An Independent Development of the Theory of $M_t/G/\infty$ Queues

Alon Jacobson

November 26, 2022

1 Introduction

In this document, I describe an independent development of the theory of $M_t/G/\infty$ queues, which is a type of queue from queueing theory. The development theory was motivated by a desire to mathematically model the movement of people into and out of a park. When I developed this theory, I was not familiar with queueing theory. Only later when I learned about queueing theory, did I discover that my work has mostly already been done, for example in the works [1, 2, 3]. I still think that my work is a valuable and interesting exposition.

This work is motivated by the following question: If you know both how many people enter a building per minute and the probability distribution of the number of minutes that someone spends in the building, what can you infer about the number of people that are in the building?

To generalize, say that we have a "state" we are interested in, and there are objects that transition into and out of this state; an object can either be in or not be in this state at a given point in time. The objects could be, for example, people in a building, cars on the highway, squirrels in your back yard, fish in a river, molecules in some subset of space, computers in a network. We are interested in, among other things, how to mathematically describe and simulate how many objects there are in that state at a given point in time. We would hopefully also like to make inferences from data using this theory that would not have been possible before. In this paper I will refer to the things that transition in and out of the state as people.

1.1 Notation and Key Results

Take the beginning of the simulation to be time 0. The number of people in the state at different points in time, call this the *count process*, can be thought of as a stochastic process $\{\tilde{n}(t), t \geq 0\}$ where $\tilde{n}(t)$ is a non-negative integer. When someone transitions into the state, call this an *entry*. When someone transitions out of the state, call this an *exit*.

The simulation consists of some number of "stays," where a stay is described by when someone entered and when they exited. Let K be the total number of stays. The ith stay is described by two numbers: t_{1i}, the time of the ith entrance, and t_{2i}, the exit time corresponding to the ith stay. The entrance times together are a sequence $(t_{1i})_{i=1}^{K}$, and the exit times together are a sequence $(t_{2i})_{i=1}^{K}$. Note that the exit times, in contrast to the entrance times, are not necessarily in order of when they happen. The portion of time that the person of the ith stay takes up is described by an interval $[t_1, t_2)$ (closed on the left, open on the right). This convention is so the math all works out. The sequence of stay intervals is $(\mathcal{S}_i)_{i=1}^{K}$ where $\mathcal{S}_i = [t_{1i}, t_{2i})$. The value of $\widetilde{n}(t)$ is defined as the number of stays that contain t:

$$\widetilde{n}(t) = \sum_{i=1}^{K} [t \in \mathcal{S}_i]$$

where $[\cdot]$ is the *iverson bracket*, which is 1 if the statement inside it is true and is 0 if the statement inside it is false:

$$[P] = \begin{cases} 1 & \text{if } P \text{ is true} \\ 0 & \text{otherwise.} \end{cases}$$

The Iverson bracket has some nice properties. The iverson bracket of the logical "and" of two statements can be decomposed by $[P \wedge Q] = [P][Q]$, and the iverson bracket of the negation of a statement is $[\neg P] = 1 - [P]$.

The total number of entries so far can be described by a counting process (that is, a stochastic process with values that are non-negative, integer, and non-decreasing, that describes the number of events over time) called the *entry process*, $\{\widetilde{\Lambda}(t), t \geq 0\}$. $\widetilde{\Lambda}(t)$ is the number of entrances that have happened up to and including time t, and is defined as

$$\widetilde{\Lambda}(t) = \sum_{i=1}^{K} [t_{1i} \leq t].$$

The tilde in $\widetilde{\Lambda}$ indicates that $\widetilde{\Lambda}$ is a stochastic process, and I will use this notation consistently. For times t_1, t_2 where $t_1 \leq t_2$, denote $\tilde{\Lambda}(t_1, t_2) = \tilde{\Lambda}(t_2) - \tilde{\Lambda}(t_1)$, which is the number of entrances that occurred in the interval $(t_1, t_2]$. Assume that the entry process is a non-homogeneous Poisson process with intensity function given by some function $\lambda : [0, \infty) \to [0, \infty)$, called the *entry rate*. (Note that whenever I write that a function maps from $[0, \infty)$, that actually means that when convenient, it maps from $\mathbb{R}$ and that its output is 0 for negative arguments.) Define the *cumulative entry rate* $\Lambda(t)$ by $\Lambda(t) = \int_0^t \lambda(s)\, ds$, and denote $\Lambda(t_1, t_2) = \Lambda(t_2) - \Lambda(t_1) = \int_{t_1}^{t_2} \lambda(s)\, ds$. Since $\tilde{\Lambda}$ is a Poisson process, we have $\widetilde{\Lambda}(t_1, t_2) \sim \text{Pois}(\Lambda(t_1, t_2))$. Recall that this implies that $\mathbb{E}(\widetilde{\Lambda}(t_1, t_2)) = \mathbb{V}(\widetilde{\Lambda}(t_1, t_2)) = \Lambda(t_1, t_2)$.

