Sains Malaysiana 55(5)(2026): 922-935
http://doi.org/10.17576/jsm-2026-5505-14

Finite Mixture of Gamma Interarrival Times and Count Distributions: Computation and Applications

(Campuran Terhingga Lat Ketibaan Gama dan Taburan Bilangan: Komputasi dan Aplikasi)

HARPREET SINGH[1], SENG HUAT ONG[1,2] & CHOUNG MIN NG[1,3,*]

[1]*Institute of Mathematical Sciences, Universiti Malaya, 50603 Kuala Lumpur, Malaysia*
[2]*Institute of Actuarial Science and Data Analytics, UCSI University, 56000 Kuala Lumpur, Malaysia*
[3]*Universiti Malaya Centre for Data Analytics, 50603 Kuala Lumpur, Malaysia*

Received: 27 April 2025/Accepted: 4 May 2026

ABSTRACT

The Poisson process is the basic renewal process with the exponential distribution as the interarrival time distribution. Many nonexponential interarrival distributions have been investigated by researchers along with the derivation of the corresponding count distributions. In many of these distributions, the computation of the count probabilities may be problematic. This paper proposes a finite mixture of gamma distributions for the interarrival times, with the corresponding count distribution derived for a renewal process, motivated by hospital patient arrivals that include both scheduled and walk-in patients. For the proposed finite gamma mixture model, the inverse Laplace transform technique is advocated to overcome the computational difficulties; this generic method is applicable to nonexponential interarrival time models. Some popular numerical inverse Laplace transforms have been compared. The Den Iseger algorithm is recommended for computing probabilities due to its simplicity, high accuracy, and stability. Analyses of both hospital and nonhospital datasets demonstrate the effectiveness of the proposed finite mixture model in accurately representing count data and providing a better fit compared to other models. This finite gamma mixture model, an extension of hyperexponential mixtures, is useful for resource allocation in healthcare, and other fields like finance and economics.

Keywords: Den Iseger algorithm; finite mixture; inverse Laplace transform; patient arrivals; renewal process

ABSTRAK

Proses Poisson ialah proses pembaharuan asas dengan taburan eksponen sebagai taburan lat ketibaan. Banyak taburan lat ketibaan yang bukan eksponen telah dikaji oleh penyelidik bersama dengan terbitan taburan bilangan yang sepadan. Dalam kebanyakan taburan tersebut, pengiraan kebarangkalian bilangan mungkin menjadi masalah. Makalah ini mencadangkan satu campuran terhingga taburan gama untuk lat ketibaan dengan taburan bilangan sepadan diterbitkan untuk satu proses pembaharuan yang didorong oleh ketibaan pesakit hospital yang merangkumi pesakit berjadual dan pesakit tanpa janji temu. Bagi model campuran gama terhingga yang dicadangkan, teknik jelmaan Laplace songsang disarankan untuk mengatasi kesukaran pengiraan; kaedah generik ini boleh digunakan untuk model lat ketibaan bukan eksponen. Beberapa jelmaan Laplace songsang berangka yang popular telah dibandingkan. Algoritma Den Iseger disyorkan untuk mengira kebarangkalian disebabkan oleh ia mudah, berketepatan tinggi dan stabil. Analisis terhadap set data hospital dan bukan hospital menunjukkan keberkesanan model campuran terhingga yang dicadangkan dalam mewakili data bilangan secara tepat dan memberikan penyuaian yang lebih baik berbanding dengan model yang lain. Model campuran gama terhingga ini yang merupakan perluasan daripada campuran hipereksponen berguna untuk pengagihan sumber dalam bidang penjagaan kesihatan serta bidang lain seperti kewangan dan ekonomi.

Kata kunci: Algoritma Den Iseger; campuran terhingga; jelmaan Laplace songsang; ketibaan pesakit; proses pembaharuan

INTRODUCTION

A renewal process is an arrival process in which the interarrival times are positive, independent, and identically distributed random variables. Renewal processes are important stochastic processes for applications in areas such as healthcare, inventory management, queueing systems, and reliability engineering. For example, in healthcare, identifying the arrival pattern of patients is crucial for resource allocation, reducing wait times, and improving service quality. In this paper, we consider a renewal counting process $\{N(t), t \geq 0\}$ where $N(t)$ represents the total number of events (arrivals) that have occurred up to time t. The distribution of $N(t)$ relates to the distribution of interarrival times $\{T_i\}$ by

$$\Pr(N(t) \geq n) = \Pr(S_n \leq t),\ n = 0, 1, 2, \ldots, \quad (1)$$

where $S_n = \sum_{i=1}^{n} T_i = S_{n-1} + T_n$, $S_0 = 0$, is the waiting time until the n-th event. From (1),

$$\Pr(N(t) = n) = \Pr(S_n \leq t) - \Pr = (S_{n+1} \leq t). \quad (2)$$

Equation (2) shows that the arrival time distribution determines the probability distribution of $N(t)$. The Poisson process is a renewal process with exponential interarrival times and Poisson counts. The arrival patterns are determined from the count distributions (Tijms 2003). The relationship (1) has been widely studied in various fields, including health science, actuarial studies, and sports education (Baker & Kharrat 2017; Jose & Abraham 2013; Nadarajah & Chan 2018; Zhang 2018).

Researchers have developed count models and their special cases based on interarrival time distributions such as the gamma, Weibull, and related lifetime distributions (Jose & Abraham 2011; McShane et al. 2008; Ong et al. 2015; Tadikamalla 1980; Winkelmann 1995). The gamma count distribution $N(t)$ (Winkelmann 1995, pp. 469–470) is versatile and can fit both under- and over-dispersed data: $N(t)$ is under-dispersed for $\alpha > 1$ and over-dispersed for $0 < \alpha < 1$, reflecting the behaviour of the hazard function $h(t) = f(t) / [1 - F(t)]$, $t > 0$, where $f(t)$ and $F(t)$ are the probability density function (pdf) and cumulative distribution function (cdf), respectively.

Negative (positive) duration dependence, $h'(t) < 0$ ($h'(t) > 0$), leads to over- (under-) dispersion of $N(t)$ as $t \to \infty$. The gamma distribution captures this duration dependence and the resulting count model can accommodate both under- and over-dispersed data. Winkelmann (1995, Eq. 14) gives a closed-form expression for gamma interarrival times with integer shape parameter α and rate parameter β:

$$\Pr(N(t) = n) = e^{-\beta t} \sum_{i=0}^{\alpha-1} \frac{(\beta t)^{\alpha n+i}}{(\alpha n + i)!},\ n = 0, 1, 2, \ldots .$$

For non-integer α, no closed form is available (Winkelmann 1995, Eq. 12) for $\Pr(N(t) = n) = G(\alpha n, \beta t) - G(\alpha(n + 1), \beta t)$, where $G(\alpha n, \beta t)$ denotes the gamma cdf,

$$G(\alpha n, \beta t) = \frac{1}{\Gamma(n\alpha)} \int_0^{\beta t} e^{-u}\, u^{n\alpha-1}\, du .$$

Interarrival times have also been modelled using lifetime distributions, with the convolution integral used to obtain the distribution of S_n, which in turn determines the distribution of $N(t)$ via (1). Obtaining the probability mass function (pmf) with a convolution integral is difficult for nonexponential distributions, as closed-form solutions for S_n are uncommon (Rausand, Barros & Hoyland 2021). Other methods include polynomial expansion, Laplace transformation, and numerical-transform inversion (Jose & Abraham 2013, 2011; McShane et al. 2008; Ong et al. 2015; Tijms 2003; Whitt 2002). Nadarajah and Chan (2018) presented thirteen discrete distributions for a variety of interarrival time distributions. It is observed that many of the probability distributions are complicated, which will result in computational issues in calculating these count probabilities, such as convergence issues and numerical underflow/overflow.

Recent work has considered hyperexponential and m-component exponential mixtures. Zhang (2018) studied actuarial applications, while Nair and Abdul (2010) examined reliability models using a mixture of two exponential distributions. Hyperexponential distributions have been employed to approximate long tailed interarrival times (Zhang 2018). For a finite mixture of exponential distributions as an interarrival time distribution, Zhang (2018) derived the probabilities in terms of a finite sum of confluent hypergeometric functions. However, many existing models fail to capture heterogeneous arrivals from multiple subpopulations, such as hospital arrivals. Finite gamma mixtures can address this by allowing additional flexibility in shape and hazard behaviour beyond exponential-based mixtures.

This paper extends the results of Nair and Abdul (2010) and Zhang (2018), motivated by heterogeneous patient arrivals consisting of two groups: scheduled (with an appointment or follow-up) and unscheduled (walk-in or emergency). Interarrival times are modelled using a finite mixture of gamma distributions, thereby generalising both previous models. The resulting count distributions are fitted to hospital clinic data, with their corresponding probabilities computed and examined. To address the computational difficulties in evaluating these probabilities, a generic method based on numerical inversion of the Laplace transform is proposed. This provides a more stable and accurate approach for computing probabilities, while accommodating heterogeneous interarrival behaviour through the gamma mixture framework. In stochastic modelling of queueing and reliability systems, the Laplace transform has been broadly utilised.

