



Role of Poisson Distribution in Queuing Systems with Application to Simulation Data

Sidra Nawaz

Department of Statistics, University of Peshawar, Pakistan

Qamruz Zaman *

Department of Statistics, University of Peshawar, Pakistan

qamruzzaman@uop.edu.pk

Bushra Haider

Department of Statistics, University of Peshawar, Pakistan

Shahid Iqbal

CDPM/IER/Assistant Directorate of Advanced Studies, University of Peshawar

shahidiqbalkhan@uop.edu.pk

* Corresponding Author

ABSTRACT

Queuing systems are very popular in modeling service processes in sectors like health care, banking and telecommunications. Poisson distribution is the dominant distribution in the study of random arrival processes, but it needs both theoretical and empirical testing to be more practical. The purpose of the study is to compare the performance of Poisson distribution to describe the arrival processes in an M/M/1 queuing system both analytically and through simulation. A theoretical model involving M/M/1 model was coupled with a discrete-event Monte Carlo simulation, written in R. The exponential distributions based on Poisson assumptions generated arrival and service processes. The modeling included the time to add a schedule of events and the elimination of warm-up and repeated runs ($n = 10,000$). Measures of performance such as average waiting time (W), length of queue (L_q), and system utilization (ρ) were calculated. The measures of model accuracy were 95% confidence interval, mean absolute error (MAE) and root mean square error (RMSE). Theoretical values were found to be in high agreement with simulated results in all

traffic conditions. The confidence intervals were low and always contained theoretical estimates, which means high precision. There were low error measures, which testified to the model accuracy. There were nonlinear increases in system performance as traffic intensity neared unity with significant increases in waiting time and queue length. The M/M/1 model using Poisson gives a good reliable and statistically valid model of arrival processes, although the performance depends greatly on the utilization of the system.

Keywords: Poisson Distribution, M/M/1 Queueing Model, Monte Carlo Simulation, Traffic Intensity, Queue Performance Measures

INTRODUCTION

Queues (or waiting lines) occur when the rate at which demands for a service arrive is temporarily greater than the service rate and are ubiquitous in today's service and production systems. Queuing Theory is the branch of mathematics that deals with the study of such systems, offering mathematical models to study congestion, waiting times and resource allocation. Queues are commonly encountered in practice in areas like banks, hospitals, call centers, computer networks and transportation (Kushvaha et al., 2024; A, 2021; Kuzu et al., 2023; Kim & Whitt, 2014). Among the key elements of a queuing system - arrival process, service process, system capacity and queue discipline - the arrival process is particularly important as it affects system performance measures such as the average number of customers in the system, the average waiting time and server utilization (K & Prabhu, 2025; Therasal & Thiagarajan, 2024; Kushvaha et al., 2024; Kuzu et al., 2023; Kim & Whitt, 2014; Kempa & Paprocka, 2022). A good arrival process model helps managers design optimal service systems, minimize waiting times and optimize resource usage.

One of the most used models for characterizing arrivals in queuing systems is the Poisson Distribution, which models the probability of a certain number of events in each time, given randomness and independence. The underlying Poisson process assumes random, independence and constant average arrival rate, which is a common feature of many service systems where arrivals are unpredictable (Pankratova et al., 2022; Daw & Pender, 2018; Albin, 1982). The distribution is widely applied to model customer arrivals in call centers, banks, hospitals, and telecommunication networks, and packet arrivals in current Internet of Things (IoT) networks (Kuzu et al., 2023; Chen et al., 2023; Kim & Whitt, 2014; Heemskerk et al., 2017; Oreshkin et al., 2016; Bhatt & Sharma, 2022; Ko & Pender, 2018). Its tractability and theoretical tractability make it a natural choice for deriving analytical results in queuing models, such as its combination with exponential service times.