The total number of exits so far can similarly be described by a counting process, called the *exit process*, $\{\widetilde{E}(t), t \geq 0\}$. It is analogously defined as

$$\widetilde{E}(t) = \sum_{i=1}^{K} [t_{2i} \leq t].$$

Knowing both a realization of the entry process and the duration of time that each person - corresponding to each "point" in the entry process - spends in the state, you can construct the exit process by displacing each point in the entry process (positively) by the duration corresponding to that point. In other words, if someone enters at time t and they spend a duration of τ, then they will exit at time $t + \tau$.

The duration of time that someone spends in the state given that they entered at time t is described by a random variable T_t which is assumed to be a continuous random variable, and we will assume that each displacement is independent. Denote the probability density function of T_t as f_{T_t}, and its survival function as S_{T_t} (the survival function means $S_{T_t}(\tau) = \Pr(T_t > \tau)$, and we have $f_{T_t}(\tau) = -S'_{T_t}(\tau)$). Knowing both the entry process $\{\widetilde{\Lambda}(t), t \geq 0\}$, and the probability distribution of the duration of time T_t, the exit process can be obtained by displacing each point t_k in $\widetilde{\Lambda}$ by a random duration from T_{t_k}. It turns out that, given that $\widetilde{\Lambda}$ is a Poisson process, $\widetilde{E}$ is also a Poisson process, with intensity function given by

$$e(t) = \int_0^t \lambda(\tau) f_{T_\tau}(t - \tau)\, d\tau.$$

The proof follows from the displacement theorem, which says that the random independent displacement of points of a Poisson process forms another Poisson process. I am not mathematically equipped to prove the displacement theorem itself, but the displacement theorem is stated in the Wikipedia article for Poisson Point Process and the result follows from that statement.

Now let's make the assumption that T_t is equal in distribution to a random variable T (with PDF f_T or simply f, and survival function S_T or simply S) for all t, in other words, the probability distribution of the duration of time that someone spends is time-invariant. This simplifying assumption of time-invariance will be held from now on. The formula for $e(t)$ then simplifies to:

$$e(t) = \int_0^t \lambda(\tau) f_T(t-\tau)\, d\tau = (f_T * \lambda)(t) = \mathbb{E}[\lambda(t-T)] \tag{1}$$

where $*$ denotes convolution: the convolution of two functions f and g, $f * g$, is defined by:

$$(f*g)(t) = \int_{-\infty}^{\infty} f(\tau)g(t-\tau)\, d\tau = \int_{-\infty}^{\infty} f(t-\tau)g(\tau)\, d\tau.$$

If f and g are supported only on $[0,\infty)$ (i.e., zero for negative arguments), then the formula becomes

$$(f*g)(t) = \int_0^t f(\tau)g(t-\tau)\, d\tau \qquad \text{for } f, g : [0,\infty) \to \mathbb{R}.$$

Note the fact that with a convolution integral, you can do a change of variables which means that which function gets the $t-\tau$ and which one gets the τ does not matter. Also note that when I write a convolution integral of the form $\int_0^t f(\tau)g(t-\tau)\, d\tau$, I can also write it as $\int_0^\infty f(\tau)g(t-\tau)\, d\tau$ because the functions that we're dealing with here are supported on $[0,\infty)$.