Numerical-transform inversion techniques, such as Fourier-series method (Abate & Whitt 1995), Piessens method (Piessens 1972), scaling procedure (Choudhury & Whitt 1997), and Stehfest algorithm (Cheng, Sidauruk & Abousleiman 1994), have been widely explored for computing probabilities and other quantities in probability models (Cheng, Sidauruk & Abousleiman 1994; Choudhury & Whitt 1997; Den Iseger 2006). However, these methods have limitations, such as sensitivity to parameter values, roundoff errors, and the need for moment evaluation (Tijms 2003). Den Iseger (2006) introduced a simple Laplace transform inversion technique, gaining attention across diverse fields like thermal engineering (Filippov, Akhmetova & Zelenova 2023). Chaudhry,

Yang and Ong (2013) computed count probabilities for several interarrival time distributions using roots and Padé approximations for complex Laplace transforms. However, Padé approximations method requires moment evaluation for an interarrival time distribution, which may lead to inaccurate approximations if the Laplace transform function is an *n*-component finite mixture.

This study extends exponential mixtures to a finite mixture of two gamma distributions with distinct parameters for interarrival times, applies numerical Laplace inversion (Den Iseger) for stable probability computation, and investigates applications in hospital, finance, and labour data. To present these contributions, the next section details the waiting time and count distributions for a finite mixture of gamma interarrival times, along with the probability generating function. It also introduces the finite-sum count distribution for the more general finite mixture of gamma interarrival times, providing the groundwork for the hazard function and subsequent analyses. Apart from special cases, the computation of the probabilities $\Pr(N(t) = n)$ based on (2) can be complicated; Laplace transform inversion is therefore adopted. The following section compares the Piessens, Fourier-series, and Den Iseger inverse Laplace transform algorithms, highlighting the stability and accuracy of Den Iseger (2006) relative to Piessens (1972) and Abate and Whitt (1995). Numerical examples then illustrate the inaccuracies in probability computations based on the derived pmf. The subsequent section describes the data collected from a specialist clinic and examines its applications to clinic arrivals, stock prices, and labour mobility. The final section provides concluding remarks.

COUNT MODELS WITH FINITE MIXTURE INTERARRIVAL DISTRIBUTIONS AND THEIR CHARACTERISTICS

This section derives the waiting time and count distributions for the proposed finite gamma mixture interarrival model. Using Laplace transforms and the convolution theorem, general expressions and tractable special cases are obtained. For the renewal process $\{N(t),\ t \geq 0\}$, let $F^{(n)}(t)$ be the distribution function of S_n. The following result is useful for computing the right-hand side of (1):

$$F^{(n)}(t) = \int_0^t \left\{ \int_0^{t-\tau} f^{(n-1)}(u)\, du \right\} f_T(\tau)\, d\tau = \int_0^t F^{(n-1)}(t-\tau)\, f_T(\tau)\, d\tau,\quad t > 0, \tag{3}$$

where $f_T(\tau)$ is the interarrival time pdf. According to Rausand, Barros and Hoyland (2021), the convolution integral (3) is practical when closed-form representations of interarrival times are available such as for exponential, Erlang, and gamma distributions. The Laplace transform is employed in deriving and computing quantities related to (2). The Laplace transform of a function $f(\tau)$ is defined as $\mathcal{L}\{f(s)\} = \int_0^\infty e^{-s\tau} f(\tau)\, d\tau$.

For functions $f(t)$ and $g(t)$, defined for $t \geq 0$, with Laplace transforms $\mathcal{F}$ and $\mathcal{G}$, the inverse Laplace transform of their convolution is given by Grewal (2016)

$$\mathcal{L}^{-1}\{\bar{f}(s)\bar{g}(s)\} = \mathcal{L}^{-1}\{\mathcal{L}\{(f * g)(t)\}\} = \int_0^t f(t-\tau)\, g(\tau)\, d\tau = \mathcal{F} * \mathcal{G}. \tag{4}$$

Arrivals, such as those at hospital clinics, are often clustered and not exponentially distributed. A proposed finite mixture model, denoted as gamma(α, λ)-gamma(β, ϕ), is suitable to capture this phenomenon. The pdf for this model is given by

$$f(t) = (1-p)\frac{\lambda^\alpha t^{\alpha-1} e^{-\lambda t}}{\Gamma(\alpha)} + p\,\frac{\phi^\beta t^{\beta-1} e^{-\phi t}}{\Gamma(\beta)},\quad t > 0,\ \lambda, \phi, \alpha, \beta > 0,\ p \in (0,1), \tag{5}$$

where α and β are shape parameters, and λ and ϕ are rate parameters. Here, λ corresponds to scheduled patients, while ϕ corresponds to unscheduled patients. We derive the waiting time S_n by considering the Laplace transform of the pdf (5) given by $\mathcal{L}_{S_n}(s) = \{(1-p)(\lambda^\alpha/(\lambda+s)^\alpha) + p(\phi^\beta/(\phi+s)^\beta)\}^n$. By using the binomial expansion, $\mathcal{L}_{S_n}(s)$ can be written as

$$\mathcal{L}_{S_n}(s) = \sum_{k=0}^{n} \binom{n}{k} (1-p)^k p^{n-k} \frac{\lambda^{\alpha k}}{(\lambda+s)^{\alpha k}} \frac{\phi^{\beta(n-k)}}{(\phi+s)^{\beta(n-k)}}. \tag{6}$$

Based on (6), the distribution of S_n may be interpreted as a finite sum of convolutions of two gamma distributions with different shape and rate parameters. Moschopoulos (1985) derived the distribution for $Y = \sum_{i=1}^n X_i$, the sum of n independent gamma random variables X_i, each with parameters (α_i, β_i), $i = 1, 2, ..., n$. The pdf of Y (Moschopoulos 1985 Theorem 1) is given by

$$g(y) = C \sum_{k=0}^{\infty} \frac{\delta_k y^{\rho+k-1} e^{-y/\beta_1}}{\left(\Gamma(\rho+k)\beta_1{}^{\rho+k}\right)}, \tag{7}$$

where $\beta_1 = \min(\beta_i)$, $C \prod_{i=1}^n \left(\frac{\beta_1}{\beta_i}\right)^{\alpha_i}$, $\rho = \sum_{i=1}^n \alpha_i$, and $\delta_k \frac{1}{k+1} \sum_{i=1}^{k+1} i\, v_i\, \delta_{k+1-i}$, $k = 0, 1, 2, ..., \delta_0 = 1$, where $v_k = \sum_{i=1}^n \alpha_i (1 - \beta_1/\beta_i)^k / k$, 1, 2, 3, ...

Let $g_n(t)$ be the pdf of S_n. By Moschopoulos (1985), $g_n(t) = \sum_{k=0}^n \binom{n}{k} (1-p)^k p^{n-k} g(t; \alpha, \lambda; \beta, \phi)$, where $g(t; \alpha, \lambda; \beta, \phi)$ is the pdf (7) with $\alpha_1 = \alpha$, $\alpha_2 = \beta$, $\beta_1 = \lambda$, and $\beta_2 = \phi$. The following properties hold for the gamma(α, λ)-gamma(β, ϕ) interarrival time (IAT) model and its special cases (including convolution-derived cases) as follows.

For all cases below t, λ, α, β, ϕ, $\mu > 0$; where applicable, $p \in (0, 1)$.

I. Proposed gamma(α, λ)-gamma(β, ϕ) model: General case with distinct parameters. pdf $f(t) = (1-p)\lambda^\alpha e^{-\lambda t} t^{\alpha-1}/\Gamma(\alpha) + p\,\phi^\beta e^{-\phi t} t^{\beta-1}/\Gamma(\beta)$;

cdf $F(t) = 1 - (1-p)\,\Gamma(\alpha, \lambda t)/\Gamma(\alpha) - p\Gamma(\beta, \phi t)/\Gamma(\beta)$;
Laplace transform $\mathcal{L}_{S_n}(s) = \left\{(1-p)\,\lambda^{\alpha}/(\lambda+s)^{\alpha} + p\phi^{\beta}/(\phi+s)^{\beta}\right\}^{n}$.

Laplace transforms for the following special cases can be obtained by substitution into the general form.

