Proceedings of the 2014 Winter Simulation Conference
A. Tolk, S. Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, eds.

SCALING AND MODELING OF CALL CENTER ARRIVALS

Xiaowei Zhang
L. Jeff Hong
Jiheng Zhang

Department of Industrial Engineering and Logistics Management
The Hong Kong University of Science and Technology
Hong Kong, CHINA

ABSTRACT

The Poisson process has been an integral part of many models for the arrival process to a telephone call centers. However, several publications in recent years suggest the presence of a significant "overdisperson" relative to the Poisson process in real-life call center arrival data. In this paper, we study the overdispersion in the context of "heavy traffic" and identify a critical factor that characterizes the stochastic variability of the arrivals to their averages. We refer to such a factor as the scaling parameter and it potentially has a profound impact on the design of staffing rules. We propose an new model to capture the scaling parameter in this paper.

1 INTRODUCTION

The Poisson process is one of the most widely used models in queueing theory and call center analysis due to its analytical tractability. Its validity is also supported by statistical analysis in Brown et al. (2005), Kim and Whitt (2013a), and Kim and Whitt (2013b). The Poisson model greatly facilitates the associated queueing analysis and produces many insightful results on staffing, including the broadly known "square root safety staffing" principle. This staffing principle stipulates that assuming the agents have a unit service rate, in order to achieve a balance between agent efficiency and service quality the safety staffing level beyond the nominal requirement should be of the order of the square root of the mean arrival rate. The square root form essentially stems from the important fact that variance of the arrivals is of the same order of magnitude as their average, an obvious property of the Poisson process.

In recent years, however, there has been substantial interest in building more sophisticated models for the arrival process to account for certain non-Poisson features, such as overdispersion and autocorrelation, which have been observed in a variety of real-life call center data; see, for example, Jongbloed and Koole (2001) and Avramidis et al. (2004). The models analyzed in these two papers are based on the doubly stochastic Poisson process (DSPP) model proposed by Whitt (1999). It assumes the uncertainty of the arrival rates is determined by a random variable, whose realized value can be interpreted as how busy a day, or a time period of interest, is. Nevertheless, this model is inadequate in capturing the correlation structure of the arrivals. This is essentially because the randomness of the arrival rate is *static*. See Soyer and Tarimcilar (2008) for an extension of this model. Zhang (2013) and Zhang et al. (2013) advocate the use of a *dynamic* DSPP which models the arrival rate as tractable stochastic process and permits a more versatile correlation structure. Another approach is to treat the arrival counts in disjoint time periods rather than the arrival process as the modeling target and attempt to build models for this multivariate random variable; see, for example, Avramidis et al. (2004) and Avramidis et al. (2009).

The uncertainty in the arrival rate obviously increases the overall uncertainty of the call center queueing system and thus may have a significant impact on the evaluation of system performances and the choice of staffing rules; see Gans et al. (2003) for a general discussion on this issue. Assume the arrival rate is of

978-1-4799-7486-3/14/$31.00 ©2014 IEEE

the multiplicative form λG, which is the case for many of the aforementioned doubly stochastic Poisson models, where G is a random variable or a stochastic process. It is then easy to show that in a heavy traffic environment of large λ, the mean of the arrival count in a fixed time period is of the order λ whereas its variance of the order λ^2. Namely, the stochastic variability of the arrival process is enormously amplified by the uncertainty of the arrival rate. Such significant overdispersion violates the assumption that underpins the square root staffing principle, and the appropriate safety staffing level is conceivably much higher; see, for example, Chen and Henderson (2001), Steckley et al. (2005), and Steckley et al. (2009). Numerical solutions that address the staffing problem in the presence of random arrival rates are discussed in Whitt (2006), Gurvich et al. (2010), and Liao et al. (2012), using the model proposed in Whitt (1999).

In this paper, we attempt to address an even more fundamental question: what is the order of the stochastic variability of the arrival process relative to its average? Is it $O(\lambda)$ as for the Poisson model, or $O(\lambda^2)$ as for the DSPP model in Whitt (1999), or something else? We will investigate the overdispersion of the arrival process in a heavy traffic environment and propose a new stochastic model to account for it. The rest of the paper is organized as follows. We provide empirical observations to motivate our investigation in Section 2. In Section 3, we propose our model that incorporates an explicit scaling parameter and further prove certain properties of our model that are consistent with the statistical evidence. Section 4 concludes with a brief discussion on future research. The proofs are collected in the Appendix.