Similar to how we did with the entry process, now that we know that $\widetilde{E}$ is a Poisson process, denote the *cumulative exit rate* by $E(t) = \int_0^t e(s)\, ds$, $E(t_1,t_2) = \int_{t_1}^{t_2} e(s)\, ds = E(t_2) - E(t_1)$, the number of exits in time interval $(t_1,t_2]$ as $\widetilde{E}(t_1,t_2) = \widetilde{E}(t_2) - \widetilde{E}(t_1)$, and we have that $\widetilde{E}(t_1,t_2) \sim \text{Pois}(E(t_1,t_2))$. It turns out that $E(t)$ can be expressed as:

$$E(t) = (\Lambda * f_T)(t) = \int_0^t \Lambda(\tau) f_T(t-\tau)\, d\tau = \mathbb{E}[\Lambda(t-T)]. \tag{2}$$

Going back to the count process $\widetilde{n}(t)$, it is easy to see that the number of people in the state equals the number of entrances so far minus the number of exits so far (although an intuitive fact, this will be proven later):

$$\widetilde{n}(t) = \widetilde{\Lambda}(t) - \widetilde{E}(t). \tag{3}$$

Now we introduce a function $n(t)$, which gives the expected number of people at time t; $n(t) := \mathbb{E}(\widetilde{n}(t))$. Taking expectations of both sides of (3), we get

$$n(t) = \Lambda(t) - E(t). \tag{4}$$

If we substitute (2) into this we find

$$n(t) = \Lambda(t) - \mathbb{E}[\Lambda(t-T)] = \mathbb{E}[\Lambda(t) - \Lambda(t-T)] = \mathbb{E}[\Lambda(t-T, t)]. \tag{5}$$

Differentiating both sides of (4), we get

$$n'(t) = \lambda(t) - e(t), \tag{6}$$

a formula that makes sense: loosely, the change in the number of people per unit time is the number of entrances per unit time minus the number of exits per unit time.

It turns out that $\widetilde{n}(t) \sim \text{Pois}(n(t))$. Although $\widetilde{n}(t) \sim \text{Pois}(n(t))$, it is not the case that $\widetilde{n}$ is a Poisson process; this fact is obvious once you notice that $\widetilde{n}$ is not necessarily non-decreasing.

An elegant formula for $n(t)$ is

$$n(t) = \int_0^\infty S_T(\tau)\lambda(t-\tau)\,d\tau = \int_0^t S_T(\tau)\lambda(t-\tau)\,d\tau = (\lambda * h)(t) \tag{7}$$

where $h(\tau)$ equals $S_T(\tau)$ for non-negative arguments and is zero for negative arguments; we need to introduce this function in order to write $n(t)$ as a convolution as defined earlier, since $S(\tau) = 1$ for $\tau < 0$.

The process of producing $n(t) = (\lambda * h)(t)$ from $\lambda(t)$ and $h(t)$ can be viewed as a linear time-invariant system (LTI); $\lambda(t)$ is the input to the system, $h(t)$ is the impulse response, and $n(t)$ is the output of the system. Similarly, the process of producing the exit rate can be viewed as a LTI, as well as for the cumulative exit rate. In fact, many of the functions that will be introduced can be expressed as convolutions of an entry-rate-related function - either the entry rate λ or the cumulative entry rate Λ, and a function regarding the distribution of durations - either f_T or h.

One last piece of terminology is what is called the *cumulative expected count*, defined by:

$$N(t) = \int_0^t n(s)\,ds = (\Lambda * h)(t) = \mathbb{E}\left[\int_{t-T}^t \Lambda(s)\,ds\right] = \mathbb{E}[\not{\Lambda}(t-T,t)]$$

where

$$\not{\Lambda}(t_1,t_2) = \int_{t_1}^{t_2} \Lambda(s)\,ds.$$

This function $N(t)$ is used in the analysis of the estimator for $\mathbb{E}[T]$ discussed below.

1.2 Difference between counts at different times

One might be interested in the unconditional (that is, not conditioned on $\widetilde{n}(t_1)$ or $\widetilde{n}(t_2)$, or anything) probability distribution of $\widetilde{n}(t_2) - \widetilde{n}(t_1)$, that is the probability distribution of the change in the number of people from time t_1 to t_2. This distribution is given by

$$\widetilde{n}(t_2) - \widetilde{n}(t_1) \sim \text{Skellam}(\overline{N}_\lambda, \overline{N}_e)$$

where Skellam means the Skellam distribution, the difference of two statistically independent Poisson-distributed random variables: $\text{Skellam}(\mu_1, \mu_2)$ is the probability distribution of the difference $N_1 - N_2$ where $N_1 \sim \text{Pois}(\mu_1)$, $N_2 \sim \text{Pois}(\mu_2)$, and N_1 and N_2 are statistically independent; parameters $\overline{N}_\lambda$ and $\overline{N}_e$ are given by:

$$\overline{N}_\lambda = \int_0^{t_2-t_1} S_T(\tau)\lambda(t_2-\tau)\,d\tau, \quad \overline{N}_e = n(t_1) - \int_{t_2-t_1}^{t_2} S_T(\tau)\lambda(t_2-\tau)\,d\tau.$$