II. Special cases of gamma(α, λ)-gamma(β, ϕ) model:

(i) Erlang(α, μ)-Erlang(β, μ) case ($\lambda = \phi = \mu$; α, $\beta \in \mathbb{Z}_+$): A finite sum of Erlang distributions with parameters ($\alpha k + \beta(n-k)$, μ). pdf $f(t) = (1-p)\mu^{\alpha} e^{-\mu t} t^{\alpha-1}/\Gamma(\alpha) + p\,\mu^{\beta} e^{-\mu t} t^{\beta-1}/\Gamma(\beta)$; cdf $F(t) = 1 - (1-p)\,\Gamma(\alpha, \mu t)/\Gamma(\alpha) - p\Gamma(\beta, \mu t)/\Gamma(\beta)$.

(ii) Exp(μ)-$E_2(\mu)$ case ($\lambda = \phi = \mu, \alpha = 1, \beta = 2, p = \mu/(1+\mu)$): A Lindley distribution (Lindley 1958) or finite sum of exponential and Erlang-2 distributions. pdf $f(t) = \mu e^{-\mu t}(1+\mu^2 t)/(1+\mu)$;
cdf $F(t) = \left(1 + \mu - e^{-\mu t} - \mu e^{-\mu t} - \mu^2 t e^{-\mu t}\right)/(1+\mu)$.

(iii) Exp(λ)-Exp(ϕ) case ($\lambda \neq \phi$, $\alpha = \beta = 1$): A hyperexponential distribution of Nair and Abdul (2010) and Zhang (2018). pdf $f(t) = (1-p)\lambda e^{-\lambda t} + p\phi e^{-\phi t}$; cdf $F(t) = 1 - (1-p)e^{-\lambda t} - pe^{-\phi t}$.

(iv) Exp(λ)-$E_j(\phi)$ case ($\lambda \neq \phi$, $\alpha = 1$, $\beta = j \in \mathbb{Z}_+$, $p = \phi/(\lambda+\phi)$): A finite sum of exponential and Erlang-j distributions. pdf $f(t) = \lambda^2 e^{-\lambda t}/(\lambda+\phi) + \phi^{j+1} t^{j-1} e^{-\phi t}/(\lambda+\phi)\,\Gamma(j)$; cdf $F(t) = \lambda(1 - e^{-\lambda t}) + \phi - (\phi\Gamma(j, \phi t)/\Gamma(j))/(\lambda+\phi)$.

III. The gamma(α, μ)-gamma(β, μ) model is another convolution case ($\lambda = \phi = \mu$), forming a finite sum of gamma distributions with parameters ($\alpha k + \beta(n-k)$, μ).

Alternatively, the expressions for the pdf of the distribution of S_n can be obtained by applying (4) to (6). The waiting time pdf for the gamma(α, λ)-gamma(β, ϕ) is obtained as

$$g_n(t) = \sum_{k=0}^{n} \binom{n}{k} \frac{(1-p)^k p^{n-k} \lambda^{\alpha k} \phi^{\beta(n-k)}}{\Gamma(\alpha k)\,\Gamma(\beta(n-k))} \left\{ \int_0^t \frac{\tau^{\alpha k-1}\,(t-\tau)^{\beta(n-k)-1}}{e^{\lambda\tau}\, e^{\phi(t-\tau)}}\, d\tau \right\}, t > 0. \quad (8)$$

By substituting $\tau = t \times h$ in (8), the integration domain becomes $h \in (0, 1)$, yielding an integral equivalent to the confluent hypergeometric function, for α, c, $x \in \mathbb{R}$ (Johnson, Kotz & Kemp 2005). Accordingly, the density of S_n, with $j = \alpha k + \beta(n-k)$, gives the waiting time pdf as

$$g_n(t) = \sum_{k=0}^{n} \binom{n}{k} (1-p)^k p^{n-k} \lambda^{\alpha k} \phi^{\beta(n-k)}\, t^{j-1} e^{-\phi t}\, {}_1F_1(\alpha k, j, (\phi-\lambda)t)/\Gamma(j). \quad (9)$$

when $\phi = \lambda = \theta$ and $\alpha = \beta = v$ are substituted into $g_n(t)$, the gamma pdf is recovered as $e^{-\theta t}\theta^{nv} t^{nv-1}/(nv-1)!$. Winkelmann (1995) shows that the reproductive property of gamma-distributed variables implies that S_n is gamma-distributed. Equation (9) is integrated to define the waiting time distribution function over (0, t) as

$$G_n(t) = \sum_{k=0}^{n} \binom{n}{k} \frac{(1-p)^k p^{n-k}}{\Gamma(j)} \left(\frac{\lambda}{\phi}\right)^{\alpha k} \sum_{m=0}^{\infty} \frac{(\alpha k)_m}{(j)_m} \frac{(1-\lambda/\phi)^m}{m!} (\Gamma(j+m) - \Gamma(j+m, \phi t)), \quad (10)$$

where $j = \alpha k + \beta(n-k)$, $(b)_m$ is the Pochhammer function and $\Gamma(z) - \Gamma(z, \tau) = \gamma(z, \tau)$, with $\Gamma(z, \tau)$ and $\gamma(z, \tau)$ representing the upper and lower incomplete gamma functions, respectively.

The following distributions are considered as special cases of the proposed model, illustrating its flexibility under different parameter structures and hazard behaviours. The Lindley distribution (Lindley 1958) is a finite mixture of Exp(μ) and $E_2(\mu)$. Ghitany, Atieh and Nadarajah (2008) explored the statistical properties and applied the Lindley distribution to fit data from bank service. By using the convolution theorem on the Laplace transform of S_n with the Lindley duration distribution (see II(ii) for the corresponding pdf), the cdf is given by

$$G_n(t) = \sum_{k=0}^{n} \binom{n}{k} \left(\frac{\mu}{\mu+1}\right)^k \left(\frac{1}{\mu+1}\right)^{n-k} Q(2n+k, 0, \mu t), \quad (11)$$

where $Q(z, 0, \tau) = (\Gamma(z) - (\Gamma(z, \tau))/\Gamma(z) = \gamma(z, \tau)/\Gamma(z)$ is the cdf of $E_z(\tau)$.

For the Exp(λ)-$E_j(\phi)$ duration distribution (see II(iv) for the corresponding pdf), we get the cdf as

$$G_n(t) = \sum_{k=0}^{n} \binom{n}{k} \left(\frac{\lambda}{\phi}\right)^{2k} \left(\frac{\phi}{\lambda+\phi}\right)^{n} \sum_{m=0}^{\infty} \frac{(k)_m (\phi-\lambda)^m}{m!\,\phi^m} Q(j(n-k)+k+m, 0, \phi t), \quad (12)$$

where $Q(z, 0, \tau) = (\Gamma(z) - (\Gamma(z, \tau))/\Gamma(z)$ is the $E_z(\tau)$ cdf.

When $\lambda = \phi$ and $j = 1$, we have $E_n(\phi)$. The cdf $G_n(t)$ for these models can be derived using methods for obtaining the count distribution of hyperexponential (Nair & Abdul 2010; Zhang 2018) and generalized Weibull (Ong et al. 2015) distributions. The inverse Laplace transform usually produces a complex formulation with double infinite summation series even when Laplace operator s is factored into three components. However, for (10) to (12), simpler cdf expressions are obtained with the convolution theorem (Equation 4) when the interarrival time distribution includes a finite mixture structure.

Count models obtained by using (10) and (12) differ in parameter count, argument complexity, and infinite summations. Models from (10) and (12) include both rate and shape parameters, making them more suitable for modelling two distinct patient groups than the model from (11). However, all models still involve finite-infinite series summations. Considering a more general case for tractability, we use a finite mixture of two gamma distributions with common parameter μ, specifically gamma(α, μ)-gamma(β, μ), where $\mu > 0$. The cdf can be expressed using (4) as

$$G_n(t) = \sum_{k=0}^{n} \binom{n}{k} (1-p)^k p^{n-k} \frac{\gamma(\alpha k + \beta(n-k), \mu t)}{\Gamma(\alpha k + \beta(n-k))}, \quad (13)$$

where γ denotes the lower incomplete gamma function and Γ denotes the gamma function. For integer α and β, this represents the cdf of an Erlang distribution with parameters $(\alpha k + \beta (n - k), \mu t)$, which is a special case of the convolution. Thus, applying the Winkelmann (1995) results, and letting $j = \alpha k + \beta (n - k)$ and $j^* = \alpha k + \beta (n + 1 - k)$, the pmf for the count distribution is given by

$$\Pr(N(t) = n) = \sum_{k=0}^{n} \binom{n}{k} (1-p)^k p^{n-k} (1 - F(j; \mu t)) - \sum_{k=0}^{n+1} \binom{n+1}{k} (1-p)^k p^{n+1-k} (1 - F(j^*; \mu t)), \quad (14)$$

where $F(x; \mu t) = \sum_{j=0}^{x} (\mu t)^j e^{-\mu t} / \Gamma(j+1)$ is the cdf of Poisson distribution. This formulation highlights that the cdf and pmf involve double finite summations, making them computationally manageable. The following plots in Figure 1 illustrate the count model from the gamma (α, λ)-gamma(β, ϕ) distribution, showing diverse count distribution shapes for five parameters sets with t =1 for simplicity.