2 MOTIVATION: WHAT IS THE APPROPRIATE SCALING IN HEAVY TRAFFIC ANALYSIS?

Let $(A(t) : t \geq 0)$ denote the arrival process, i.e. $A(t)$ is the cumulative number of arrivals up to time t. In a typical heavy traffic analysis such as the "quality and efficiency driven" (QED) regime (Gans et al. 2003), $A(t)$ is properly scaled and the limiting process is then derived by increasing the mean arrival rate in order to simplify the subsequent queueing analysis. In particular, assuming $\mathbb{E}A(t) = \lambda t$ and $\mathrm{Var}(A(t)) = \Theta(\lambda^p)$ as $\lambda \to \infty$, where the notation $f = \Theta(g)$ means f is bounded both above and below by g up to a constant asymptotically, we study the limit of the scaled arrival process $\frac{A(t)-\mathbb{E}A(t)}{\sqrt{\mathrm{Var}(A(t))}}$. Or equivalent, we are interested in the convergence as follows

$$\frac{A_\lambda(t) - \lambda t}{\lambda^{\frac{p}{2}}} \Rightarrow Z(t) \tag{1}$$

as $\lambda \to \infty$ for some non-degenerate stochastic process $Z(t)$, where we have rewritten $A(t)$ as $A_\lambda(t)$ to emphasize its dependence on λ. For example, if $A(t)$ is a Poisson process with rate λ, then the limit in (1) reads

$$\frac{N(\lambda t) - \lambda t}{\lambda^{\frac{1}{2}}} \Rightarrow B(t)$$

as $\lambda \to \infty$, where $N(t)$ is a Poisson process with unit rate and $B(t)$ is a standard Brownian motion. On the other hand, if $A(t)$ is a DSPP whose arrival rate is of the form λG for some random variable G with $\mathbb{E}G = 1$, then one can show easily via the characteristic function that

$$\frac{N(\lambda G t) - \lambda t}{\lambda} \Rightarrow (G-1)t$$

as $\lambda \to \infty$. Hence, the parameter p in (1) is critical for the limiting process, the sequent queueing analysis and even the design of staffing rules.

In order to identify the parameter p, we conduct statistical analysis on a real-life call center dataset. This dataset is from a large call center of an anonymous bank in U.S. which operates 24/7. It contains phone records in July 2001. (We have also examined the data in other months in 2001 - 2002 and they all give similar findings as we present here in this paper.) We have removed weekends and public holidays since the traffic of phone calls during these days have an obviously different pattern and finally obtain 21 days' records. The left panel of Figure 1 shows the arrival count of each 10-minute time period estimated based on daily average. The “time-of-day” effect facilitates our study of the parameter p since it creates a sequence of scenarios of increasingly heavier traffic (i.e. λ).

Besides the “time-of-day” effect, the arrival process also exhibits significant overdispersion. The actual variance is overwhelmingly larger than what the Poisson model implies. Additionally, it appears that the overdispersion is more significant as the arrival rate increases. We therefore compare the variance of the arrival count of each time period against the mean on the logarithmic scale; see the right panel of Figure 1. It suggests that the order of the variance relative to the mean lies between 1 and 2. Namely, the overdispersion relative to the Poisson process does exist in real-life call center arrivals, but it is not as severe as what the DSPP in Whitt (1999) implies.

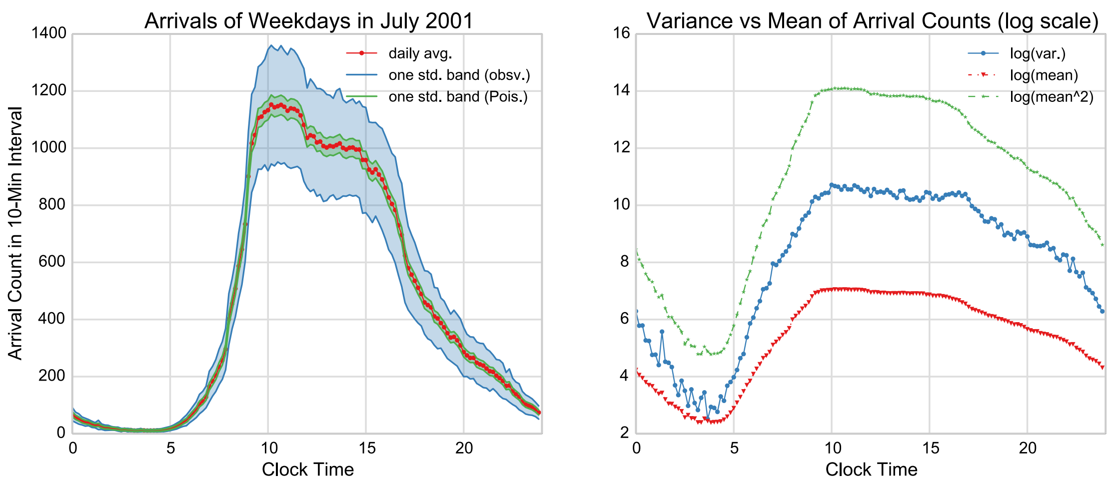

Figure 1: Time-of-day effect and overdispersion of the arrival process. Left: the mean and the one standard deviation band of the arrival count in 10-minute time periods. Right: the mean and the variance of the arrival count in 10-minute time periods on the logarithmic scale.

