω (Ali & Neamah, 2022; Sivanandam et al., 2007).

3. Fuzzy and Crisp Set

If Ω is a set that all sets must include, and A is a subset of it, then each element x in A may or might not be a member of set A. The binary characteristic function of set A, denoted as $\mu_A(x)$, assigns a degree of membership to set A to each element in set A. {0,1}:(Ali & Neamah, 2022; Neamah & Ali, 2020)

$$\mu_A(x) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{if } x \notin A \end{cases}$$
(1)

The fuzzy subset \widetilde{A} of Ω is defined by a membership function $\mu_{\widetilde{A}}(x)$ that, for every value of x in the fuzzy sample space, generates a value between 0 and 1. The set of ordered pairs is called the fuzzy set:

$$\hat{A} = \{ (x_i, \mu_{\widetilde{A}}(x_i)), x \in \Omega, i = 1, 2, 3, \dots, n, 0 \le \mu_{\widetilde{A}}(x)$$
(2)

4. Membership Functions

A positive value is assigned to the function that transfers the element's significance (its degree of belonging) from the universal set to the fuzzy set. For use in creating fuzzy set memberships, it takes a value between 0 and 1 to indicate the degree to which each fuzzy set element is a member of the conventional universal set (Abboudi et al., 2020; Sun & Choi, 2023). A membership function graph would show the normal values of the fuzzy variable on the X-axis and the degree of group membership on the Y-axis. The most fundamental requirement for these functions is that their range be from 0 to 1. Values that have a belonging degree of 1 are considered to be part of the group, whilst values with a belonging degree of 0 are considered to be outside of the group. The degrees of membership of the items in the group are defined by the intermediate value between the two values (0, 1). A variety of fuzzy sets are generated from various kinds of belonging functions, which are generalizations of the standard set's characteristic function. The membership functions have three main properties:(Chaira, 2019; Yahia et al., 2012)

1- Core: If \widetilde{A} is a fuzzy set, its core is when its degree of membership is complete and equals 1.

2- Support (base): If \tilde{A} is a fuzzy set, then the elements included these set whose degree of membership it is greater than zero i.e.,

Support
$$(\widetilde{A}) = \{x \in \Omega / \mu_{\widetilde{A}}(x) > 0\}$$
 (3)

3- Boundary: the elements included set \widetilde{A} whose membership degree greater than zero and are incomplete, i.e.,

Boundary
$$(\widetilde{A}) = \{x \in \Omega; 0 < \mu_{\widetilde{A}}(x) < 1\}$$
 (4)



Figure 1: Properties of Membership Function (Neamah & Ali, 2020)

5. Triangular Membership Function

It is one of the linear MF recognized by three parameters: a lower limit (a), an upper limit (b), and is conditional on a central value (m). Its formula is as follows: (Akman et al., 2023; Ali & Neamah, 2022; Barros et al., 2017; Monfared & Yang, 2004; Neamah & Ali, 2020; Yahia et al., 2012)

Figure 2: Triangular MF

6. Trapezoid MF

It is one of the linear functions that has four parameters, a minimum (a) and an upper limit (d) and is conditional on two central values (b) and (c) respectively. Its formula is as follows: (Ali & Neamah, 2022; Neamah & Ali, 2020)

7. Fuzzy Discrete Probability Distribution

Let Ω universal set with probability measure $(\Omega, \sigma, P_{\theta})$ and (Ω, σ) is measurable space and σ is sigma field, $P_{\theta}, \theta \in \Theta$ probability measure, and let A be a subset in Ω . The probability of point x as follows (Burak Parlak & Cağrı Tolga, 2016),

54 Najm, A. A: Modeling Biological Uncertainty Using the Double Fuzzy Poisson Distribution

 $P_{\theta}(x_i) = \left\{ \mathsf{p}, \mu_{P_{\theta}(x_i)}(\mathsf{p}) \mid P_{\theta} \in [0,1] \right\}$

$$P_{\theta}(x_i) = \sum_{i=1}^{n} P_{\theta} \cdot \mu_{P_{\theta}(x_i)}(P_{\theta}) = \mathbb{E}(\mu_{P_{\theta}(x_i)}(P_{\theta}))$$
(7)

Where $\mu_{P_{\theta}(x_i)}(p)$ membership function for P_{θ} , \tilde{n} sample size of fuzzy set.

7.1 Poisson Distribution (PD)

Let $\underline{X} = (x_1, x_2, ..., x_n)$ be a crisp random sample having the following probability mass function with crisp parameter θ , (El-Dawoody et al., 2023; Zhao et al., 2020)

$$p(x,\theta) = \frac{\theta^{x}e^{-\theta}}{x!}$$
; $x = 0, 1, 2, ...$ (8)

Where x and θ is measured accurately and expressed in precise numbers, then the distribution is crisp with data and parameters. Figure (1) showed the probability mass function of Poisson distribution under different values of θ (Neamah & Ali, 2020).