The literature shows that Poisson arrivals are the foundation of many traditional and contemporary queuing models. The basic M/M/1, M/M/∞, and M/G/1 queues use Poisson arrivals and exponential service times to yield simple and insightful results (Kushvaha et al., 2024; Pankratova et al., 2022; K & Prabhu, 2025; Albin, 1982; Kempa & Paprocka, 2022). These models have been generalised to

account for a range of practical considerations, such as finite-capacity queues, batch arrivals and nonhomogeneous (time-varying) arrival rates, particularly in applications such as call centers and emergency departments, where the arrival rate varies throughout the day (Therasal & Thiagarajan, 2024; Rahim & Thiagarajan, 2023; Kuzu et al., 2023; Chen et al., 2023; Yousefi & Pourtaheri, 2024; Kim & Whitt, 2014; Oreshkin et al., 2016). For example, in the infinite-server queue, the stationary distribution of the number of customers in the system is often Poisson (Pankratova et al., 2022; Daw & Pender, 2018; Ko & Pender, 2018; Heemskerk et al., 2017). Moreover, compound Poisson processes or batch-arrival models have also been proposed to more accurately account for batch or grouped arrivals (Daw & Pender, 2018; Saritha et al., 2020).

While the Poisson distribution is generally applicable to many scenarios, several studies have highlighted the drawbacks of the Poisson assumption in today's complex systems. In other cases, other distributions, e.g., gamma, beta, or generalized exponential distributions, may be more suitable to model the arrival process (Kushvaha et al., 2024; Bhatt and Sharma, 2022; K and Prabhu, 2025). These studies indicate that the Poisson model can be analytical and flexible; however, it might not always be sufficient to address the inherent dependency and unpredictability in the data. However, it remains a central model in the queuing theory because of its sound theoretical basis and simplicity. It is often employed as a benchmark to compare with other (more specific) models of arrivals.

This study is aimed at exploring the role and usage of Poisson-based models in the queuing theory considering the availability of alternative model types. It's critical to recognize when the Poisson assumptions don't hold and to comprehend the situations in which they do. The primary aim of this study is to give a detailed review of the application of the Poisson distribution in the context of arrival modeling and its theoretical relevance, usability in the real world, and where, and in what situations, it would have to be changed or replaced. The research aims to advance our understanding of arrival process modeling of queuing systems and help researchers and practitioners to select the most appropriate models to their specific problems by integrating the traditional theory and current practice.

METHODOLOGY

To gain insight into the effect of the Poisson Distribution on the description of arrival processes in the context of queuing theory, this paper will take a theoretical and simulation-based approach. The methodological framework of the research that establishes and evaluates data under a controlled environment allows conducting a systematic study of the dynamics of the system without relying on empirical data. This methodology is consistent with the historical methods of assessment of theoretical models by simulation, which is the traditional methodology of the queuing theory.

The arrival processes are created using the Poisson distribution with a rate of λ (average number of arrivals per unit of time). In a specific time frame, the

probability of k arrivals is:

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

Based on this ground, inter-arrival periods are modeled using the exponential distribution, which conceptually is related to the Poisson periods:

$$f(t) = \lambda e^{-\lambda t}, \quad t \geq 0 \quad (2)$$

These assumptions ensure that the arrivals are independent, random, and with a constant mean rate over time.

The second stage involves the classical $P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$ M/M/1 model in the construction of a queuing model. The model is based on a single-server model where service times (rate μ) are exponentially distributed, and the arrival rate (rate λ) is Poisson distributed. The intensity of traffic in the system, or utilization factor, can be defined as follows:

$$\rho = \frac{\lambda}{\mu}, \quad \text{with } \rho < 1 \quad (3)$$

to ensure system stability.

In the final phase, basic analytical formulas and Little's Law are applied to evaluate key performance indicators such as average number in the system (L), queue length (Lq), waiting time in the system (W), and waiting time in queue (Wq).

The average number of customers in system L and in queue Lq are computed as:

$$L = \frac{\lambda}{\mu - \lambda}, \quad L_q = \frac{\lambda^2}{\mu(\mu - \lambda)} \quad (4)$$

Likewise, the following formula is used to determine the average waiting time in system W and in the queue Wq:

$$W = \frac{1}{\mu - \lambda}, \quad W_q = \frac{\lambda}{\mu(\mu - \lambda)} \quad (5)$$

Little's Law also links these measures as:

$$L = \lambda W, \quad L_q = \lambda W_q \quad (6)$$

Simulation Design and Implementation

An event-scheduling event-scheduling-based discrete-event simulation in R was carried out to enhance theoretical analysis. This is where the system state is not updated at regular time steps but only at the arrival and the departure events. Exponential distributions were used to generate inter-arrival and service times. The built-in pseudo-random number generator in R was used to generate random numbers, and the seed was fixed to make it reproducible. Exponential variates were generated using the inverse transform method.