We take a heuristic perspective in this paper for estimating the parameter p. Note that $\mathrm{Var}(A(t)) \sim c(\mathbb{E}A(t))^p$ for some constant $c > 0$, so $\log(\mathrm{Var}(A(t))) \sim p\log(\mathbb{E}A(t))$ as $\lambda \to \infty$. Hence, we assume the following linear relationship

$$\log(\mathrm{Var}(A(t))) = p\log(\mathbb{E}A(t)) + c, \tag{2}$$

and estimate p via a linear regression. Conceivably, this estimation approach is not rigorous and the result depends on the choice of t, i.e. the length of the time period. But we are more interested in demonstrating the existence of p than in estimating it accurately. Figure 2 shows the linear relationship between the variance and the mean of the arrival count on the logarithmic scale. Obviously, the linear model (2) fits the data remarkably well with $R^2 = 0.99$. The parameter p is estimated as $\hat{p} = 1.63$, which verifies our speculation that the level of overdispersion of the arrival process may lie between those implied by the Poisson model and the DSPP model in Whitt (1999). It also motivates us to develop a new stochastic model that explicitly captures this phenomenon.

3 DOUBLY STOCHASTIC POISSON MODEL WITH SCALING PARAMETER

We model the arrival process $A(t)$ as a DSPP with the arrival rate $X(t)$, i.e.

$$A(t) = N\left(\int_0^t X(s)\,\mathrm{d}s\right),$$

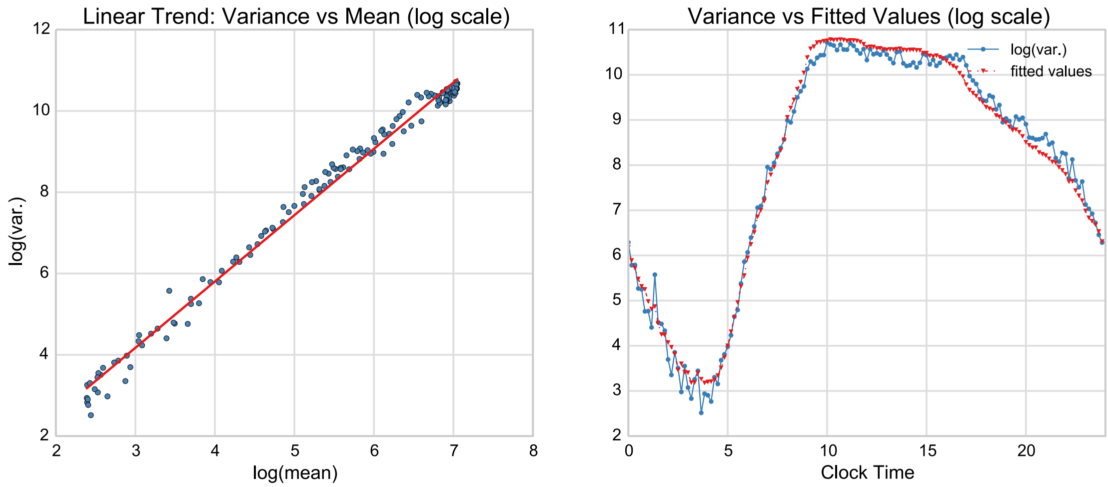

Figure 2: Fitting result of the linear model (2). Left: almost perfect linear relationship between the variance and the mean of the arrival count in 10-minute time periods on the logarithmic scale. Right: comparison between the variance and the fitted curve.