It produces diverse distribution shapes, with parameter variations shifting the mode and altering skewness. Figure 1(a) and 1(b) depicts right-skewed distributions with long tails, showing zero and nonzero modes, respectively. Figure 1(c) shows symmetrical shape with large λ and ϕ while Figure 1(d), show bimodal distribution, occurring when $\phi \to 0$ and $\beta = 0.02$. Figure 1(e) displays a left-skewed distribution with small β and large λ. As ϕ decreases, the mode moves to zero, influenced by α and β.

HAZARD FUNCTION FOR gamma(α, μ)-gamma(β, μ) MODEL

As stated in the Introduction, the hazard rate of interarrival times affects the dispersion of count distribution. For example, a mixture of exponentials with a decreasing hazard function is suitable for over-dispersed data, indicating negative duration dependence, while an increasing hazard function suggests positive duration dependence, leading to under-dispersion (Shortle et al. 2018). The $h'(t)$ for the gamma(α, λ)-gamma(β, ϕ) model is complex due to the involvement of multiple layers of products, exponentials, polynomials, and gamma functions. To simplify, we set $\lambda = \phi = \mu$ and obtain the hazard function

$$h(t) = \frac{e^{-\mu t}\left[(1-p)\mu^{\alpha} t^{\alpha-1} + p\mu^{\beta} t^{\beta-1}\right]}{\Gamma(\alpha)\,\Gamma(\beta)\left[p\Gamma(\beta, \mu t)/\Gamma(\beta) - (1-p)\,\Gamma(\alpha, \mu t)/\Gamma(\alpha)\right]},$$

where $\alpha, \beta, \mu > 0$. This function helps determine the under- or over-dispersion of the gamma(α, μ)-gamma(β, μ) count distribution. Plotting $h(t)$ for various values of α, β, μ, and p shows patterns in the hazard function. The exponential term $e^{-\mu t}$ controls the decay rate, while polynomial terms $t^{\alpha-1}$ and $t^{\beta-1}$ are influenced by α and β, and the factors μ^{α} and μ^{β}. The mixing parameter p balances the contributions of the different shape parameters.

As $t \to \infty$ and $\mu \to 0$, the normalised incomplete gamma function scaled by $1/\Gamma(\alpha)$ approaches one. Several cases were analysed to observe trends in the model's hazard functions. Figure 2 highlights how α and β influence the hazard function's initial behaviour and trend over time. For $\alpha, \beta > 1$ (Figure 2(a)), it is monotonically increasing (solid line) or follows an increasing-decreasing-increasing pattern (non-solid lines). When $\alpha > 1$ and $\beta < 1$ (Figure 2(b)), it takes a bathtub shape. If $\alpha, \beta < 1$ (Figure 2(c)), it is monotonically decreasing, whereas for fixed $\beta > 1$ and $\alpha < 1$ (Figure 2(d)), it again forms a bathtub shape. When $\alpha = \beta = 1$ (Figure 2(e)), the hazard function remains constant. Additionally, for $\alpha, \beta > 1$ (Figure 2(f)), a unimodal pattern appears within $t \in (0, 10)$.

COMPUTATION OF THE COUNT PROBABILITIES

To obtain the pmf (2) for the count distribution of the gamma(α, λ)-gamma(β, ϕ) model, one can subtract two consecutive terms of the cdf,

$$F^{(n)}(t) - F^{(n+1)}(t) = \Pr(N(t) = n).$$

However, this method is cumbersome due to the complex expression of $F^{(n)}(t)$. A straightforward approach involves the numerical inversion of the Laplace transform, one such method is outlined by Den Iseger (2006). This method uses the Gaussian quadrature rule to efficiently approximate the infinite sum of Laplace transform values by computing function values at $f(k\Delta)$, $k = 0, 1, ..., M - 1$, for $\Delta > 0$ and $M > 1$, where $f(k\Delta)$ refers to the value of the function $f(t)$ at time $t = k\Delta$. The fast Fourier transform (FFT) is employed for this computation, offering a time complexity of $O(M \log(M))$ time. The discrete FFT method is advantageous for its speed, numerical stability, and simplicity. Den Iseger's algorithm, which belongs to the Fourier-series methods, efficiently inverts the Laplace transform and returns values at regular intervals. For further mathematical details, refer to Appendix A of Den Iseger (2006) and Appendix F of Tijms (2003).

COMPARISON, VALIDATION, AND IMPLEMENTATION OF ALGORITHMS

In computational probability models, it is well-known that numerical errors sometimes cause the total probabilities to exceed one. This common issue, often arising from rounding approximations, highlights the importance of ensuring probabilities sum to one as one of the fundamental rules in probability theory. The following analysis compares the Piessens, Fourier-series, and Den

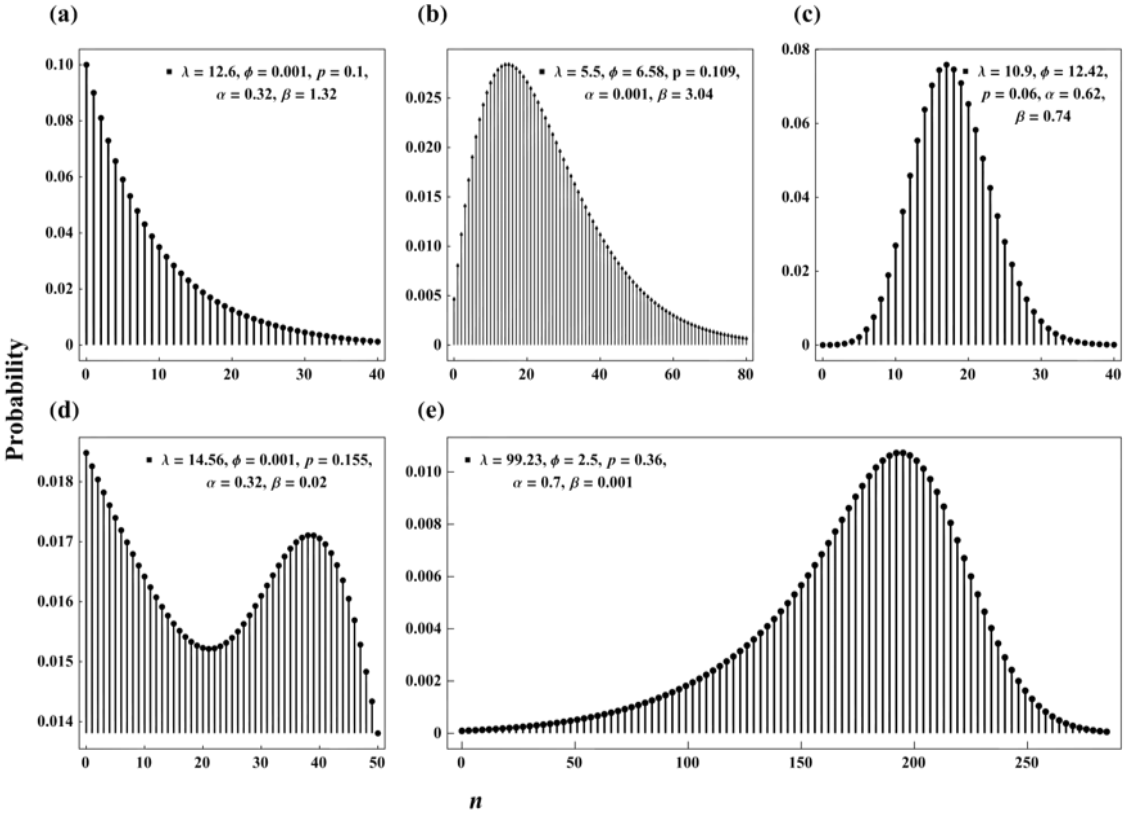

FIGURE 1. Probability plots for the count distribution of the gamma(α, λ)-gamma(β, ϕ) model show: (a) peak at 0 with a long tail; (b) long tail right-skewed distribution; (c) symmetrical distribution; (d) bimodal distribution; and (e) long tail left-skewed distribution