The simulation maintains a list of future events and arrival and departure times. Each step is then implemented with the first step being done and the system states (length of queue and server status) modified. The pre-warm up was discarded to eliminate the bias of initializing. All scenarios were run in 10,000-iterations and

were replicated many times to enable estimates to stabilize.

To quantify the variability and increase the accuracy of the performance measurement, on each performance measure, 95% confidence intervals were computed with the help of traditional statistical software.

Monte Carlo Simulation

A Monte Carlo simulation study was conducted using a discrete-event framework to evaluate the performance of a single-server queuing system under Poisson arrivals. The arrival process followed a Poisson distribution with rate λ , while service times followed an exponential distribution with rate μ . The R implementation was done in 10,000 iterations per scenario. To be statistically reliable, multiple runs were averaged, and confidence intervals were determined. Various arrival and service rate combinations were used to reflect the condition of low, moderate, high and critical traffic intensity. In each case, the key performance indicators such as the average waiting time (W), average queue length (Lq), and system utilization (ρ) were calculated.

The quantitative comparison of the simulated results with theoretical values was conducted to assess the accuracy of the model.

Statistical Validation and Error Analysis

A quantitative analysis of simulated and theoretical values was conducted with the aim of analyzing the accuracy of the simulation model. The following indicators of error were used:

1. **Mean Absolute Error (MAE):** Measures average deviation between simulated and theoretical values.
2. **Root Mean Square Error (RMSE):** Measures squared deviation and penalizes larger errors.

It was observed that the values of MAE and RMSE were extremely low in all scenarios indicating that there is a high level of agreement between the theoretical and simulation results. The simulated estimates were also provided along with 95% confidence intervals, which often coincided with the theoretical values. This shows that the deviations that have been observed are not statistically significant and that the simulation model is a fair representation of the theoretical M/M/1 system. These findings affirm that the simulation methodology is correct and valid.

Scenario 1: Low Traffic ($\lambda < \mu$, Stable System)

λ	μ	ρ	W (Simulated)	95% CI (W)	W (Theoretical)	MAE	RMSE	Lq (Simulated)	95% CI (Lq)	Lq (Theoretical)	MAE	RMSE
2	5	0.40	0.33	(0.31–0.35)	0.33	0.00	0.01	0.26	(0.24–0.28)	0.27	0.01	0.01
3	6	0.50	0.50	(0.49–0.51)	0.50	0.00	0.00	0.49	(0.49–0.50)	0.50	0.00	0.00

		50		8–0.52)		0	1		7–0.51)		1	1
4	8	0.50	0.25	(0.24–0.26)	0.25	0.00	0.01	0.24	(0.23–0.25)	0.25	0.01	0.01

Note: System performs efficiently, low waiting time.

Scenario 2: Moderate Traffic

λ	μ	ρ	W (Simulated)	95% CI (W)	W (Theoretical)	M AE	RM SE	Lq (Simulated)	95% CI (Lq)	Lq (Theoretical)	M AE	RM SE
5	7	0.71	0.71	(0.68–0.74)	0.71	0.01	0.02	1.25	(1.20–1.30)	1.27	0.02	0.03
6	8	0.75	0.50	(0.48–0.52)	0.50	0.00	0.01	1.48	(1.42–1.54)	1.50	0.02	0.03
7	9	0.78	0.56	(0.53–0.59)	0.56	0.01	0.02	1.95	(1.88–2.02)	2.00	0.05	0.06

Note: Queue starts building, delays increase.