where $N(\cdot)$ is a Poisson process with unit rate. Further, we model $X(t)$ as a stochastic process that satisfies the following stochastic differential equation (SDE)

$$\mathrm{d}X(t) = \kappa(\lambda - X(t))\,\mathrm{d}t + \sigma\lambda^{\alpha}X(t)^{\frac{1}{2}}\,\mathrm{d}B(t), \tag{3}$$

where κ, λ, σ, and α are constants and $B(t)$ is a standard Brownian motion that is independent of $N(t)$. In fact, the SDE (3) is a reparameterized Cox-Ingersoll-Ross (CIR) process. We choose this model because it is positive and fairly tractable, and has a stationary distribution. The parameter λ is the long-run average of the process so this is a stationary model but one can easily modify this model to incorporate predictable time-varying patterns such as the time-of-day effect. We refer to α as the *scaling* parameter because it controls the order of the variance of the arrival process relative to its mean as will be shown in Theorem 1. Intuitively, the volatility term of the SDE (3) is $\sigma\lambda^{\alpha}X(t)^{\frac{1}{2}}$, which is roughly equal to $\sigma\lambda^{\alpha+\frac{1}{2}}$ since the long-run average of $X(t)$ is λ. It follows that the contribution to the variance of the arrival process $A(t)$ from the arrival rate is of the order $\lambda^{2\alpha+1}$. So it suffices to assume $\alpha \in (0, \frac{1}{2})$ so that the overdispersion lies in the range that reconciles with our speculation and empirical evidence.

Lemma 1 (Cox et al. 1985) The process $X(t)$ has a unique stationary distribution π, which is a gamma distribution with mean λ and variance $\frac{\sigma^2\lambda^{2\alpha+1}}{2\kappa}$.

Theorem 1 Suppose $X(t)$ is initialized with the stationary distribution π. Then,

$$\mathbb{E}_{\pi}A(t) = \lambda t, \tag{4}$$

and

$$\mathrm{Var}_{\pi}(A(t)) = t\left[\lambda + \frac{\sigma^2\lambda^{2\alpha+1}}{\kappa^2}\left(1 - \frac{1-e^{-\kappa t}}{\kappa t}\right)\right]. \tag{5}$$

We then immediately have the following result regarding the order of magnitude of $\mathrm{Var}(A(t))$ relative to $\mathbb{E}A(t)$.

Corollary 1 Suppose $X(t)$ is initialized with the stationary distribution π and $\alpha \in (0, \frac{1}{2})$. Then, $\mathrm{Var}_{\pi}(A(t))) = \Theta(\lambda^{2\alpha+1})$ as $\lambda \to \infty$.

Therefore, the relationship between the parameter p defined in Section 2 and the scaling parameter α in our model is simply $p = 2\alpha + 1$. Moreover, by assuming $\alpha \in (0, \frac{1}{2})$, our model well captures the level of stochastic variation in the arrival process relative to its average observed in the data.

We now turn to the more fundamental property with regard to the scaling scheme in heavy traffic analysis, namely the convergence (1). Thanks to the analytical tractability of the CIR process as well as the doubly stochastic Poisson structure, we can show the convergence (in the sense of marginal distribution) of the scaled arrival process under our model as follows. Note that we write $A_\lambda(t)$ when it is necessary to emphasize its dependence on λ.

Theorem 2 Suppose $X(t)$ is initialized with the stationary distribution π and $\alpha \in (0, \frac{1}{2})$. Then, for any given $t > 0$,

$$\frac{A_\lambda(t) - \lambda t}{\lambda^{\alpha + \frac{1}{2}}} \Rightarrow \int_0^t u(s)\,\mathrm{d}s, \tag{6}$$

as $\lambda \to \infty$, where $u(t)$ is an Ornstein-Uhlenbeck (OU) process

$$\mathrm{d}u(t) = -\kappa u(t)\,\mathrm{d}t + \sigma \mathrm{d}B(t),$$

with initial distribution being its unique stationary distribution, i.e. normal distribution with mean 0 and variance $\frac{\sigma^2}{2\kappa}$.

Note that $\int_0^t u(s)\,\mathrm{d}s$ has normal distribution with

$$\mathbb{E}_\phi\left(\int_0^t u(s)\,\mathrm{d}s\right) = 0 \quad \text{and} \quad \mathrm{Var}_\phi\left(\int_0^t u(s)\,\mathrm{d}s\right) = \frac{\sigma^2}{\kappa^2}\left(t - \frac{1}{\kappa}(1 - e^{-\kappa t})\right),$$

where ϕ is the stationary distribution of $u(t)$. See Section A.2 for the proof. Hence, for any given $t > 0$,

$$\frac{A_\lambda(t) - \lambda t}{\lambda^{\alpha + \frac{1}{2}}\sqrt{\mathrm{Var}_\phi\left(\int_0^t u(s)\,\mathrm{d}s\right)}} \Rightarrow \mathcal{N}(0,1),$$

as $\lambda \to \infty$, where $\mathcal{N}(0,1)$ denotes a standard normal random variable. On the other hand, Equation (5) implies that if $\alpha \in (0, \frac{1}{2})$,

$$\mathrm{Var}_\pi(A_\lambda(t)) \sim \frac{\sigma^2 \lambda^{2\alpha+1}}{\kappa^2}\left(t - \frac{1}{\kappa}(1 - e^{-\kappa t})\right) = \lambda^{2\alpha+1}\mathrm{Var}_\phi\left(\int_0^t u(s)\,\mathrm{d}s\right)$$

as $\lambda \to \infty$. Consequently, we have the following weak convergence.

