Proceedings of ICEMGD 2025 Symposium: The 4th International Conference on Applied Economics and Policy Studies
DOI: 10.54254/2754-1169/2025.BJ24761

# *Poisson Distribution and Its Applications*

**Yan Zhang**

*The University of Nottingham Ningbo China, Ningbo, China*
*zy080805@163.com*

***Abstract.*** Poisson distribution, a foundational discrete probability distribution, describes the probability of observing k independent events within a fixed temporal or spatial interval under conditions of constant mean occurrence rate (λ). Therefore, it can be used to address a variety of real-world challenges across diverse domains, particularly in the application of low-probability phenomena. This paper investigates the application of the Poisson distribution through rigorous theoretical analysis and case studies specifically spanning on two fields, which are healthcare resource allocation and communication network optimization. Although the Poisson distribution has inherent limitations in some phenomena, our study reveals that as long as the model is optimized based on statistical principles and specific domain contexts, this framework can effectively provide valuable and actionable insights to inform data-driven decision-making across diverse domains ranging from healthcare resource allocation to telecommunications network optimization in complex systems.

***Keywords:*** Poisson distribution, application, limitation, optimization

## 1. Introduction

The Poisson distribution, a discrete probability model describing the occurrence of independent events within fixed intervals, has been widely applied across disciplines since its formulation in 1837 [1]. Defined by the probability mass function:

$$P\left(X = k\right) = \frac{\lambda^k e^{-\lambda}}{k!} \tag{1}$$

it relies on core assumptions of event independence, a constant rate $\lambda$ and negligible simultaneous occurrences. While classical studies validate its utility in static systems—such as defect detection in manufacturing or routine patient admissions in healthcare—modern applications increasingly confront dynamic environments where $\lambda$ fluctuates temporally or spatially (e.g. urban traffic peaks, pandemic-driven medical surges). Despite its mathematical elegance, critical gaps persist in understanding its adaptability to non-stationary conditions, cross-domain performance variations, and real-world violations of independence assumptions. This study addresses these limitations by systematically evaluating Poisson models through theoretical analysis and empirical case studies spanning medical resource allocation, disaster relief resource allocation and communication network optimization. Through empirical interrogation of multimodal datasets (bike-sharing flows, healthcare admissions, manufacturing throughput), this study not only demonstrates the boundaries and constraints of Poisson distributions but also proposes adaptive optimizing strategies (e.g., real-time $\lambda(t)$ adjustments for

DOI: 10.54254/2754-1169/2025.BJ24761

traffic management) and validates domain-specific guidelines for over-dispersed data or time-dependent $\lambda$ , this work bridges theoretical rigor with practical relevance. It advances stochastic modeling as a decision-making tool for urban planners, healthcare administrators, and industrial engineers, while highlighting future research directions in event-dependent systems and machine learning-integrated Poisson extensions[2].

## 2. Theoretical foundations of the poisson distribution

### 2.1. Definition and formula

The Poisson distribution is a discrete probability distribution that models the number of independent random events occurring within a fixed unit of time or space[3]. Its probability mass function (PMF) is defined as:

$$p = \left(X = k\right) = \frac{e^{-\lambda}\lambda^{k}}{k!},\ k = 0,1,2 \tag{2}$$

where $\lambda$ represents the average event rate (expected number of occurrences), $e$ is the base of the natural logarithm (~2.71828), and $k!$ denotes the factorial of $k$ . A key characteristic of the Poisson distribution is that its mean and variance are equal ( $E(x) = Var(X) = \lambda$ ), making it uniquely suited for modeling "equidispersed" count data[4]. At the same time, the relationship between the Poisson distribution and Poisson distribution is:

$$Bin(n, p) \approx Po(np) \tag{3}$$

### 2.2. Applicability conditions

The validity of the Poisson distribution relies on three core assumptions:

Independence: Each event occurs independently of the others.

Constant Rate: Events occur at a constant average rate ( $\lambda$ ) over the observed interval.

Rarity: The probability of two or more events occurring in an infinitesimally small sub-interval is small.

Violations of these assumptions—such as time-varying  or interdependence between events—can significantly compromise the accuracy of the model.