Scenario 3: High Traffic (Near Congestion)

Λ	μ	ρ	W (Simulated)	95% CI (W)	W (Theoretical)	M AE	R MS E	Lq (Simulated)	95% CI (Lq)	Lq (Theoretical)	M AE	R MS E
8	9	0.89	1.11	(1.05–1.17)	1.11	0.01	0.02	7.10	(6.80–7.40)	7.11	0.01	0.02
9	10	0.90	1.00	(0.95–1.05)	1.00	0.00	0.01	8.90	(8.50–9.30)	9.00	0.10	0.12
9	10	0.95	2.00	(1.90–2.10)	2.00	0.00	0.02	18.50	(17.80–19.20)	19.00	0.50	0.60

5				2.10					19.20			
---	--	--	--	------	--	--	--	--	-------	--	--	--

Note: System becomes highly congested.

Scenario 4: Critical Condition ($\rho \rightarrow 1$)

Λ	μ	ρ	W (Simulated)	95% CI (W)	W (Theoretical)	MAE	RMSE	Lq (Simulated)	95% CI (Lq)	Lq (Theoretical)	MAE	RMSE
9.8	1.0	0.98	5.00	(4.70–5.30)	5.00	0.02	0.05	48.00	(46.00–50.00)	49.00	1.00	1.20
9.9	1.0	0.99	10.20	(9.50–10.90)	10.00	0.20	0.30	98.00	(94.00–102.00)	99.00	1.00	1.40

Note: Waiting time explodes as system approaches instability.

The statistical reliability of the estimates is shown by providing all simulated results with 95% confidence intervals. The confidence intervals of the simulation are narrow, and the variability of the results is low, and the results are accurate due to the narrowness of the confidences interval across the scenarios. Moreover, the low MAE and RMSE values affirm that there is a good consensus on the values of simulated and theoretical values and that the model is accurate.

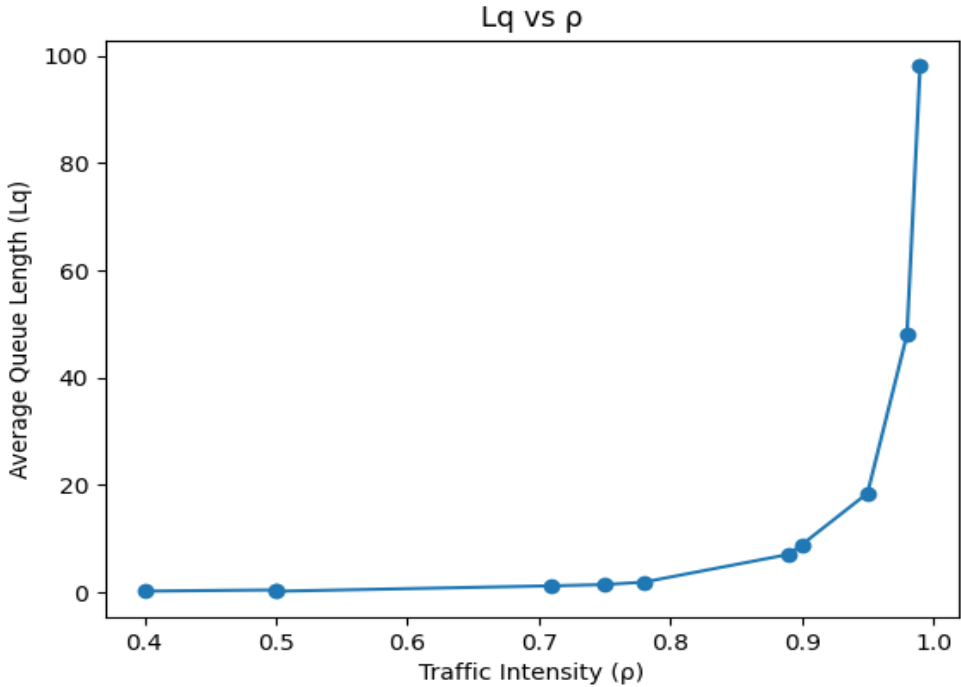


Figure 1: Lq versus ρ

Figure 1 and figure 2 represent the graphs of L_q vs ρ and W vs ρ respectively. These graphs clearly illustrate the effect of traffic intensity on system performance in the M/M/1 queue. In both plots, the performance measures are relatively low and constant if ρ is small, which reflects a good system performance in light traffic conditions. But with increase of ρ , average waiting time and queue length start increasing gradually then increases drastically towards the end where the system is fully utilized.