Corollary 2 Suppose $X(t)$ is initialized with the stationary distribution π and $\alpha \in (0, \frac{1}{2})$. Then, for any given $t > 0$,

$$\frac{A_\lambda(t) - \mathbb{E}_\pi A_\lambda(t)}{\mathrm{Var}_\pi(A_\lambda(t))} \Rightarrow \mathcal{N}(0,1), \tag{7}$$

as $\lambda \to \infty$, where $\mathcal{N}(0,1)$ is a standard normal random variable.

Interestingly, the weak convergence (7) is also well supported by the data. Since we use the arrival counts in 10-minute time periods for our study, there are $6 \times 24 = 144$ such time periods in a day, among which we select five with increasingly larger average arrival counts. For each of the five selected periods, we examine the distribution of its arrival count in the 21 days of our dataset and apply the Gaussian kernel smoothing method (see e.g. Hastie et al. 2009) to estimate its probability density function. We then can see a clear sign of convergence of such estimated densities to the standard normal density as the mean arrival rate increases; see Figure 3

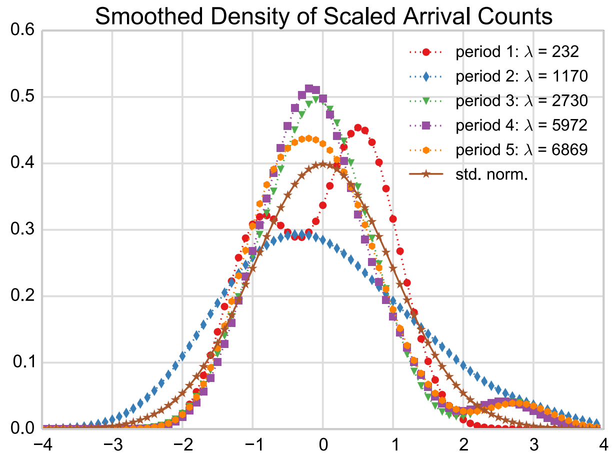

Figure 3: Evidence in data of the convergence (7) of the scaled arrival process. Five 10-minute time periods are selected and ranked in the ascending order based their average arrival counts. We use λ to denote the average number of arrivals per hour. The density functions of the arrival counts are estimated via the Gaussian kernel smoothing method.

4 CONCLUDING REMARKS

In this paper, we have investigated the scaling scheme of the arrival process which is fundamental in typical heavy traffic analysis of queueing systems. By a careful statistical analysis, we have identified that the level of the stochastic variation of the arrival process relative to its average is neither as low as that in the Poisson process, nor as high as that in the DSPP proposed by Whitt (1999) and widely used in the literature. It in fact lies somewhere in between. To accommodate this new and significant observation, we have developed a stochastic model that allows an explicit control over the level of overdispersion in a heavy traffic environment. Our model is fairly tractable and we have proved two important properties to demonstrate that our model aligns with the data very well.

Note that our arrival model is a DSPP with arrival rate being a (reparameterized) CIR process. Its simulation is well studied and can be implemented easily; see Giesecke and Kim (2007) and Giesecke et al. (2011). Hence, our arrival model can be readily used as a more accurate input model to drive simulation of complex queueing systems.

We believe the presented results in this paper would potentially lead to many interesting open questions. Here are some examples. First of all, given the excessive random fluctuations in the arrival process, the necessary staffing level should be higher than it would be for Poisson arrivals in order to achieve similar service quality. Indeed, the non-conventional scaling scheme in Theorem 2 suggests that the safety staffing level should conceivably be of the order $\Theta(\lambda^{\alpha+\frac{1}{2}})$ for our arrival model, instead of $\Theta(\lambda^{\frac{1}{2}})$ as for the Poisson arrival model. However, the associated heavy traffic analysis requires that the scaled arrival process should have a stronger version of the convergence than the "convergence in marginal distribution" in Theorem 2. In particular, it should converge in Skorohod topology (see e.g. Whitt 2002). This is part of our on-going research.