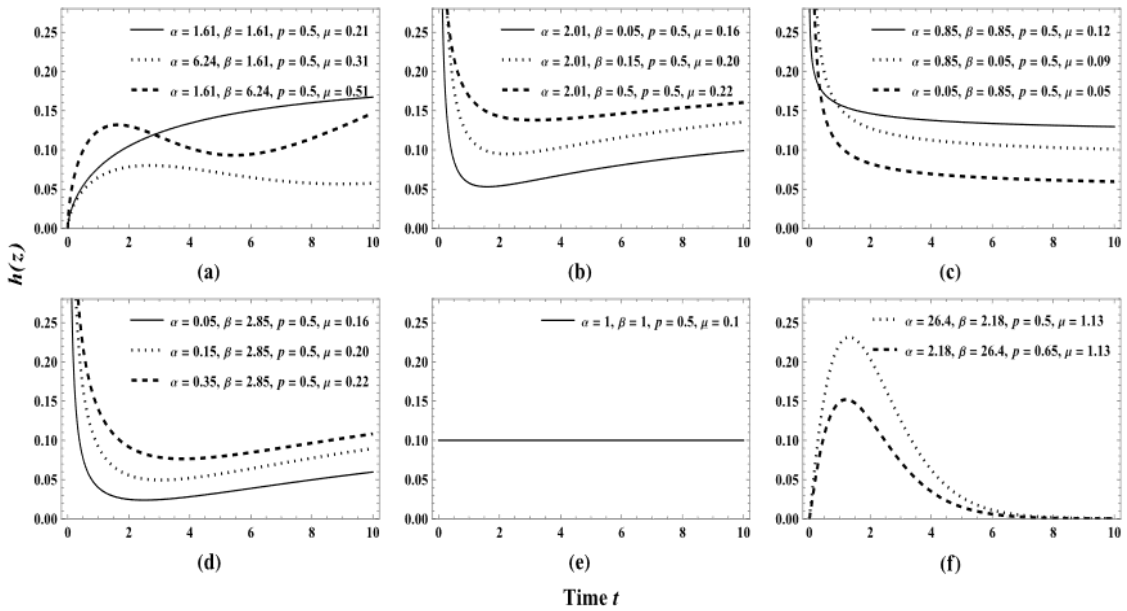

FIGURE 2. Hazard function plots for the gamma(α, λ)-gamma(β, ϕ) IAT model with common labels for the axes show: (a) varying α and β, when $\alpha > 1$, $\beta > 1$; (b) fixed α but varying β, when $\alpha > 1$, $\beta < 1$; (c) varying α and β when $\alpha < 1$, $\beta < 1$; (d) fixed β but varying α, when $\alpha < 1$, $\beta > 1$; (e) when $\alpha = \beta = 1$; and (f) when $\alpha > 1$, $\beta > 1$ with common μ and different p

Iseger algorithms to identify a stable and accurate method for approximating probabilities, targeting at least six significant digits of accuracy, and focusing on rounding errors. This involves using the Laplace transforms of the hyperexponential (Exp(λ) − Exp(ϕ)) and the finite mixture of two Erlang distributions (Erlang(α, μ)-Erlang(β, μ)). These IAT models are special cases of the more general gamma(α, μ)-gamma(β, μ) model, providing flexibility and finite sums.

We computed probabilities from the count distribution for the hyperexponential distribution (abbreviated as CHE) using Zhang (2018, Eq. 3.13) as a benchmark to compare the three algorithms. The probabilities were computed using the direct pmf and approximation methods for two different sets of parameters, covering values of n from 0 to 50. Table 1 presents results for parameter settings (a) and (b), with vertical ellipses (⋮) indicating omitted intermediate values of n. Differences in approximations are underlined, showing deviation in tail probabilities.

Table 1 shows that all three Piessens, Fourier-series, and Den Iseger algorithms provide probabilities consistent with the direct pmf up to six significant digits for the initial values of n (up to n = 10)). However, rounding errors emerge as n increases, reflecting the approximate nature of these methods. Particularly, the Piessens and Fourier-series methods exhibit deviations in the tail probabilities. For example, at n = 30, Piessens method shows a negative value, while the Fourier-series produces a larger deviation at higher n values.

The total probability values also reflect the rounding errors, with slight deviations observed in Piessens and Fourier-series (1.00000176 and 1.00000001, respectively). Den Iseger maintains accuracy in the initial values and ensures that the total probability sums to 1. In contrast, Piessens and Fourier-series show small deviations, which may be acceptable depending on the required precision. The results from Table 1 indicate that the Den Iseger algorithm demonstrates better approximations to the direct pmf, highlighting its stability.

This section examines and implements the Den Iseger algorithm using the Erlang(α, μ)-Erlang(β, μ) count distribution (Equation 14), which has greater shape flexibility compared to the Exp(λ)-Exp(ϕ) count distribution. The analysis uses the Laplace transform of Erlang(α, μ)-Erlang(β, μ) with a single set of parameters to evaluate the Den Iseger algorithm's accuracy and stability for various t. Table 2 shows probabilities calculated using the direct pmf (14) and Den Iseger algorithm, respectively. Single values for each n and t indicate probabilities from both the direct pmf and the Den Iseger method. In cells with two values, the first is from the direct pmf, while the underlined value is from the Den Iseger method.

Probabilities are presented up to three significant digits to capture most variations with changing t. Most approximated probabilities align, with discrepancies underlined in Table 2. The last row of each table shows the total probabilities summing to either exactly one or approaching one. In contrast, the Den Iseger algorithm approximated slightly higher probabilities for $n \geq 5$ (Table 2) compared to the direct pmf method, with notable differences when t = 1. While direct pmf calculations can face underflow issues, especially at $t \geq 10$, the Den Iseger algorithm provides stable approximations across different values of t.

Table 3 compares the count probabilities for the gamma(α, λ)-gamma(β, ϕ) model computed directly using (10) in $G_n(t) - G_{n+1}(t)$ and the Den Iseger algorithm. To exemplify the tendency of inaccuracies, Table 3 shows the count probabilities for different t, with two different sets of parameters. When n equals 1–5 and 8 (Table 3(a)) and when n equals 3–8 (Table 3(b)) respectively, there are numerical errors in the computation of the probabilities directly from the pmf formula.

DATA COLLECTION AND APPLICATIONS OF THE COUNT DISTRIBUTION

To assess performance on heterogeneous arrival patterns, we analyse data collected from a specialist clinic in Malaysia. Elective admissions were recorded daily from Monday to Friday, 8:00 a.m. to 5:00 p.m., with queues starting as early as 7:00 a.m. Patient arrivals, both scheduled and emergency, were recorded in 15-minute intervals from 7:15 a.m. to 2:00 p.m., using a queue management system (QMS). Data were collected over 98 days, with each day comprising 27 intervals of 15 minutes each. The number of arrivals for each interval was obtained from the QMS report, with zero recorded for intervals with no arrivals. To determine the pattern of a mixture of clustered arrivals (a group of two subpopulations of patients), this study aims to model clustered arrivals during peak and nonpeak periods for effective resource allocation. Statistical analysis was performed using Mathematica (v12).

Hospital arrival modelling has traditionally employed homogeneous or non-homogeneous Poisson processes (Bhattacharjee & Ray 2014; Feller 1971). Since patient arrivals often occur in clusters and exhibit over-dispersion, Singh et al. (2024) recently proposed Lévy-type processes based on Poisson-stopped sum distributions as an alternative to the non-homogeneous Poisson process, fitting these to a mixture of clustered patient arrivals across weekdays over the course of 98 days.

HOSPITAL DATASETS WITH OVER-DISPERSION AND LONG TAIL

We analyse 98 days of hospital data across peak and nonpeak periods (Tables 4 & 5). Using the Den Iseger algorithm, we approximate probabilities from the count models of (i) gamma(α, λ)-gamma(β, ϕ) (CG); (ii) Exp-E_{20} (CEE_{20}) with j set to 20 for comparable fitting; (iii) Lindley (CL); (iv) gamma(α, μ)-gamma(β, μ) (CGG; see next section); and (v) generalized Weibull distribution (CGW). Model parameters are estimated through likelihood function maximisation using simulated annealing.