Figure 4: Traditional Poisson Distribution for Different Valued of Lambda

8. Proposed Fuzzy Poisson Distribution (FPD)

FPD can be found by suggesting that the data is fuzzy and the parameters are also fuzzy, as follows:

8.1 Fuzzy Data

Assuming that the measured data is inaccurate and is expressed in fuzzy numbers, whether triangular, trapezoidal, etc., and let $\underline{\tilde{A}} = (\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_n)$ is fuzzy set where,

$$\underline{\tilde{A}} = \{ \tilde{x}, \tilde{x} = (\tilde{x}_i, \mu_A(\tilde{x}_i)); 0 < \mu_A(\tilde{x}_i) < 1 \}$$
(9)

The traditional Poisson distribution becomes a fuzzy distribution with the fuzzy data according to the following fuzzy probability mass function:

$$p(\tilde{x}_i) = \frac{\theta^{\tilde{x}_i} e^{-\theta}}{\tilde{x}_i!} \quad ; \tilde{x}_i = 0, 1, 2, \dots$$
 (10)

8.2 Fuzzy Parameters

Assuming that the distribution parameter is inaccurate and that there are previous parameters for it resulting from previous studies or according to the opinion of those specialized in the studied phenomenon and be as fuzzy number, Let us have k available parameters each have the membership degree as fuzzy number obtain with membership function,

$$\underline{\tilde{\theta}} = \left\{ \tilde{\theta}_{i} \left(\tilde{\theta}_{i}, \mu_{\theta}(\tilde{\theta}_{i}) \right); \ 0 < \mu_{\theta}(\tilde{\theta}_{i}) < 1 \right\}$$
(11)

After that, the parameter of FPD obtained by extract the weighted average for $\tilde{\theta}_i$ by assuming the membership degrees as weigh on parameter. The Fuzzy Poisson distribution becomes a fuzzy distribution with the fuzzy parameters according to the following fuzzy probability mass function:





Figure 5: Traditional and Fuzzy Poisson Distribution for Different Valued of Lambda

9. Maximum Likelihood Estimation

Let $\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_n$ be a fuzzy observation from Poisson distribution with fuzzy parameter $\tilde{\theta}$ with the following FPMF.

$$p(\tilde{x}_i, \tilde{\theta}) = \frac{\tilde{\theta}^{\tilde{x}_i} e^{-\tilde{\theta}}}{\tilde{x}!} \quad i = 0, 1, 2, \dots$$
(13)

The likelihood function $L(\theta)$ for observations as following,

$$L(\tilde{\Theta}) = \prod_{i=1}^{\tilde{n}} p(\tilde{x}_i, \tilde{\Theta})$$
$$= \prod_{i=1}^{\tilde{n}} \frac{\tilde{\Theta}^{\tilde{x}_i} e^{-\tilde{\Theta}}}{\tilde{x}!}$$

Then the Log $L(\tilde{\theta})$ is,

$$Log\left(L(\tilde{\theta})\right) = Log\left(\prod_{i=1}^{\tilde{n}} \frac{\tilde{\theta}^{\tilde{x}_i} e^{-\tilde{\theta}}}{\tilde{x}!}\right)$$
$$= \sum_{i=1}^{\tilde{n}} \tilde{x}_i Log(\tilde{\theta}) - \tilde{n}\tilde{\theta} - \tilde{n}Log(\tilde{x}!)$$
(14)

By derivative the equation (3) according to parameter $\tilde{\theta}$ and equal to zero we obtain,

$$\frac{\partial Log\left(L\left(\tilde{\Theta}\right)\right)}{\partial \tilde{\Theta}} = \frac{\sum_{i=1}^{\tilde{n}} \tilde{x}_i}{\tilde{\Theta}} - \tilde{n} = 0$$
$$\frac{\partial Log\left(L(\tilde{\Theta})\right)}{\partial \tilde{\Theta}} = \frac{\sum_{i=1}^{\tilde{n}} \tilde{x}_i}{\tilde{\Theta}} - \tilde{n} = 0$$
$$\tilde{\Theta} = \frac{\sum_{i=1}^{\tilde{n}} \tilde{x}_i}{\tilde{n}}$$
(15)

10. Simulation Study

We used the Monte-Carlo simulation to evaluate the suggested distribution and compared it with traditional distribution. First, we generated the Poisson samples (n=25, 50, 75, 150, 200), and extract the values of memberships by

55

using triangular membership function with (a, b, c). Second we assume that different theoretical values of traditional Poisson distribution (θ =2, 3.5, 5) and simulate the theoretical parameters of fuzzy by generated Poisson random variable and compute the mean for each r.v. and extract the membership values from triangular membership function with (a, b, c) where a minimum value, b centroid value and c, maximum value of parameter , and then compute the weighted average for these parameters. The maximum likelihood method was used to estimate the parameter of TPD and FPD and compared the distributions by using MSE M the results as the following tables:

 Table 1: Values of the Theoretical Probability Density Function Estimated by the Maximum Likelihood Method for the

 Traditional and Fuzzy Distribution for all Assumed Sample Sizes. The PMF in Tables Represent the mean of PMF under

 each Sample Size for TPD and FPD for First Example



Figure 5: Traditional and Fuzzy Poisson Distribution at True and MLE parameter for a) n=25, b) n=50, c) n=75, d) n=150 Respectively for First Example for Second Example

 Table 2: Values of the Theoretical Probability Density Function Estimated by the Maximum Likelihood Method for the

 Traditional and Fuzzy Distribution for all Assumed Sample Sizes. The PMF in Tables Represent the mean of PMF under

 each Sample Size for TPD and FPD for Second Example

θ =3.5 Mean (θ =2) =3.1											
Ν	TPMF	MLE TPMF	MSE	FPMF	MLE FPMF	MSE	Best				
25	0.21568	0.23874	0.00053	0.21269	0.21856	0.00003	FPMF				
50	0.21157	0.22923	0.00031	0.12357	0.12769	0.00002					
75	0.20317	0.21378	0.00011	0.19988	0.19744	0.00001					
150	0.26921	0.26134	0.00006	0.17855	0.17750	0.00000					

LETTERS IN BIOMATHEMATICS



Figure 6: Traditional and Fuzzy Poisson Distribution at True and MLE parameter for a) n=25, b) n=50, c) n=75, d) n=150 Respectively for Second Example

 Table 3: Values of the Theoretical Probability Density Function Estimated by the Maximum Likelihood Method for the

 Traditional and Fuzzy Distribution for all Assumed Sample Sizes. The PMF in Tables Represent the mean of PMF under

 each Sample Size for TPD and FPD for Third Example.





Figure 7: Traditional and Fuzzy Poisson Distribution at True and MLE parameter for a) n=25, b) n=50, c) n=75, d) n=150 Respectively for Third Example

The simulations results shows that the fuzzy Poisson distribution (FPMF) outperforms the conventional distribution in all examples and case studies in terms of accuracy and data fit, making it a better choice for dealing with data uncertainty.

11. Real Data Set

The data set represent the number of HIV infections in for the period from the year 2023 until end of the fifth month of the year 2024, which is approximate numbers with (25) observation from WHO site. The traditional Poisson distribution and the proposed fuzzy Poisson distribution were applied to it, and the results are in table (3).

Traditional x	Traditional λ	Traditional PMF	Fuzzy X	Fuzzy λ	FPMF
4	1.52	0.04864	3	0	0.00000
5	2.63	0.07558	3	1.31	0.10110
6	2.63	0.03313	4	1.31	0.03311
6	2.64	0.03355	3	1.32	0.10240
3	5.64	0.10623	3	1.33	0.10370
6	2.65	0.03398	1	3.34	0.11836
5	2.65	0.07694	2	3.44	0.18972
5	2.66	0.07763	2	4.36	0.12146
3	6.66	0.06308	2	3.88	0.15544
2	3.67	0.17157	1	4.37	0.05529
5	2.67	0.07831	0	3.38	0.03405
9	2.68	0.00135	3	1.39	0.11149
6	2.68	0.03528	2	3.66	0.17235
6	2.69	0.03572	3	1.41	0.11406
7	2.69	0.01373	3	1.42	0.11535
3	4.7	0.14094	3	0.43	0.00862
5	2.7	0.08036	1	3.11	0.13871
7	2.71	0.01417	2	3.44	0.17972
1	2.71	0.18031	4	1.45	0.04320
6	2.72	0.03705	0	3.46	0.03143
2	4	0.14653	2	3.47	0.18734
4	2.73	0.15094	0	2.48	0.08374
7	2.73	0.01462	2	3.9	0.15394
4	2.74	0.15164	2	3.8	0.16152
1	3.74	0.08884	1	3.74	0.08884

Table 4: the traditional Poisson Distribution and the Proposed Fuzzy Poisson Distribution with HIV Infections in Iraq

58



Figure 8: The Traditional Poisson Distribution and the Proposed Fuzzy Poisson Distribution with HIV Infections in Iraq

Table (1) and Figure (8) show a comparison between the traditional Poisson distribution and the fuzzy distribution for estimating the probability of HIV infection. In the traditional distribution, the calculations are based on fixed values for the infection rate (λ) and the expected number of infections (x), which leads to calculating the probabilities accurately but without taking into account any ambiguity or uncertainty in the data. In the fuzzy distribution, the vague or uncertain data are dealt with, as the infection rate and the number of infections is in a range of possible values instead of fixed values, which provides flexibility in predicting the probabilities. Through the results that have been reached, we find that when the value of the traditional observation is equal to (2) and the distribution parameter is equal to (3.670), the probability of infection is (0.172), which is the highest probability of infection. When the traditional observation is equal to (2.7), we find that the probability of infection is (0.0014), which is the lowest probability of infection. While when we take into account the fuzzy parameter for the observation and the parameter at the same time, the highest probability of infection is (0.176) and the lowest probability is (0.032). This indicates that the fuzzy distribution gives more stability in estimating rare incidents.