Moreover, given its critical role in determining the proper staffing level, it is of great interest and importance to develop a credible approach to efficiently estimate the parameter α. (The linear regression estimation (2) is heuristic and can not be used in any rigorous sense.) One possible approach is to adopt the Bayesian inference via Markov chain Monte Carlo; see e.g. Gelman et al. (2004) for an extensive coverage on this topic, or Zhang (2013) for the discussion on a similar problem.

ACKNOWLEDGMENTS

The authors thank Avi Mandelbaum and the SEE Center at Technion for providing the data.

A APPENDICES

The proofs of both Theorem 1 and Theorem 2 heavily rely on the explicit calculation of the Laplace transform of $A(t)$, thanks to the analytical tractability of the CIR process.

Lemma 2 For any $\theta > 0$,

$$\mathbb{E}_x \exp\left(-\theta \int_0^t X(s)\,\mathrm{d}s\right) = \exp[-f(\theta) - g(\theta)x],$$

where $\mathbb{E}_x(\cdot) \triangleq \mathbb{E}(\cdot|X(0) = x)$,

$$f(\theta) = -\frac{2\kappa\lambda^{1-2\alpha}}{\sigma^2} \log\left(\frac{2\gamma e^{(\gamma+\kappa)t/2}}{(\gamma+\kappa)e^{\gamma t} + \gamma - \kappa}\right), \tag{8}$$

and

$$g(\theta) = \frac{2\theta(e^{\gamma t} - 1)}{(\gamma+\kappa)e^{\gamma t} + \gamma - \kappa}, \tag{9}$$

where $\gamma = \sqrt{\kappa^2 + 2\theta\sigma^2\lambda^{2\alpha}}$.

Proof. See Equation (3.76) on page 129 of Glasserman (2003). □

Lemma 3 For any $\theta > 0$,

$$\psi(\theta) \triangleq \mathbb{E}_\pi \exp(-\theta A(t)) = \exp\left[-f(1-e^{-\theta})\right] \cdot \left[1 + \frac{\sigma^2\lambda^{2\alpha}}{2\kappa} \cdot g(1-e^{-\theta})\right]^{-\frac{2\kappa\lambda^{1-2\alpha}}{\sigma^2}}, \tag{10}$$

where the functions f and g are given by (8) and (9).

Proof. By the doubly stochastic Poisson structure,

$$\psi(\theta) = \mathbb{E}_\pi\left[\mathbb{E}\left(\exp[-\theta A(t)]|X(s) : 0 \le s \le t\right)\right] = \mathbb{E}_\pi \exp\left(-(1-e^{-\theta})\int_0^t X(s)\mathrm{d}s\right).$$

It the follows from Lemma 2 that

$$\begin{aligned}
\psi(\theta) &= \int_0^\infty \mathbb{E}_x \exp\left(-(1-e^{-\theta})\int_0^t X(s)\mathrm{d}s\right)\pi(\mathrm{d}x) \\
&= \int_0^\infty \exp\left[f\left(1-e^{-\theta}\right) - g\left(1-e^{-\theta}\right)x\right]\pi(\mathrm{d}x) \\
&= \exp\left[-f\left(1-e^{-\theta}\right)\right] \cdot \left[1 + \frac{\sigma^2\lambda^{2\alpha}}{2\kappa} \cdot g\left(1-e^{-\theta}\right)\right]^{-\frac{2\kappa\lambda^{1-2\alpha}}{\sigma^2}},
\end{aligned}$$

since π is a gamma distribution. □

A.1 Proof of Theorem 1

The calculation of $\mathbb{E}A(t)$ is a simple implication of the doubly stochastic Poisson structure.

$$\mathbb{E}_\pi A(t) = \mathbb{E}_\pi\left[\mathbb{E}\left(A(t)|X(s), 0 \le s \le t\right)\right] = \mathbb{E}_\pi\left[\int_0^t X(s)\,\mathrm{d}s\right] = \int_0^t \mathbb{E}_\pi X(s)\,\mathrm{d}s = \lambda t.$$

The derivation of $\mathbb{E}A^2(t)$, however, utilizes the fact that $\mathbb{E}A^2(t) = \psi''(0)$, where $\psi(\theta) = \mathbb{E}_\pi \exp(-\theta A(t))$ is given by (10). The calculation is straightforward but fairly lengthy and we omit the details.