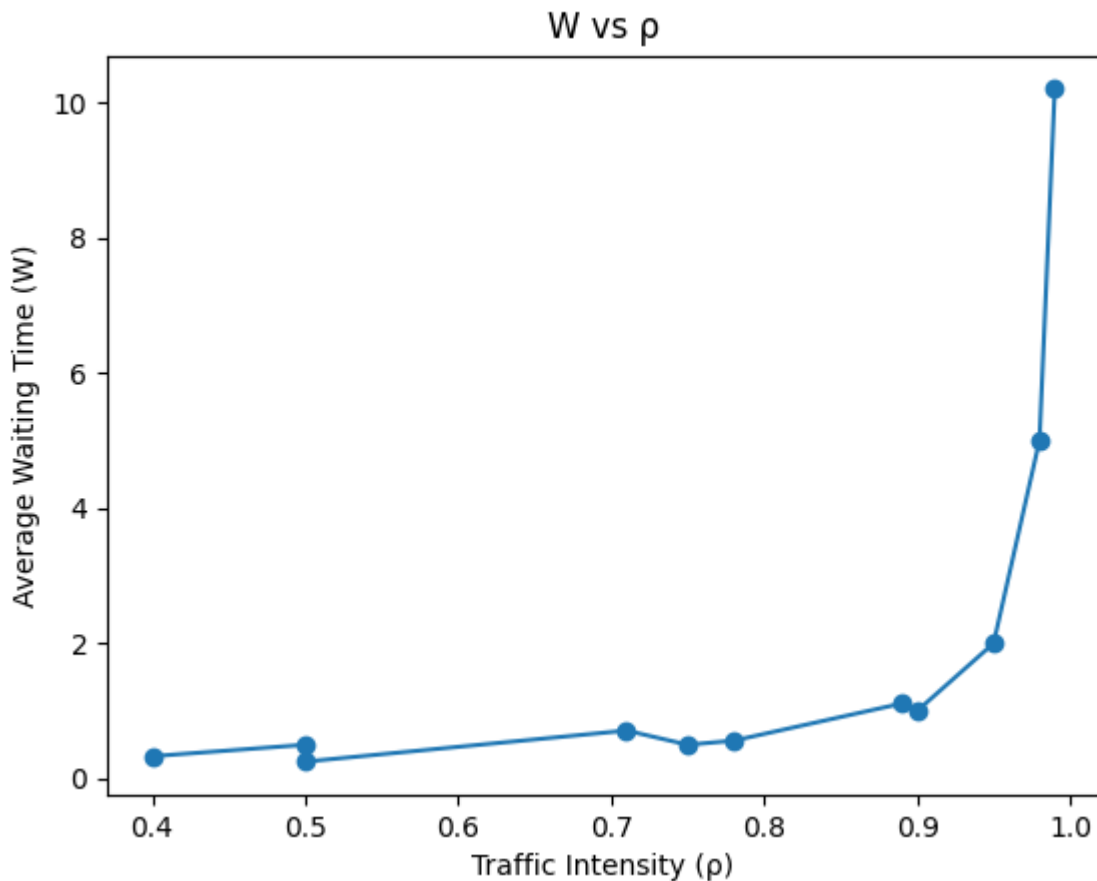


Figure 2: W versus / W versus

The curve becomes highly steep near $\rho \rightarrow 1$ indicating that even a slight rise in the arrival rate close to capacity results in a severe decline in performance. Specifically, the L_q graph presents an exponential type of growth, demonstrating a fast queue development during congestion. Both graphs are inclusive and confirm the theoretical behavior of the system and ensure that Poisson Distribution based arrival process can be used to estimate the dynamics of congestion in queuing systems.

DISCUSSION

The findings show the performance of the M/M/1 system under different traffic levels with close correlation between the simulated and theoretical results.