1.3 Steady-state

Something of interest is the limiting, steady-state behavior of this system. By “steady state,” I mean that $\lambda(t)$ is constant and has been constant for infinitely long; $\lambda(t) = \lambda$. In this case, we have $e(t) = \lambda(t) = \lambda$, and also we have that $n(t) = \lambda\mathbb{E}[T]$. More formally, this is saying that assuming λ is bounded and continuous with a finite limit at ∞, and $\mathbb{E}[T]$ exists, then

$$\lim_{t\to\infty} e(t) = \lim_{t\to\infty} \lambda(t), \quad \lim_{t\to\infty} n(t) = \Big(\lim_{t\to\infty} \lambda(t)\Big)\mathbb{E}[T].$$

1.4 Approximation to $n(t)$ and $e(t)$

It turns out that if T is small in comparison with the rate of fluctuation of $\lambda(t)$, then $n(t)$ looks like $\lambda(t)$ but with a short time lag, and scaled up by $\mathbb{E}[T]$. Thus it is natural to look for an approximation of the form $n(t) \approx \mathbb{E}[T]\lambda(t-c)$, in order to get a sense of the "time lag" of $n(t)$ from $\lambda(t)$. The best choice of c, in a sense, results in

$$n(t) \approx \mathbb{E}[T]\lambda\left(t - \frac{\mathbb{E}[T^2]}{2\mathbb{E}[T]}\right). \tag{8}$$

Similarly, $e(t)$ looks like a shifted version $\lambda(t)$, and finding a best formula of the form $e(t) \approx \lambda(t-c)$ (this method will be discussed later) results in

$$e(t) \approx \lambda(t - \mathbb{E}[T]). \tag{9}$$

If instead of directly approximating $n(t)$ to get (8), we plug (9) into (7), we get

$$n(t) \approx \Lambda(t - \mathbb{E}[T], t), \tag{10}$$

which seems to have comparable performance as (8) as an approximation for $n(t)$.

1.5 Estimator for $\mathbb{E}[T]$ based on entry times and exit times

Now, given all this theory, you might be wondering, *are there any practical uses of this?* Well, there is one, and it is the only way I have come up with to use data to make inferences using this theory. And here it is. Suppose you have timestamps of all the entrances $t_1^\lambda, \ldots, t_{N_\lambda}^\lambda$ and exits $t_1^e, \ldots, t_{N_e}^e$ in an interval of time $(t_1, t_2]$ (where $t_1 < t_i^\lambda \le t_2$ for all i and $t_1 < t_i^e \le t_2$ for all i). Then a good estimator of $\mathbb{E}[T] = \mu$ is given by

$$\widehat{\mu}_{t_1,t_2,\alpha} = \frac{\int_{t_1}^{t_2} \widetilde{n}(t)\,dt}{\alpha \widetilde{E}(t_1,t_2) + (1-\alpha)\widetilde{\Lambda}(t_1,t_2)}$$

where the numerator is given by

$$\begin{aligned}\int_{t_1}^{t_2} \widetilde{n}(t)\,dt &= \widetilde{n}(t_1)(t_2 - t_1) + \sum_{k=1}^{N_\lambda}(t_2 - t_k^\lambda) - \sum_{k=1}^{N_e}(t_2 - t_k^e) \\ &= \widetilde{n}(t_1)(t_2 - t_1) + \sum_{i=1}^{K}[t_{1i} \in (t_1, t_2]](t_2 - t_{1i}) - \sum_{i=1}^{K}[t_{2i} \in (t_1, t_2]](t_2 - t_{2i}),\end{aligned}$$

$\widetilde{E}(t_1,t_2) = N_e$, $\widetilde{\Lambda}(t_1,t_2) = N_\lambda$, and α is a chosen parameter where you probably want $0 \le \alpha \le 1$, and it is usually best if $\alpha \ge 1/2$ due to some results, and a reasonable choice of α would be around 0.7, although the specific choice of α doesn't affect the accuracy too much. The validity of this estimator and justification for the choice of α will be discussed later in this paper.