## 3. Typical application scenarios of the poisson distribution

### 3.1. Healthcare: medical resource allocation

The Poisson distribution plays an important role in optimizing healthcare resource allocation, particularly in dynamic environments such as emergency departments , surgical units, and hospital bed management. By modeling patient arrival rates and service demands, healthcare administrators can reduce the occurrence of overcrowding, shorten patient waiting time, and allocate staff and equipment more effectively.

The Poisson distribution is often used to model the occurence of rare and independent events within a fixed time interval. Under ideal conditions, the number of patients arriving in a fixed interval $t$ can denoted as $X(t)$, while the average rate of arriving is labeled as $\lambda$. Then the probability of $k$ patients arriving in $t$ is:

DOI: 10.54254/2754-1169/2025.BJ24761

$$p\left(X\left(t\right) = k\right) = \frac{e^{-\lambda}(\lambda t)^{k}}{k!} \tag{4}$$

To apply this pratically, hospitals typically record daily patient volumes, staff schedules and resource requirements (e.g., numbers of beds, ventilators, and medications). For instance, during non-pandemic periods, a hospital emergency departments may observe an average of 5 patient arrivals per hour ( $\lambda = 5$ ) . The probability of patient arrivals exceeding the department's capacity (e.g. $P(x > 10)$ ) can be computed so that the additional staff can be deployed preemptively.

A case study analyzing the length of hospital stay for 3,589 cardiovascular patients applied Poisson regression to evaluate which variables that are related to the length of hospital stay, including the type of procedure ((CABG vs. PTCA), sex (male vs. female), admission type (urgent vs. elective), and age (>75 vs. ≤75). Let $\mu_i$ denote the expected mean value of the hospital length of stays for $i$ patients, which is modeled through an exponential link function:

$$\mu_i = \exp(\gamma_0 + \gamma_1 x_{1i} + \gamma_2 x_{2i} + \gamma_3 x_{3i} + \gamma_4 x_{4i}) \tag{5}$$

where $\gamma_0$ labeled as the intercept which is the baseline log-rate of hospital length of stay when all co-variates are at their reference levels., and $\gamma_1$ , $\gamma_2$ , $\gamma_3$ and $\gamma_4$ represent the regression coefficients for the co-variates $x_{1i}$ (CABG procedure), $x_{2i}$ (male sex), $x_{3i}$ (urgent admission) and $x_{4i}$ (age >75 years). The exponential function ensures $\mu_i > 0$ , aligning with the Poisson distribution's requirement for positive count-based outcomes. Length of hospital stay is treated as a count variable, where the Poisson model assumes the variance equals the mean. However, the initial analysis revealed significant over dispersion (dispersion index = 5.432), violating the Poisson assumption of equal mean and variance. Despite the over-dispersion, Poisson regression identified CABG procedures( $\gamma_1 = -0.96$ ), urgent admissions( $\gamma_3 = 0.33$ ), and older age ( $\gamma_4 = 0.12$ ) as significant factors (all $p < 0.001$ ). But model fit metrics (AIC = 22389.8; BIC = 22420.7) were inferior to alternative models. A negative binomial (NB) regression and a non-parametric generalized linear (NPGL) model were compared, with NPGL achieving superior fit (AIC = 21,138.5) and better residual diagnostics.

This case demonstrates that while Poisson regression a useful baseline for analyzing count data in healthcare, its limitations, especially under over-dispersion, necessitate the use of more flexible models such as NB or NPGL for accurate inference [5].

### 3.2. Communication network optimization

The Poisson distribution also finds practical application in the optimization of communication networks. In modern wireless communication systems, cellular networks represent one of the foundational architectures for mobile connectivity. These networks divide geographic regions into infinite "cells" (often modeled as hexagons), each served by at least one base station such as 4G/5G towers, enabling seamless mobility for users [6]. The primary objective of cellular networks is to efficiently manage real-time, user-driven interactions, such as voice calls, video conferencing and instant messaging by dynamically allocating resources like frequency bands and transmission power. The inherent randomness and independence of initiated requests by users align naturally with the statistical assumptions of the Poisson distribution, which models events occurring at a constant average rate with no dependence.