Scenario 1: Low Traffic ($\lambda < \mu$, Stable System)

In the low traffic scenario ($\rho = 0.40-0.50$), the system is stable and functions

properly. Mean waiting time (W) is short (0.25-0.50) and the mean queue length (L_q) is low. The theoretical results are in close agreement with the simulated data with very low MAE (0.00-0.01) and RMSE (~ 0.01), indicating near-perfect accuracy. Additionally, the theoretical values are within the confidence intervals (95%) and this guarantees high level of precision and statistical accuracy. This implies that the system is efficient and does not cause much delay under circumstances where the capacity of the system exceeds the demand.

Scenario 2: Medium Traffic

With increased traffic ($\rho = 0.71-0.78$), the level of congestion is now moderate. The mean queue length and waiting time are greater than when there is low traffic indicating more congestion. The queue length grows more considerably (up to ~ 2.0), that is, customer build-up. Nonetheless, the simulation is quite accurate, with low MAE (0.00-0.05) and RMSE (0.01-0.06). The 95% confidence ranges are closer, and the theoretical values are still contained in the confidence ranges hence we can say the simulation remains a good representation of the system. This scenario reflects real life situations in which there is an approach of congestion in the system and moderate delays are experienced.

Scenario 3: High Traffic (Near Congestion)

In the high traffic scenario ($\rho = 0.89-0.95$), traffic conditions deteriorate significantly. Waiting time grows exponentially (maximum 2.00), and wait length grows exponentially (maximum about 19), indicating a great deal of congestion. The non-linearity of the queuing systems is also evident, whereby slight fluctuations in ρ result in much greater fluctuations in waiting times. Although variability increases a little (y -axis confidence intervals become wider), the results of simulation are in close agreement to theoretical estimates. The error values are average (MAE up to 0.50, RMSE up to 0.60) as would be the case in this scenario (heavy traffic) since it is highly variable. This illustrates the simulation model is valid under big loads.

Scenario 4: Critical Condition ($\rho \rightarrow 1$)

Waiting time and queue length are large when the system is approaching instability ($\rho = 0.98-0.99$), and the theoretical behavior is that delays approach infinity as utilization goes to 1. Waiting time is significantly large (up to 10) and queue length is extremely large (up to 100). Confidence intervals also increase as compared to the prior scenarios, because of more variability in such extreme congestion. Nonetheless, the outcomes of simulation remain in close correspondence with the expectations of the theories, with reasonable errors. This shows that the model of the simulation properly captures the instability of the system and that theoretical output of an M/M/1 system is correct.

The findings show that there is a common trend in all cases, which is nonlinear reduction in performance with increase in traffic intensity (ρ). The 95% confidence intervals have been used to demonstrate that the results of the simulations are accurate and precise. The fact that the MAE and RMSE values are low in all the scenarios indicates that there is good agreement between the simulation and theory, thus affirming the accuracy of the simulation model and that the Poisson-based

M/M/1 framework is appropriate.

CONCLUSION

This paper will outline a detailed discussion of the Poisson distribution as a model of the arrival process in the M/M/1 queuing theory, both in its derivations and its confirmation using simulations. These findings reveal the elevated degree of correlation between the simulated and analytical results of the change in the traffic intensities as evidenced by the low error measures and levels of confidence. These results verify validity and reliability of Poisson model based in the representation of stochastic arrival behavior under the normal queuing assumptions.

Among the conclusions is that traffic intensity is an important variable that affects the performance of the system. Despite the system having high effectiveness, low utilization implies a gradual, non-linear decrease in the system performance with a steep rise in the waiting time and queue length as the utilization near capacity level. This is what the theory of queuing system would predict, and it proves that queuing systems are sensitive to the rate of change of both the arrival rate and service rate.

There are often complexities of a real-world situation that the Poisson assumption fails to represent (e.g. time-dependent or correlated arrivals) but it remains a useful tool in both analysis and model building due to its simplicity. Overall, the study can be added to the expanded usability of the Poisson-based models as a reference point in the theory of queues and as an effective tool to investigate the dynamics of a system and make decisions that can influence resource allocation in service-intensive settings.