$\widehat{\mu}_{t_1,t_2,\alpha}$ converges in probability to μ as $t_2 \to \infty$. $\widehat{\mu}_{t_1,t_2,\alpha}$ is in general a biased estimator of μ. The advantage of this estimator is that you do not need to know which entrances correspond to which exits; if you did, then the task would be trivial. If you do not know this correspondance, then as far as I know, the only thing you can estimate about the probability distribution of T is $\mathbb{E}[T]$, using this method, and you also need to know the value of $\widetilde{n}(t_1)$ (or alternatively, knowing $\widetilde{n}(t)$ for some $t \in [t_1, t_2]$ would technically be sufficient).

1.6 Probability distributions of durations

Two more things that you might want to know about this system are the probability distribution of how long someone currently in the state at time t has been in the state since they entered; call this random variable H_t; and also the probability distribution of the duration of time that someone in the state will have ended up staying in the visit they are in once they leave, call this T'_t. In case you do not understand what I mean by T'_t, this is the probability distribution of, if you pick a random person in the state at time t, their exit time (which has not yet happened) minus the time that they entered. The density functions for H_t and T'_t are given by:

$$f_{H_t}(\tau) = \frac{\lambda(t-\tau)S_T(\tau)}{n(t)}, \quad f_{T'_t}(\tau) = \frac{f_T(\tau)\Lambda(t-\tau,t)}{n(t)}$$

which reduce to the following when in steady state (if in steady state, T'_t is renamed T' and H_t is renamed H):

$$f_H(\tau) = \frac{S_T(\tau)}{\mathbb{E}[T]}, \quad f_{T'}(\tau) = \frac{\tau f_T(\tau)}{\mathbb{E}[T]} = \mathbb{E}[1/T']\tau f_T(\tau)$$

and we can invert the second equation to get

$$f_T(\tau) = \frac{f_{T'}(\tau)}{\mathbb{E}[1/T']\tau}.$$

The moments of T' and T can be defined in terms of each other like so:

$$\mathbb{E}[T'^k] = \frac{\mathbb{E}[T^{k+1}]}{\mathbb{E}[T]}, \quad \mathbb{E}[T^k] = \frac{\mathbb{E}[T'^{k-1}]}{\mathbb{E}[1/T']}.$$

We also have:

$$\mathbb{E}[H^k] = \frac{\mathbb{E}[T^{k+1}]}{(k+1)\mathbb{E}[T]}.$$

In the case where $k=1$, from the equations above we get the expected values of T' and H:

$$\mathbb{E}[T'] = \frac{\mathbb{E}[T^2]}{\mathbb{E}[T]}, \quad \mathbb{E}[H] = \frac{\mathbb{E}[T^2]}{2\mathbb{E}[T]} = \frac{\mathbb{E}[T']}{2}.$$

Notice that the approximation for $n(t)$ in (8) earlier now simplifies to:

$$n(t) \approx \mathbb{E}[T]\lambda(t - \mathbb{E}[H])$$

which says that the approximate "time lag" of the expected number of people function from the entry rate function is the expected value of the steady-state duration that someone in the state has been in the state so far. I don't have an intuitive or logical explanation for this and the discrepancy of the "time lag" of $n(t)$ vs. $e(t)$, but this is what the math says.

2 Proofs

Theorem 1. *If $\widetilde{\Lambda}$ is a Poisson process, then $\widetilde{E}$ is a Poisson process, with intensity function given by*

$$e(t) = \int_0^t \lambda(\tau) f_{T_\tau}(t-\tau)\, d\tau.$$

If time-invariance of T_τ is assumed, i.e., $T_\tau \stackrel{d}{=} T$ for all τ, then this simplifies to

$$e(t) = \int_0^t \lambda(\tau) f(t-\tau)\, d\tau = (f * \lambda)(t).$$

Proof. Follows from the the displacement theorem.

According to Wikipedia, the displacement theorem states the following: Let N be a Poisson point process on $\mathbf{R}^d$ with intensity function $\lambda(x)$. It is then assumed the points of N are randomly displaced somewhere else in $\mathbf{R}^d$ so that each point's displacement is independent and that the displacement of a point formerly at x is a random vector with a probability density $\rho(x, \cdot)$. Then the new point process N_D is also a Poisson point process with intensity function

$$\lambda_D(y) = \int_{\mathbf{R}^d} \lambda(x)\rho(x, y)\, \mathrm{d}x.$$

In our case, we have $d = 1$, $\rho(x, y) = f_{T_x}(y - x)$, and the result easily follows. $\square$