For example, consider a practical scenario involving voice call requests in a 5G cellular network, there is a base station in a densely populated metropolitan area handles an average of $\lambda = 500$ call requests per minute during peak hours. Assuming these requests are independent and occur at a stable rate, the number of calls per minute, denotes as $X(t)$ , follows a Poisson distribution:

DOI: 10.54254/2754-1169/2025.BJ24761

$$p\left(X\left(t\right)=k\right)=\frac{e^{-500}500^{k}}{k!},\ \ k=0,1,2... \tag{6}$$

By calculating the probability of exceeding the capacity of base station, network operators can assess the risk of service disruption. If the probability exceeds than 1%, staff will take action immediately, such as deploying temporary small cell stations or rerouting traffic to adjacent towers. This model not only quantifies the distribution of random requests in cellular networks but also provides actionable insights for resource allocation, demonstrating the practical value of Poisson modeling in managing high priority, real-time traffic.

In cellular networks, background traffic, such as software updates, push notifications or sensor data, often exhibits characteristics aligned with Poisson assumptions. Each user equipment generates flow independently. The time between consecutive signals arriving corresponds to the hallmark of Poisson distribution, especially for the time that the traffic arrival rate per user equipment is low.

A study comparing Poisson approximation and Gaussin approximations found that while the Gaussin model performs well under under heavy traffic, it is inaccurate under light traffic, which is more common in many real-world scenarios. Therefore, the Poisson distribution is more accurate than the Gaussin distribution. At the same time, with the appearance of the small cell networks, there will be fewer user equipment on a base station, which means the traffic will be discrete and sparse so that the modified Poisson distribution will be an accurate approximation in current and future cellular networks [7].

Certainly, there are still some limitations of the Poisson distribution. It is difficult to cope with unexpected situations and changes of behaviors of users, which enables the approximation to be inaccurate. To address this problem, additional parameters can be introduced into the modified Poisson distribution to adjust shape, scale and variance to enhance the flexibility of Poisson distribution while retaining its core framework.

## 4. Limitation of the poisson distribution

Although the Poisson distribution and Poisson process are widely used in modeling events across diverse fields like epidemiology and telecommunications, their effectiveness hinges on several strict assumptions that may not hold true in the real world. This section will focus on three main limitations of the Poisson distribution, which are dependent events, time varying rate and high frequency events.

The Poisson model assumes that events occur independently within a fixed time or space interval. For example, its application in healthcare resource allocation relies on the assumption that patient arrivals are statistically independent. However, this assumption is completely disproved with the outbreaks of infectious diseases because the probability of new infections increases rapidly, which directly contradicts to the Poisson independence axiom.

For instance, during the COVID-19 pandemic, transmission chains caused the fail of Poisson models to capturing the temporal clustering of cases. A study that used Poisson models for COVID-19 projections underestimated peak case loads by 40–60% during early 2020, which is a failure of application of Poisson model [8]. In such cases, self-exciting processes such as the Hawkes process—an extension of an inhomogeneous Poisson process—provide better alternatives [9].

Furthermore, the Standard Poisson processes assume a constant event rate $\lambda$, which is unrealistic in systems with periodic or trending intensity changes such like the traffic accidents during rush hours, emergency calls at night and the website visits during promotional periods all exhibit that the value of $\lambda(t)$ is time-dependent. This situation can may be addressed by the inhomogeneous Poisson process, which allows $\lambda$ to vary with the function of time. For example, modeling highway accidents might use

DOI: 10.54254/2754-1169/2025.BJ24761

a piecewise $\lambda(t)$ with peaks at 8 AM and 6 PM. Mathematically, the expected number of events in interval $[t_1, t_2]$ becomes:

$$\Lambda\left(t_1, t_2\right) = \int_{t_1}^{t_2} \lambda(t)\, dt \tag{7}$$

Nevertheless, the inhomogeneous Poisson process requires precise estimation of the rate function $\lambda(t)$, which means it demands larger datasets. Advanced implementations combine inhomogeneous Poisson process with machine learning to predict $\lambda(t)$ from covariates like weather or holidays [10].