TABLE 1. Count probabilities and their approximations using the Piessens, Fourier-series, and Den Iseger inverse Laplace transform algorithms for the hyperexponential model. Two parameter setting are: (a) $\lambda = 0.356$, $\phi = 2.37$, $p = 0.25$, $t = 1$ and (b) $\lambda = 3.271$, $\phi = 0.482$, $p = 0.25$, $t = 1$. The direct pmf represents the exact count probabilities for comparison

n	From direct pmf	Piessens	Fourier-series	Den Iseger
0	(a) 0.548725	(a) 0.548726	(a) 0.548725	(a) 0.548725
	(b) 0.182863	(b) 0.182865	(b) 0.182863	(b) 0.182863
⋮	⋮	⋮	⋮	⋮
10	(a) 1.24048×10^{-8}	(a) 1.24048×10^{-8}	(a) 1.24053×10^{-8}	(a) 1.24048×10^{-8}
	(b) 0.000167739	(b) 0.000167739	(b) 0.000167739	(b) 0.000167739
⋮	⋮	⋮	⋮	⋮
20	(a) 3.77922×10^{-21}	(a) 3.77922×10^{-21}	(a) 7.23969×10^{-21}	(a) 3.94591×10^{-21}
	(b) 3.03410×10^{-12}	(b) 3.03411×10^{-12}	(b) 3.17950×10^{-12}	(b) 3.03976×10^{-12}
⋮	⋮	⋮	⋮	⋮
30	(a) 7.42503×10^{-36}	(a) -4.24322×10^{-35}	(a) 3.49890×10^{-31}	(a) 2.64291×10^{-32}
	(b) 3.49211×10^{-22}	(b) 3.45287×10^{-22}	(b) 8.87622×10^{-19}	(b) 5.46502×10^{-20}
⋮	⋮	⋮	⋮	⋮
50	(a) 3.06610×10^{-69}	(a) -2.48711×10^{-58}	(a) 1.36350×10^{-53}	(a) 1.91643×10^{-55}
	(b) 4.89170×10^{-46}	(b) -4.39447×10^{-36}	(b) 1.02779×10^{-32}	(b) 6.60778×10^{-34}
Total probability for cases (a) and (b)	1	1.00000176	1.00000001	1

TABLE 2. Count distribution probabilities of Erlang(α, μ)-Erlang(β, μ) with $\alpha = 10$, $\beta = 5$, $\mu = 4.1$, $p = 0.71$ using direct pmf (14) and Den Iseger algorithm at different values of t

n	$t = 1$	$t = 2$	$t = 3$	$t = 4$	$t = 5$	$t = 6$	$t = 7$
0	0.720	0.264	0.0673	0.0105	0.00109	0.0000838	5.18×10^{-6}
1	0.275	0.572	0.430	0.210	0.0724	0.0189	0.00386
2	0.00481	0.157	0.399	0.440	0.314	0.164	0.0662
3	9.43×10^{-6}	0.00754	0.0962	0.272	0.382	0.353	0.242
4	3.87×10^{-9}	0.0000879	0.00697	0.0616	0.186	0.307	0.343
5	4.77×10^{-13}	3.25×10^{-7}	0.000177	0.00556	0.0404	0.127	0.236
	4.78×10^{-13}						
6	2.23×10^{-17}	4.67×10^{-10}	1.81×10^{-6}	0.000218	0.00419	0.027	0.0873
	2.95×10^{-16}						
7	4.61×10^{-22}	3.01×10^{-13}	8.53×10^{-9}	4.08×10^{-6}	0.000218	0.00307	0.0182
	8.45×10^{-17}						
8	4.73×10^{-27}	9.69×10^{-17}	2.03×10^{-11}	3.94×10^{-8}	6.05×10^{-6}	0.000194	0.00222
	1.58×10^{-17}	6.7×10^{-16}					
9	2.62×10^{-32}	1.69×10^{-20}	2.65×10^{-14}	2.11×10^{-10}	9.49×10^{-8}	7.15×10^{-6}	0.000162
	1.83×10^{-18}	6.65×10^{-16}					
10	8.40×10^{-38}	1.72×10^{-24}	2.01×10^{-17}	6.61×10^{-13}	8.83×10^{-10}	1.59×10^{-7}	7.35×10^{-6}
	1.34×10^{-19}	5.39×10^{-16}	5.39×10^{-16}				
Total	1	1	1	1	1	1	0.9999

TABLE 3. Count distribution probabilities from the direct pmf of the gamma(α, λ)-gamma(β, ϕ) and the Den Iseger algorithm for (a) $\lambda = 10.3$, $\phi = 0.3$, $p = 0.2$, $\alpha = 6.5$, $\beta = 1.5$, $t = 2$, and (b) $\lambda = 4.7$, $\phi = 0.2$, $p = 0.65$, $\alpha = 3.05$, $\beta = 0.05$, $t = 5$

(a)			(b)		
n	From direct pmf	Den Iseger approximation	n	From direct pmf	Den Iseger approximation
0	0.131078	0.150672	0	0.00760742	0.00749042
1	−2.67405	0.177428	1	0.0231873	0.00929222
2	21.6198	0.322882	2	0.311812	0.0113858
3	−26.3772	0.283422	3	−0.85744	0.0138594
4	7.36265	0.0617494	4	0.741977	0.0168625
5	0.918036	0.00376612	5	2.02347	0.0205644
6	0.019366	0.0000803667	6	−3.94727	0.0250716
7	0.000305081	7.44229×10^{-7}	7	2.41194	0.0303490
8	-1.58797×10^{-7}	3.59664×10^{-9}	8	8.72925	0.0361846

TABLE 4. ML estimates and log-likelihood values for hospital and nonhospital datasets (CG = Count gamma(α, λ)-gamma(β, ϕ); CEE_{20} = Count Exp-E_{20}; CL = Count Lindley; CHE = Count Hyperexponential; CGW = Count generalized Weibull; and CGG = Count gamma(α, μ)-gamma(β, μ))

Dataset	ML estimates (model)	Log-likelihood
Hospital peak period	$\hat{\lambda} = 14.0778, \hat{\phi} = 0.0132, \hat{p} = 0.0955, \hat{\alpha} = 0.3588, \hat{\beta} = 1.3336$ (CG)	−7030.88
	$\hat{\lambda} = 36.9197, \hat{\phi} = 3.8349$ (CEE_{20})	−7033.06
	$\hat{\lambda} = 10.035$ (CL)	−15286.9
	$\hat{\lambda} = 36.478, \hat{\phi} = 0.149, \hat{p} = 0.104$ (CHE)	−7067.37
Hospital nonpeak period	$\hat{\lambda} = 1.1187, \hat{\phi} = 0.9943, \hat{p} = 0.6067, \hat{\alpha} = 8.0856, \hat{\beta} = 0.013$ (CG)	−5337.12
	$\hat{\lambda} = 16.2578, \hat{\phi} = 10.5246$ (CEE_{20})	−5401.29
	$\hat{\lambda} = 2.126$ (CL)	−8008.07
	$\hat{\lambda} = 12.879, \hat{\phi} = 0.008, \hat{p} = 0.391$ (CHE)	−5449.32
IBM stock price	$\hat{\mu} = 4.8198, \hat{p} = 0.1488, \hat{\alpha} = 5.4996, \hat{\beta} = 1.1925$ (CGG)	−22444.14
	$\hat{a} = 0.34135, \hat{\alpha} = 0.95724, \hat{\lambda} = 0.4017$ (CGW)	−22483.40
	$\hat{\lambda} = 0.6589, \hat{\phi} = 0.0018$ (CEE_{20})	−23109.80
Labour mobility	$\hat{\mu} = 0.0278, \hat{p} = 0.4797, \hat{\alpha} = 0.0556, \hat{\beta} = 1.5206$ (CGG)	–981.075
	$\hat{a} = 0.7942, \hat{\alpha} = 0.1539, \hat{\lambda} = 0.9445$ (CGW)	–994.577
	$\hat{\lambda} = 8.9539, \hat{\phi} = 11.2014$ (CEE_{20})	–993.4

For the hospital datasets, Table 4 shows varying estimates of arrival rate λ. The CG model suggests a moderate rate, while CEE_{20} and CHE estimate higher rates. The CG model estimates a relatively low arrival rate for ϕ compared to λ, whereas CEE_{20} estimates a higher ϕ value than both the CG and CHE models, indicating that the CG model suggests a lower proportion of arrivals for one subpopulation.

Figure 3(a) and 3(b) illustrates that the CG model provides a satisfactory fit, whereas the CL model proves inadequate for the given datasets. Table 5(a) shows that the CG and CEE_{20} models provide nearly identical fits for peak arrivals, with AIC values of 14069.8 and 14070.1, respectively. As the AIC difference is less than 2, both models are effectively equivalent (Burnham & Anderson 2004). However, the BIC for CEE_{20} (14081.5) is lower

TABLE 5. Observed (ObsF) and expected (ExpF) frequencies for the hospital's (a) peak and (b) nonpeak datasets (CG = Count gamma(α, λ)-gamma(β, ϕ) model; CEE_{20} = Count Exp-E_{20}; CL = Count Lindley, with expected frequencies grouped; CHE = Count Hyperexponential)