Theorem 2. *Assuming $\widetilde{\Lambda}$ is a Poisson process and the time-invariant property holds, then*

$$n(t) = \mathbb{E}[\widetilde{n}(t)] = \int_0^t \lambda(\tau)S(t-\tau)\, d\tau = (h * \lambda)(t).$$

Proof. Starting with (6), and using Theorem 1, we first rewrite $n'(t)$ using integration by parts; note that $f(x) = -S'(x)$:

$$\begin{aligned} n'(t) = \lambda(t) - e(t) &= \lambda(t) - \int_0^t \lambda(\tau) f_T(t-\tau)\, d\tau \\ &= \lambda(t) + \int_0^t \lambda(\tau)S'(t-\tau)\, d\tau \\ &= \lambda(t) - \lambda(\tau)S(t-\tau)|_{\tau=0}^t + \int_0^t \lambda'(\tau)S(t-\tau)\, d\tau \\ &= \lambda(t) + \lambda(0)S(t) - \lambda(t)\underbrace{S(0)}_{1} + \int_0^t \lambda'(\tau)S(t-\tau)\, d\tau \\ &= \lambda(0)S(t) + \int_0^t \lambda'(\tau)S(t-\tau)\, d\tau. \end{aligned}$$

Next we integrate $n'(t)$ to find $n(t)$:

$$\begin{aligned} n(t) &= \int_0^t n'(u)\, du \\ &= \int_0^t \left(\lambda(0)S(u) + \int_0^u \lambda'(\tau)S(u-\tau)\, d\tau\right) du \\ &= \lambda(0)\int_0^t S(u)\, du + \underbrace{\int_0^t \int_0^u \lambda'(\tau)S(u-\tau)\, d\tau\, du}_{c} \end{aligned}$$

Now apply a u-substitution and a multivariable change of variables:

$$\begin{aligned}
c &= \int_0^t \int_0^u \lambda'(u-v)S(v)\,dv\,du \\
&= \int_0^t \int_v^t \lambda'(u-v)S(v)\,du\,dv \\
&= \int_0^t S(v) \int_v^t \lambda'(u-v)\,du\,dv \\
&= \int_0^t S(v) \int_0^{t-v} \lambda'(w)\,dw\,dv \\
&= \int_0^t S(v)(\lambda(t-v)-\lambda(0))\,dv \\
&= \int_0^t S(v)\lambda(t-v)\,dv - \lambda(0)\int_0^t S(v)\,dv
\end{aligned}$$

And thus we obtain

$$n(t) = \int_0^t S(\tau)\lambda(t-\tau)\,d\tau.$$

□

2.1 Intuitive, looser, and more direct derivation of the convolution formula for $n(t)$

The impulse response $h(\tau)$ is the expected number of people at the reservoir at time τ in a situation where there are no people at the reservoir and then one person comes at time $\tau = 0$, and this person is the only person who ever comes to the reservoir for all of time. The impulse response can also be thought of as the probability that this single person is at the reservoir at time τ, and if the function is shifted we have the following:

$$h(\tau) = \Pr(\text{a person who came at time } t-\tau \text{ is still there})$$

where the current time is t.

The value of the impulse response is the survival function $S(\tau)$ (the probability that a duration is greater than τ) for positive arguments and zero for negative arguments, or alternatively, $h(\tau) = u(\tau)S(\tau)$ where $u(\tau)$ is the unit step function. This is because for $\tau < 0$, the probability that the person is there is zero as they didn't come yet, and for $\tau \geq 0$, the probability the person is there is the probability that they stay longer than a duration of τ, which is the survival function. (The survival function is 1 for negative arguments as the probability for a duration being greater than a negative value is 1, which is why the survival function needs to be multiplied by the unit step function.)

Using this theory from linear time-invariant systems, a simpler and more intuitive, but less rigorous, way of deriving the convolution formula for $n(t)$ is as follows (there may be a way of using this method to prove the formula rigorously, but I have not found a way to do so):

$$\begin{aligned}
n(t) &= \mathbb{E}[\text{number of people here}] \\
&= \int_{\tau=0}^{\infty} \mathbb{E}[\text{number of people who came } \tau \text{ ago and are still here}] \\
&= \int_{\tau=0}^{\infty} \Pr(\text{person who came } \tau \text{ ago here})\mathbb{E}[\text{num entries at time } t-\tau]
\end{aligned}$$