12. Conclusions

The research dealt with a proposed method that both the measured observations and the Poisson distribution parameter should be fuzzy to deal with imprecise and ambiguous data. Monte-Carlo simulation experiments were used to test the proposed distribution and compare it with the traditional Poisson distribution. Also, a real data set was used on HIV/AIDS documented in Iraq, which is a rare occurrence. The proposed distribution proved its suitability for this data and gave more accurate and flexible results in dealing with such data.

13. References

- Abboudi, E. H., Al-Mashhadani, A. K., & Salman, M. S. (2020). Estimation of the fuzzy reliability function using twoparameter exponential distribution as prior distribution. *Periodicals of Engineering and Natural Sciences (PEN)*, 8(2), 613-625.
- Akman, O., Arriola, L., Schroeder, R., & Ghosh, A. (2023). Quantum Mechanics for Population Dynamics. *Letters in Biomathematics*, *10*(1), 105–115.
- Ali, B. K., & Neamah, M. W. (2022). A new robust fuzzy informative standard Bayes estimator for exponential distribution. *Journal of Algebraic Statistics*, 13(1), 431-442.
- Barros, L. C. d., Bassanezi, R. C., & Lodwick, W. A. (2017). A first course in fuzzy logic, fuzzy dynamical systems, and biomathematics: theory and applications. (*No Title*).
- Burak Parlak, I., & Cağrı Tolga, A. (2016). Fuzzy Probability Theory I: Discrete Case. Fuzzy Statistical Decision-Making: Theory and Applications, 13-34.

Chaira, T. (2019). Fuzzy set and its extension: The intuitionistic fuzzy set. John Wiley & Sons.

EI-Dawoody, M., Eliwa, M. S., & EI-Morshedy, M. (2023). An Extension of the Poisson Distribution: Features and Application for Medical Data Modeling. *Processes*, 11(4), 1195.

59

- Garg, H., Sharma, S., & Rani, M. (2013). Weibull fuzzy probability distribution for analysing the behaviour of pulping unit in a paper industry. *International Journal of Industrial and Systems Engineering*, *14*(4), 395-413.
- Jamkhaneh, E. B., & Gildeh, B. S. (2011). Chain sampling plan using fuzzy probability theory. *Journal of applied sciences*, 11(24), 3830-3838.
- Monfared, M., & Yang, J.-B. (2004). Design of an intelligent manufacturing scheduling and control system using fuzzy logic: sensitivity analysis and parameter optimization. *Journal of Intelligent & Fuzzy Systems*, 15(2), 89-104.
- Moraes, R. M., & Machado, L. S. (2015). A fuzzy poisson naive bayes classifier for epidemiological purposes. 2015 7th International Joint Conference on Computational Intelligence (IJCCI),
- Neamah, M. W., & Ali, B. K. (2020). Fuzzy reliability estimation for Frechet distribution by using simulation. *Periodicals of Engineering and Natural Sciences (PEN), 8*(2), 632-646.
- Shebib, H., Jaafar, Z., & Ali, B. (2023). Choose Best Formula for Lindley Distribution for Modeling of Rainfall Data in Iraq in 2020. Journal of Al-Rafidain University College For Sciences (Print ISSN: 1681-6870, Online ISSN: 2790-2293)(1), 547-556.
- Sirbiladze, G., Kacprzyk, J., Manjafarashvili, T., Midodashvili, B., & Matsaberidze, B. (2022). New fuzzy extensions on binomial distribution. Axioms, 11(5), 220.
- Sivanandam, S., Sumathi, S., & Deepa, S. (2007). Introduction to fuzzy logic using MATLAB. Springer.
- Sun, L., & Choi, E. (2023). Substrate Transport in Cylindrical Multi-Capillary Beds with Axial Diffusion. *Letters in Biomathematics*, 10(1), 63–74.
- Yahia, N. B., Bellamine, N., & Ghezala, H. B. (2012). Integrating fuzzy case-based reasoning and particle swarm optimization to support decision making. *International Journal of Computer Science Issues (IJCSI)*, 9(3), 117.
- Zhao, J., Zhang, F., Zhao, C., Wu, G., Wang, H., & Cao, X. (2020). The properties and application of poisson distribution. Journal of Physics: Conference Series,