(a) Peak period

	ObsF	ExpF					ObsF	ExpF			
n		CG	CEE_{20}	CL	CHE	*n*		CG	CEE_{20}	CL	CHE
0	360	205.37	202.87	0.18	193.18	21	33	26.16	25.78	1.18	26.79
1	221	185.83	183.78	1.82	177.51	22	27	23.85	23.55		24.29
2	211	168.17	166.49	8.74	162.92	23	26	21.75	21.58		22.07
3	155	152.17	150.82	27.48	149.36	24	27	19.83	19.87		20.10
4	127	137.70	136.63	64.21	136.78	25	26	18.07	18.37		18.35
5	83	124.60	123.78	119.28	125.14	26	23	16.46	17.07		16.79
6	65	112.74	112.13	183.93	114.36	27	27	14.97	15.93		15.38
7	61	102.02	101.58	242.43	104.41	28	21	13.59	14.89		14.09
8	63	92.33	92.02	279.05	95.23	29	17	12.31	13.92		12.89
9	55	83.56	83.36	285.09	86.78	30	12	11.13	12.99		11.73
10	51	75.64	75.52	261.84	79.00	31	20	10.03	12.06		10.61
11	36	68.49	68.41	218.42	71.85	32	14	9.00	11.10		9.51
12	43	62.03	61.97	166.90	65.28	33	15	8.05	10.11		8.43
13	38	56.21	56.14	117.65	59.27	34	16	7.17	9.09		7.38
14	44	50.96	50.86	76.97	53.76	35	14	6.36	8.05		6.35
15	47	46.22	46.08	46.98	48.73	36	5	5.61	7.02		5.39
16	23	41.96	41.76	26.88	44.13	37	9	4.92	6.01		4.48
17	31	38.12	37.85	14.47	39.94	38	4	4.30	5.05		3.67
18	30	34.66	34.32	7.35	36.13	39	2	3.74	4.16		2.94
19	27	31.53	31.15	3.54	32.69	40	6	3.23	3.36		2.31
20	28	28.71	28.31	1.62	29.58	41+	13	16.47	10.19		6.43

	ObsF	ExpF			
Total	2156	2156	2156	2156	2156
Number of classes	42	42	42	22	42
Number of estimated parameters		5	2	1	3
d.f.		36	39	20	38
χ^2		325.278	303.322	860037	381.16
p-value		$< 10^{-3}$	$< 10^{-3}$	$< 10^{-3}$	$< 10^{-3}$
AIC		14069.8	14070.1	30575.8	14138.7
BIC		14092.5	14081.5	30581.5	14157.8

(b) Nonpeak period

	ObsF	ExpF					ObsF	ExpF			
n		CG	CEE_{20}	CL	CHE	*n*		CG	CEE_{20}	CL	CHE
0	1827	1233.94	1225.03	628.64	1216.41	7	46	37.26	37.65	2.86	39.56
1	403	748.42	750.11	1056.81	745.55	8	34	22.60	22.91		24.07
2	263	453.94	456.61	821.90	456.68	9	28	13.71	13.94		14.44

continue to next page

continue from previous page

3	159	275.33	277.19	414.37	279.69	10	22	8.31	8.46		8.46
4	121	166.99	168.13	154.50	171.38	11	21	5.04	5.09		4.80
5	89	101.29	101.99	45.71	105.14	12	6	3.06	3.03		2.63
6	87	61.43	61.93	11.21	64.55	13+	30	4.68	3.92		2.65

	ObsF	ExpF			
Total	3136	3136	3136	3136	3136
Number of classes	14	14	14	8	14
Number of estimated parameters		5	2	1	3
d.f.		8	11	6	10
χ^2		834.33	883.15	15642.79	1004.01
p-value		$< 10^{-3}$	$< 10^{-3}$	$< 10^{-3}$	$< 10^{-3}$
AIC		10682.2	10806.6	16018.1	10902.6
BIC		10706.4	10818.7	16024.2	10922.8

CL model expected frequencies are grouped from $n = 22$ onwards and $n = 8$ onwards for the peak and nonpeak periods, respectively

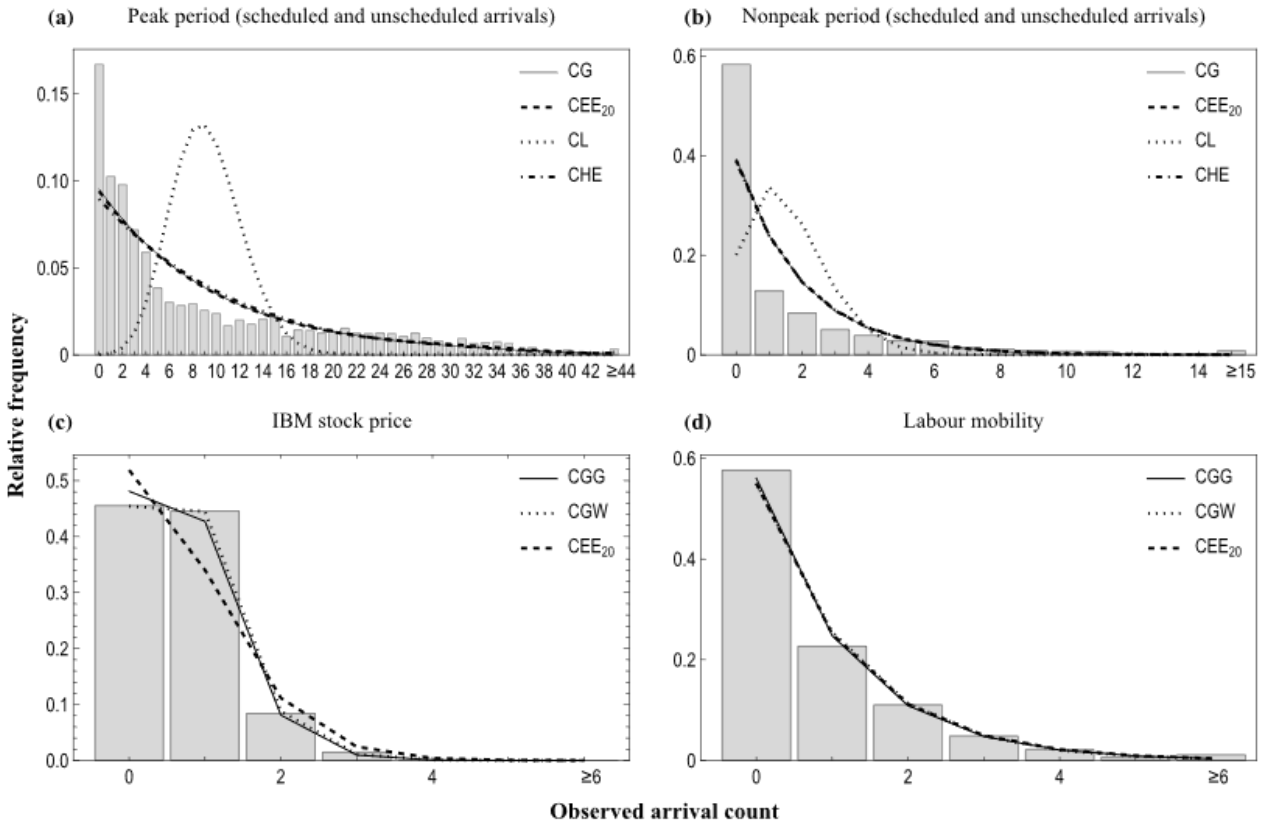

FIGURE 3. Relative frequency plots for (a) peak over-dispersed dataset and (b) nonpeak over-dispersed dataset, fitted with CG (solid line), CEE_{20} (dashed line), CL (dotted line), and CHE (dot dashed line) models, (c) IBM stock price under-dispersed dataset and (d) labour mobility over-dispersed dataset, fitted with CGG (solid line), CGW (dotted line), and CEE_{20} (dashed line)

than CG (14092.5), favouring CEE_{20} as the simpler model. During nonpeak periods (Table 5(b)), characterised by low admissions and a high frequency of zeros, CG has the lowest AIC (10682.2) and BIC (10706.4), indicating a better fit overall. While the CL model performs poorly across both datasets, the CEE_{20} model emerges as a practical alternative for implementation due to its simplicity, comparable fit, and reduced computational time.

NONHOSPITAL DATASETS WITH UNDER-DISPERSION AND OVER-DISPERSION

To assess applicability beyond healthcare settings, we analyse two nonhospital datasets with differing dispersion characteristics: IBM stock price data and labour mobility data (Ong et al. 2015). Figure 3(c) and 3(d) shows that the IBM dataset has a very short tail with a mode at zero, whereas the labour mobility dataset has a slightly longer tail, also with a mode at zero. Both have shorter tails than the hospital datasets, one of which has a notably longer tail than the other. The generalized Weibull count distribution (CGW), introduced by Ong et al. (2015), fits these datasets well. Low and Ong (2016) reviewed the CHE model specifically for labour mobility data. Table 4 also shows the ML estimates and log-likelihood values for the IBM and labour mobility datasets, respectively. The CG model, which has five parameters, is unsuitable for modelling IBM stock price data with such short tail. Instead, the count distribution of gamma(α, μ)-gamma(β, μ) model, denoted as CGG, derived from (13) in $G_n(t) - G_{n+1}(t)$, is selected. This model provides a better fit compared to both CGW and CEE_{20} models.