The instantaneous expected entry rate function, λ, is *roughly* defined as follows (not sure if this is correct but this is how I defined it before I knew about Poisson processes):

$$\lambda(t) := \lim_{\Delta t \to 0} \frac{\mathbb{E}[\text{number of people who came during interval } [t, t+\Delta t]]}{\Delta t}$$

which means

$$\mathbb{E}[\text{num people came at time } t-\tau] = \lambda(t-\tau)\, d\tau$$

which makes $n(t)$ the following:

$$n(t) = \int_0^\infty h(\tau)\lambda(t-\tau)\, d\tau = \int_0^t h(\tau)\lambda(t-\tau)\, d\tau = (h * \lambda)(t).$$

2.2 More Proofs

Lemma 1.

$$E(t) = (f * \Lambda)(t).$$

Proof.

$$\begin{aligned}
E(t) &= \int_0^t e(s)\, ds \\
&= \int_0^t \int_0^\infty f(\tau)\lambda(s-\tau)\, d\tau\, ds \\
&= \int_0^\infty \int_0^t f(\tau)\lambda(s-\tau)\, ds\, d\tau \\
&= \int_0^\infty f(\tau) \int_0^t \lambda(s-\tau)\, ds\, d\tau \\
&= \int_0^\infty f(\tau)(\Lambda(t-\tau) - \underbrace{\Lambda(-\tau)}_{0})\, d\tau \\
&= \int_0^\infty f(\tau)\Lambda(t-\tau)\, d\tau = (f * \Lambda)(t).
\end{aligned}$$

□

Lemma 2. *If $a, b \in \mathbb{R}$ and $a \leq b$, then for any $x \in \mathbb{R}$,*

$$[x \in [a, b)] = [a \leq x] - [b \leq x].$$

Proof.

$$\begin{aligned}
[x \in [a, b)] = [a \leq x < b] &= [a \leq x \wedge b > x] \\
&= [a \leq x \wedge \neg(b \leq x)] \\
&= [a \leq x](1 - [b \leq x]) \\
&= [a \leq x] - [a \leq x \wedge b \leq x] \\
&= [a \leq x] - [b \leq x]
\end{aligned}$$

□

Lemma 3.

$$\widetilde{n}(t) = \widetilde{\Lambda}(t) - \widetilde{E}(t).$$

Proof.

$$\begin{aligned}
\widetilde{n}(t) &= \sum_{i=1}^{K} [t \in [t_{1i}, t_{2i})] \\
&= \sum_{i=1}^{K} ([t_{1i} \le t] - [t_{2i} \le t]) && \text{by Lemma 2} \\
&= \sum_{i=1}^{K} [t_{1i} \le t] - \sum_{i=1}^{K} [t_{2i} \le t] = \widetilde{\Lambda}(t) - \widetilde{E}(t)
\end{aligned}$$

□

Lemma 4. *If T is a constant τ, then*

$$\widetilde{n}(t) = \widetilde{\Lambda}(t - \tau, t),$$

which (assuming $\widetilde{\Lambda}$ is a Poisson process) implies

$$\widetilde{n}(t) \sim \mathrm{Pois}(\Lambda(t - \tau, t)).$$

Proof. Recall that the number of people at time t, $\widetilde{n}(t)$, is the number of stays $[t_{1i}, t_{2i})$ such that $t \in [t_{1i}, t_{2i})$, i.e. $t_{1i} \le t < t_{2i}$. Since T is a constant τ, $t_{2i} = t_{1i} + \tau$. Thus we obtain two inequalities: $t_{1i} \le t$, and $t < t_{1i} + \tau$. The second one is equivalent to $t - \tau < t_{1i}$. Now we can combine the two to form $t - \tau < t_{1i} \le t$, i.e. $t_{1i} \in (t - \tau, t]$. Hence, $\widetilde{n}(t)$ equals the number of entrances in the interval $(t - \tau, t]$, which is given by $\widetilde{\Lambda}(t - \tau, t)$. □