Additionally, if events occur with extreme frequency, simple Poisson models will become inadequate. For example, insurance claims during natural disasters have different financial impacts for different groups of people, thousands of claims are high frequency. This is an extreme case so that the Poisson distribution is not well suited. In this situation, compound Poisson processes might address this by introducing a random variable $Y_i$ representing the magnitude of each event. The total loss becomes:

$$S\left(t\right) = \sum_{i=1}^{N(t)} Y_i \tag{8}$$

where $N(t)$ (the number of events occur) still follows a Poisson process. This framework is also fundamental in risk management. In real world, for extremely high-frequency systems such as the stock trades for every millisecond, continuous approximations like diffusion models may be more appropriate for them.

## 5. Conclusion

As a cornerstone of probability theory, the Poisson distribution has shown extraordinary capability in prediction and estimation across scientific and engineering disciplines. It is able to model rare events through a surprisingly simple formula and only requires the average event rate ( $\lambda$ ) as its parameter. The Poisson model reduce complex stochastic process to the quantification of unit measure of events, which enables it to be indispensable in various fields, such as quantum physics and modern epidemiology. During the application, there are still some limitations of the use of the Poisson model, because the applying scenario has to rigorous comply with three fundamental assumptions, which are event independence, constant occurrence rate and countability of events.

These prerequisites may cause challenges in practical implementations. Therefore, the evolution of Poisson modeling is necessary. When confronted with over dispersed data where variance exceeds mean ( $Var(x) > E(x)$ ), researchers often adopt negative binomial distributions which effectively is a Poisson-gamma mixture , accounts for rate heterogeneity. In actuarial science, compound Poisson processes combine event frequency (Poisson) with severity distributions to model total claim amounts. Additionally, the Bayesian revolution can further expand Poisson's horizons through hierarchical modeling, where prior distributions update dynamically with incoming data streams. After the evolution of the Poisson model, the Poisson distribution has great potential in cross-disciplinary application. Telecommunications engineers employ it to design cellular networks, ecologists use it to estimate species distribution and financial analysts leverage it for credit risk assessment. Ultimately, whether the Poisson distribution will play a bigger role in the future depends on the ability to adapt its elegant formalism to complexities of future technology while preserving the statistical rigor that has made it endure.

DOI: 10.54254/2754-1169/2025.BJ24761

## References

[1] Poisson, Siméon D. (1837). Probabilité des jugements en matière criminelle et en matière civile, précédées des règles générales du calcul des probabilités [Research on the Probability of Judgments in Criminal and Civil Matters] (in French). Paris, France: Bachelier.
[2] Rasch, Georg (1963). The Poisson Process as a Model for a Diversity of Behavioural Phenomena 17th International Congress of Psychology. Vol. 2. Washington, DC: American Psychological Association.
[3] Haight, Frank A. (1967). Handbook of the Poisson Distribution. New York, NY, US: John Wiley & Sons.
[4] Bektashi, X. , Rexhepi, S. , & Limani-Bektashi, N. (2022). Dispersion of count data: A case study of Poisson distribution and its limitations. Asian Journal of Probability and Statistics, 19(2), 18–28.
[5] Altun, E. 2021. A new two-parameter discrete Poisson-generalized Lindley distribution with properties and applications to healthcare data sets. Computational Statistics 36 (4):2841–2861. doi:10. 1007/s00180-021-01097-0.
[6] EITC. "Cellular Networks, Cells, and Base Stations — EITC". Retrieved 22 November 2024.
[7] Zhang, S. , Zhao, Z. , Guan, H. , & Yang, H. (2017). A modified poisson distribution for smartphone background traffic in cellular networks. International Journal of Communication Systems, 30(6), e3117.
[8] Endo, A. , et al. (2020). Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. Wellcome Open Research, 5(67)
[9] Chiang W. -H. , Liu X. (2022). Mohler G. H awkes process modeling of COVID-19 with mobility leading indicators and spatial covariates. Int J Forecast, 38(2), 505-520.
[10] Cook, R. J. , & Lawless, J. F. (2007). The Statistical Analysis of Recurrent Events. Springer.