Table 6 presents the fit of all candidate models using model selection criteria and goodness-of-fit tests. In Table 6(a), CGG achieves a lower χ^2, AIC and BIC, indicating it fits the data better than the other models, though it uses more parameters, making it a better choice for this dataset. Conversely, the CEE_{20} and CGW models exhibit a poorer fit for the under-dispersed IBM stock price data, as indicated by their higher AIC and BIC values. In Table 6(b), the CGG model also demonstrates a closer fit compared to the CGW and CEE_{20} models based on AIC and BIC values.

TABLE 6. Observed (ObsF) and expected (ExpF) frequencies for (a) IBM stock price, and (b) labour mobility data (CGG = Count gamma(α, μ)-gamma(β, μ); CGW = Count generalized Weibull; CEE_{20} = Count Exp-E_{20})

	(a) IBM stock price data				(b) labour mobility data			
	ObsF	ExpF			ObsF	ExpF		
n		CGG	CGW	CEE_{20}		CGG	CGW	CEE_{20}
0	10333	10914.6	11369.7	11769.7	465	451.91	443.44	443.56
1	10108	9693.38	9155.8	7722.98	183	199.62	205.74	202.38
2	1913	1838.2	1894.25	2536.34	89	87.71	91.37	90.2
3	335	220.68	244.15	555.44	39	38.35	39.18	39.99
4	0	20.48	23.23	91.23	17	16.69	16.32	17.68
5	0	1.70	1.88	11.99	5	7.24	6.64	7.73
6+	0			1.37	9	5.47	4.3	5.46
Total	22689	22689	22689	22689	807	807	807	807
Number of classes		6	6	7		7	7	7
Number of estimated parameters		4	3	2		4	3	2
d.f.		1	2	4		2	3	4
χ^2		133.171	251.368	1257.19		4.7697	9.1935	6.2182
p-value		$< 10^{-3}$	$< 10^{-3}$	$< 10^{-3}$		0.0921	0.0268	0.1834
AIC		44896.3	44972.8	46226.4		1970.15	1995.15	1990.8
BIC		44928.4	44996.9	46239.7		1988.92	2009.23	2000.19

CONCLUSION

This paper introduces a general gamma(α, λ)-gamma(β, ϕ) model, which nests several special convolution cases within finite mixture interarrival time models and offers a simplified cumulative distribution function through the convolution theorem. The gamma(α, μ)-gamma(β, μ) model, used for the hazard rate analysis, highlights the utility of the gamma(α, λ)-gamma(β, ϕ) model across various count data scenarios, particularly in healthcare resource management. While Piessens, Fourier-series, and Den Iseger methods generally achieve similar accuracy, differences in tail probabilities and rounding errors are observed. With the finite-sum count model of Erlang(α, μ)-Erlang(β, μ), the Den Iseger algorithm provides a stable approximation across different values of t. It stands out for its reliability in capturing tail probabilities with greater accuracy, minimising rounding errors, and ensuring consistency in total probability. This makes it particularly suitable for long-tailed distributions.

Among the models, although having more parameters, CG and CGG exhibit the lowest AIC values for hospital and nonhospital datasets, respectively. Meanwhile, the CEE_{20} model also shows potential in modelling over-dispersed data, especially in scenarios with high zero counts and long tails. These findings underscore the strengths and potential applications of each model. Overall, the gamma(α, λ)-gamma(β, ϕ) model and its special cases perform well across different datasets. While this study focuses on stochastic models, integrating machine learning techniques through hybrid approaches, such as combining neural networks, regression models, and traditional methods, could improve arrival time predictions. Future studies could test the model in areas like telecommunications, finance, and logistics to assess its broader applicability. Exploring these areas will help refine renewal process models and enhance their efficiency across industries.

ACKNOWLEDGEMENTS

Seng Huat Ong is supported by the Ministry of Higher Education, Malaysia grant 596 FRGS/1/2020/STG06/SYUC/02/1 and UCSI University grant REIG-FBM-2022/050. We sincerely thank the reviewers for their insightful comments, which greatly helped to improve this paper.

REFERENCES

Abate, J. & Whitt, W. 1995. Numerical inversion of laplace transforms of probability distributions. ORSA Journal on Computing 7(1): 36-43.

Baker, R. & Kharrat, T. 2017. Event count distributions from renewal processes: Fast computation of probabilities. IMA Journal of Management Mathematics 29(4): 415-433.

Bhattacharjee, P. & Ray, P.K. 2014. Patient flow modelling and performance analysis of healthcare delivery processes in hospitals: A review and reflections. Computers and Industrial Engineering 78: 299-312.

Burnham, K.P. & Anderson, D.R. 2004. Multimodel inference: Understanding AIC and BIC in model selection. Sociological Methods & Research 33(2): 261-304.

Chaudhry, M.L., Yang, X. & Ong, B. 2013. Computing the distribution function of the number of renewals. American Journal of Operations Research 3: 380-386.

Cheng, A.H., Sidauruk, P. & Abousleiman, Y. 1994. Approximate inversion of the laplace transform. Mathematica Journal 4(2): 76-82.

Choudhury, G.L. & Whitt, W. 1997. Probabilistic scaling for the numerical inversion of nonprobability transforms. INFORMS Journal on Computing 9(2): 175-184.

Den Iseger, P. 2006. Numerical transform inversion using Gaussian quadrature. Probability in the Engineering and Informational Sciences 20(1): 1-44.

Feller, W. 1971. An Introduction to Probability Theory and Its Applications. Vol. 2. New York: John Wiley & Sons.

Filippov, A.I., Akhmetova, O.V. & Zelenova, M.A. 2023. Influence of the position of a horizontal hydraulic fracture on the pressure field in the stratum. Journal of Engineering Physics and Thermophysics 96(2): 301-311.

Ghitany, M.E., Atieh, B. & Nadarajah, S. 2008. Lindley distribution and its application. Mathematics and Computers in Simulation 78(4): 493-506.

Grewal, B. 2016. Higher Engineering Mathematics. 43rd ed. Khanna Publishers.

Johnson, N.L., Kotz, S. & Kemp, A.W. 2005. Univariate Discrete Distributions. 3rd ed. John Wiley & Sons.

Jose, K. & Abraham, B. 2013. A counting process with gumbel interarrival times for modeling climate data. Journal of Environmental Statistics 4(5): 1-13.

Jose, K.K. & Abraham, B. 2011. A count model based on Mittag-Leffler interarrival times. Statistica 71(4): 501-514.

Lindley, D.V. 1958. Fiducial distributions and Bayes' theorem. Journal of the Royal Statistical Society: Series B (Methodological) 20(1): 102-107.

Low, Y.C. & Ong, S.H. 2016. Count distribution for mixture of two exponentials as renewal process duration with applications. AIP Conference Proceedings 1739(1): 020078.

McShane, B., Adrian, M., Bradlow, E.T. & Fader, P.S. 2008. Count models based on Weibull interarrival times. Journal of Business & Economic Statistics 26(3): 369-378.

Moschopoulos, P.G. 1985. The distribution of the sum of independent gamma random variables. Annals of the Institute of Statistical Mathematics 37: 541.

Nadarajah, S. & Chan, S. 2018. Discrete distributions based on interarrival times with application to football data. Communications in Statistics-Theory and Methods 47(1): 147-165.

Nair, M.T. & Abdul, S. 2010. Finite mixture of exponential model and its applications to renewal and reliability theory. Journal of Statistical Theory and Practice 4(3): 367-373.

Ong, S.H., Biswas, A., Peiris, S. & Low, Y.C. 2015. Count distribution for generalized Weibull duration with applications. Communications in Statistics-Theory and Methods 44(19): 4203-4216.

Piessens, R. 1972. A new numerical method for the inversion of Laplace transforms. Journal of the Institute of Mathematics and Its Applications 10: 185-192.

Rausand, M., Barros, A. & Hoyland, A. 2021. System Reliability Theory: Models, Statistical Methods, and Applications. 3rd ed. John Wiley & Sons.

Shortle, J.F., Thompson, J.M., Gross, D. & Harris, C.M. 2018. Fundamentals of Queueing Theory. 5th ed. John Wiley & Sons.

Singh, H., Ong, S.H., Ng, C.M. & Ratnavelu, K. 2024. Poisson-stopped sum Lévy-type processes with application to stochastic modeling of hospital arrivals. Communications in Statistics-Simulation and Computation 54(11): 4712-4725.

Tadikamalla, P.R. 1980. A look at the Burr and related distributions. International Statistical Review 48: 337-344.

Tijms, H.C. 2003. A First Course in Stochastic Models. John Wiley & Sons.

Whitt, W. 2002. Stochastic-Process Limits: An Introduction to Stochastic-Process Limits and Their Application to Queues. Springer.

Winkelmann, R. 1995. Duration dependence and dispersion in count-data models. *Journal of Business and Economic Statistics* 13(4): 467-474.

Zhang, Z. 2018. Renewal sums under mixtures of exponentials. Applied Mathematics and Computation 337: 281-301.

*Corresponding author; email: ngcm@um.edu.my