A.2 Proof of Theorem 2

Note that the convergence (6) is about the marginal distribution of the scaled arrival process. So it suffices to show the convergence of the corresponding Laplace transform, namely

$$\mathbb{E}_\pi \exp\left(-\theta\lambda^{-\alpha-\frac{1}{2}}(A(t)-\lambda t)\right) \to \mathbb{E}_\phi \exp\left(-\theta\int_0^t u(s)\,\mathrm{d}s\right) \tag{11}$$

as $\lambda \to \infty$, where ϕ is the stationary distribution of the OU process $u(t)$. Note that the integrated OU process is normally distributed, so

$$\mathbb{E}\exp\left(-\theta\int_0^t u(s)\,\mathrm{d}s\right) = \exp\left[-\theta\mathbb{E}\left(\int_0^t u(s)\,\mathrm{d}s\right) + \frac{\theta^2}{2}\mathrm{Var}\left(\int_0^t u(s)\,\mathrm{d}s\right)\right].$$

By Equation (3.51) and (3.54) on page 113 of Glasserman (2003),

$$\mathbb{E}\left[\exp\left(-\theta\int_0^t u(s)\,\mathrm{d}s\right)|u(0)\right] = \exp\left[-\frac{\theta}{\kappa}(1-e^{-\kappa t})u(0) + \frac{\theta^2\sigma^2}{2\kappa^2}\left(t + \frac{1}{2\kappa}(1-e^{-2\kappa t}) - \frac{2}{\kappa}(1-e^{-\kappa t})\right)\right].$$

It then follows that, letting U is a random variable distributed as ϕ (i.e. normal distribution with mean 0 and variance $\frac{\sigma^2}{2\kappa}$),

$$\begin{aligned}\mathbb{E}_\phi \exp\left(-\theta\int_0^t u(s)\,\mathrm{d}s\right) =& \mathbb{E}\exp\left[-\frac{\theta}{\kappa}(1-e^{-\kappa t})U + \frac{\theta^2\sigma^2}{2\kappa^2}\left(t + \frac{1}{2\kappa}(1-e^{-2\kappa t}) - \frac{2}{\kappa}(1-e^{-\kappa t})\right)\right] \\ =& \exp\left[\frac{\theta^2\sigma^2}{4\kappa^3}(1-e^{-\kappa t})^2 + \frac{\theta^2\sigma^2}{2\kappa^2}\left(t + \frac{1}{2\kappa}(1-e^{-2\kappa t}) - \frac{2}{\kappa}(1-e^{-\kappa t})\right)\right] \\ =& \exp\left[\frac{\theta^2\sigma^2}{2\kappa^2}\left(t - \frac{1}{\kappa}(1-e^{-\kappa t})\right)\right].\end{aligned} \tag{12}$$

On the other hand, Lemma 3 implies that

$$\mathbb{E}_\pi \exp\left(-\theta\lambda^{-\alpha-\frac{1}{2}}(A(t)-\lambda t)\right) = \psi(-\theta\lambda^{-\alpha-\frac{1}{2}}) \cdot \exp(\theta t\lambda^{\frac{1}{2}-\alpha}). \tag{13}$$

Hence, Equation (11) is equivalent to

$$\psi(-\theta\lambda^{-\alpha-\frac{1}{2}}) \cdot \exp(\theta t\lambda^{\frac{1}{2}-\alpha}) \to \exp\left[\frac{\theta^2\sigma^2}{2\kappa^2}\left(t - \frac{1}{\kappa}(1-e^{-\kappa t})\right)\right]$$

as $\lambda \to \infty$, because of Equations (12) and (13), which can be shown by an elementray calcualtion and we omit the details.

REFERENCES

Avramidis, A. N., N. Channouf, and P. L'Ecuyer. 2009. "Efficient Correlation Matching for Fitting Discrete Multivariate Distributions with Arbitrary Marginals and Normal-Copula Dependence". *INFORMS Journal on Computing* 21 (1): 88–106.

Avramidis, A. N., A. Deslauriers, and P. L'Ecuyer. 2004. "Modeling Daily Arrivals to a Telephone Call Center". *Management Science* 50 (7): 896–908.

Brown, L., N. Gans, A. Mandelbaum, A. Sakov, H. Shen, S. Zeltyn, and L. Zhao. 2005. "Statistical Analysis of a Telephone Call Center: A Queueing-science Perspective". *J. Amer. Statist. Assoc.* 100 (1): 36–50.

Chen, B. P. K., and S. G. Henderson. 2001. "Two Issues in Setting Call Centre Staffing Levels". *Annals of Operations Research* 108:175–192.

Cox, J. C., J. E. Ingersoll, and S. A. Ross. 1985. "A Theory of the Term Structure of Interest Rates". *Econometrica* 53 (2): 385–407.