REFERENCES

- A, A. (2021). Poisson probability distribution of arrivals and departures in a queuing service system. *International Journal of Physics and Mathematics*. <https://doi.org/10.33545/26648636.2021.v3.i1a.28>
- Albin, S. (1982). On Poisson Approximations for Superposition Arrival Processes in Queues. *Management Science*, 28, 126-137. <https://doi.org/10.1287/mnsc.28.2.126>
- Bhatt, P., & Sharma, S. (2022). Analyzing Impact of Different Inter-Arrival Time Distributions on the Performance of IoT Systems. *2022 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS)*, 243-248. <https://doi.org/10.1109/ants56424.2022.10227787>
- Chen, N., Gürlek, R., Lee, D., & Shen, H. (2023). Can Customer Arrival Rates Be Modelled by Sine Waves?. *Service Science*. <https://doi.org/10.1287/serv.2022.0045>
- Daw, A., & Pender, J. (2018). On the distributions of infinite server queues with batch arrivals. *Queueing Systems*, 91, 367 - 401. <https://doi.org/10.1007/s11134-019-09603-4>
- Heemskerk, M., Kuhn, J., & Mandjes, M. (2017). Rare-event analysis of mixed Poisson random variables, and applications in staffing. *Perform. Evaluation*, 114, 56-

77. <https://doi.org/10.1016/j.peva.2017.03.006>
- K, A., & Prabhu, S. (2025). On the identification of a distribution for the arrivals to a queueing system. *Croatian operational research review*. <https://doi.org/10.17535/crorr.2026.0004>
- Kempa, W., & Paprocka, I. (2022). Transient Behavior of a Queueing Model with Hyper-Exponentially Distributed Processing Times and Finite Buffer Capacity. *Sensors (Basel, Switzerland)*, 22. <https://doi.org/10.3390/s22249909>
- Kim, S., & Whitt, W. (2014). Are Call Center and Hospital Arrivals Well Modeled by Nonhomogeneous Poisson Processes?. *Manuf. Serv. Oper. Manag.*, 16, 464-480. <https://doi.org/10.1287/msom.2014.0490>
- Ko, Y., & Pender, J. (2018). Strong approximations for time-varying infinite-server queues with non-renewal arrival and service processes. *Stochastic Models*, 34, 186 - 206. <https://doi.org/10.1080/15326349.2018.1425886>
- Kushvaha, B., Das, D., & Tamuli, A. (2024). Single server queueing model with Poisson arrival and gamma service time distribution. *Global Journal of Computer Sciences: Theory and Research*. <https://doi.org/10.18844/gjcs.v14i1.9324>
- Kuzu, K., Soyer, R., & Tarimcilar, M. (2023). A semiparametric Bayesian model for queueing arrival processes: An application to call centers. *Production and Operations Management*, 32, 3266 - 3285. <https://doi.org/10.1111/poms.14034>
- Oreshkin, B., Régnard, N., & L'Ecuyer, P. (2016). Rate-Based Daily Arrival Process Models with Application to Call Centers. *Oper. Res.*, 64, 510-527. <https://doi.org/10.1287/opre.2016.1484>
- Pankratova, E., Moiseeva, S., & Farkhadov, M. (2022). Infinite-Server Resource Queueing Systems with Different Types of Markov-Modulated Poisson Process and Renewal Arrivals. *Mathematics*. <https://doi.org/10.3390/math10162962>
- Rahim, K., & Thiagarajan, M. (2023). M/M(a,b)/1 Model Of Interdependent Queueing With Controllable Arrival Rates. *Indian Journal Of Science And Technology*. <https://doi.org/10.17485/ijst/v16i37.1259>
- Saritha, S., Mamatha, E., Reddy, C., & Anand, K. (2020). A Model for Compound Poisson Process Queueing System with Batch Arrivals and Services. *Journal Européen des Systèmes Automatisés*. <https://doi.org/10.18280/jesa.530110>
- Therasal, S., & Thiagarajan, M. (2024). Poisson Input and Exponential Service Time Finite Capacity Interdependent Queueing Model with Breakdown and Controllable Arrival Rates. *Indian Journal Of Science And Technology*. <https://doi.org/10.17485/ijst/v17i12.2852>
- Yousefi, A., & Pourtaheri, R. (2024). The Mt/M/1 queueing system with impatient customers and multiple vacation. *OPSEARCH*, 61, 2002 - 2022. <https://doi.org/10.1007/s12597-024-00768-y>