Gans, N., G. Koole, and A. Mandelbaum. 2003. "Telephone Call Centers: Tutorial, Review, and Research Prospects". *Manufacturing & Service Operations Management* 5 (2): 79–141.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 2004. *Bayesian Data Analysis*. 2 ed. Chapman and Hall/CRC.

Giesecke, K., and B. Kim. 2007. "Estimating tranche spread by loss process simulation". In *Proceedings of the 2007 Winter Simulation Conference*, edited by S. G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J. D. Tew, and R. R. Barton, 967–975: Institute of Electrical and Electronics Engineering, Inc.

Giesecke, K., B. Kim, and S. Zhu. 2011. "Monte Carlo Algorithms For Default Timing Problems". *Management Science* 57:2115–2129.

Glasserman, P. 2003. *Monte Carlo Methods in Financial Engineering*. Springer.

Gurvich, I., J. Luedtke, and T. Tezcan. 2010. "Staffing Call Centers with Uncertain Demand Forecasts: A Chance-Constrained Optimization Approach". *Management Science* 56 (7): 1093–1115.

Hastie, T., R. Tibshirani, and J. Friedman. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2 ed. Springer.

Jongbloed, G., and G. Koole. 2001. "Managing Uncertainty in Call Centers Using Poisson Mixtures". *Appl. Stochastic Models Bus. Indust.* 17:307–318.

Kim, S.-H., and W. Whitt. 2013a. "Are Call Center and Hospital Arrivals Well Modeled by Nonhomogeneous Poisson Processes?". To appear in *Manufacturing & Service Operations Management*.

Kim, S.-H., and W. Whitt. 2013b. "Choosing Arrival Process Models for Service Systems: Tests of a Nonhomogeneous Poisson Process". To appear in *Naval Research Logistics*.

Liao, S., G. Koole, C. van Delft, and O. Jouini. 2012. "Staffing a Call Center with Uncertain Non-stationary Arrival Rate and Flexibility". *OR Spectrum* 34:691–721.

Soyer, R., and M. M. Tarimcilar. 2008. "Modeling and Analysis of Call Center Arrival Data: A Bayesian Approach". *Management Science* 54 (2): 266–278.

Steckley, S. G., S. G. Henderson, and V. Mehrotra. 2005. "Performance Measures For Service Systems With A Random Arrival Rate". In *Proceedings of the 2005 Winter Simulation Conference*, edited by M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, 566–575.

Steckley, S. G., S. G. Henderson, and V. Mehrotra. 2009. "Forecast Errors in Service Systems". *Probability in the Engineering and Informational Sciences* 23:305–332.

Whitt, W. 1999. "Dynamic Staffing in a Telephone Call Center Aiming to Immediately Answer All Calls". *Oper. Res. Lett.* 24:205–212.

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

Whitt, W. 2006. "Staffing a Call Center with Uncertain Arrival Rate and Absenteeism". *Production and Operations Management* 15 (1): 88–102.

Zhang, X. 2013. “A Bayesian Approach for Modeling and Analysis of Call Center Arrivals”. In *Proceedings of the 2013 Winter Simulation Conference*, edited by R. Pasupathy, S.-H. Kim, A. Tolk, R. Hill, and M. E. Kuhl.

Zhang, X., L. J. Hong, and P. W. Glynn. 2013. “Timescales in Modeling Call Center Arrivals”. Working paper.

AUTHOR BIOGRAPHIES

XIAOWEI ZHANG is an Assistant Professor in the Department of Industrial Engineering and Logistics Management at Hong Kong University of Science and Technology. He received his Ph.D. in Operations Research from Stanford University in 2011. He is a member of INFORMS and his research interests include rare event simulation and stochastic modeling in service engineering. His email address is xiaoweiz@ust.hk.

L. JEFF HONG is a professor in the Department of Industrial Engineering and Logistics Management at the Hong Kong University of Science and Technology. His research interests include Monte Carlo method, financial engineering, and stochastic optimization. He is currently an associate editor for *Operations Research*, *Naval Research Logistics* and *ACM Transactions on Modeling and Computer Simulation*. His email address is hongl@ust.hk.

JIHENG ZHANG is an Assistant Professor in the Department of Industrial Engineering and Logistics Management at Hong Kong University of Science and Technology. He received both Ph.D. and M.S. degrees in operations research from Georgia Institute of Technology (2009, 2005), M.S. degree in mathematics from the Ohio State University (2003), and B.S. degree in mathematics in Nanjing University (2002). His research interests are in the areas of service operations management, queueing system and applied probability. His email address is jiheng@ust.hk.