Lemma 5. *If $\widetilde{\Lambda}$ is a Poisson process and T is a discrete probability distribution, then for a given value of t,*

$$\widetilde{n}(t) \sim \mathrm{Pois}(\mathbb{E}[\Lambda(t - T, t)]).$$

Proof. For now assume that T can take on a finite number of values $\tau_1, \ldots, \tau_k$ with respective probabilities $p_1, \ldots, p_k$. Split $\widetilde{\Lambda}$ into k Poisson processes $\widetilde{\Lambda}_1, \ldots, \widetilde{\Lambda}_k$, with $\widetilde{\Lambda}_i$ being the entrances t_{1j} in $\widetilde{\Lambda}$ where the person spent a duration of τ_i, i.e. where $t_{2j} - t_{1j} = \tau_i$. Since the duration that someone spends is independent, it follows that by the thinning theorem, all the $\widetilde{\Lambda}_i$ are independent Poisson processes with cumulative entry rate given by $\Lambda_i(t) = p_i \Lambda(t)$. By Lemma 3, it follows that $\widetilde{n}_i(t) \sim \mathrm{Pois}(\Lambda_i(t - \tau_i, t))$. The superposition of the $\widetilde{n}_i$s gives back $\widetilde{n}$:

$$\widetilde{n}(t) = \sum_{i=1}^{k} \widetilde{n}_i(t).$$

Since the sum of independent Poisson random variables is itself Poisson distributed, we have

$$\widetilde{n}(t) \sim \mathrm{Pois}\left(\sum_{i=1}^{k} p_i \Lambda(t - \tau_i, t) \right)$$

i.e.

$$\widetilde{n}(t) \sim \mathrm{Pois}(\mathbb{E}[\Lambda(t - T, t)]).$$

□

Theorem 3. *If $\widetilde{\Lambda}$ is a Poisson process, then for a given value of t,*

$$\widetilde{n}(t) \sim \mathrm{Pois}(n(t)).$$

Proof. Let $T_1, T_2, \ldots$ be a sequence of random variables where

$$\begin{aligned} \mathrm{P}(T_i = 0) &= \mathrm{P}(0 < T \leq 1/i), \\ \mathrm{P}(T_i = 1/i) &= \mathrm{P}(1/i < T \leq 2/i), \\ \mathrm{P}(T_i = 2/i) &= \mathrm{P}(2/i < T \leq 3/i), \\ &\vdots \end{aligned}$$

T_i is a discrete approximation of T. T_i has probability values that sum to 1 as it should:

$$\sum_{j=0}^{\infty} \mathrm{P}(T_i = j/i) = \sum_{j=0}^{\infty} \mathrm{P}(j/i < T \leq (j+1)/i) = \mathrm{P}(T > 0) = 1.$$

This sequence converges in distribution to T. Let $\widetilde{n}_1(t), \widetilde{n}_2(t), \ldots$ be the sequence of random variables of the value of the number of people at time t if T were actually distributed according to T_i. We have, by Lemma 5, that for all i,

$$\widetilde{n}_i(t) \sim \mathrm{Pois}(\mathbb{E}[\Lambda(t - T_i, i)])$$

where

$$\mathbb{E}[\Lambda(t - T_i, t)] = \sum_{j=0}^{\infty} \mathrm{P}\left(\frac{j}{i} < T \leq \frac{j+1}{i}\right) \Lambda\left(t - \frac{j}{i}, t\right).$$

Let $\tau_j = j/i$, $\Delta = 1/i$. So we have

$$\begin{aligned} \mathbb{E}[\Lambda(t - T_i, t)] &= \sum_{j=0}^{\infty} \mathrm{P}(\tau_j < T \leq \tau_j + \Delta) \Lambda(t - \tau_j, t) \\ &= \sum_{j=0}^{\infty} \Delta \frac{\mathrm{P}(\tau_j < T \leq \tau_j + \Delta)}{\Delta} \Lambda(t - \tau_j, t) \end{aligned}$$

Now the limit is

$$\lim_{i \to \infty} \mathbb{E}[\Lambda(t - T_i, i)] = \lim_{\Delta \to 0} \mathbb{E}[\Lambda(t - T_i, i)].$$

Let F denote the CDF of T.

$$\lim_{\Delta \to 0} \frac{\mathrm{P}(\tau_j < T \leq \tau_j + \Delta)}{\Delta} = \lim_{\Delta \to 0} \frac{F(\tau_j + \Delta) - F(\tau_j)}{\Delta} = F'(\tau_j) = f(\tau_j).$$

Thus

$$\begin{aligned} \lim_{\Delta \to 0} \mathbb{E}[\Lambda(t - T_i, t)] &= \lim_{\Delta \to 0} \sum_{j=0}^{\infty} \Delta f(\tau_j) \Lambda(t - \tau_j, t) \\ &= \int_0^{\infty} f(\tau) \Lambda(t - \tau, t)\, d\tau \\ &= \mathbb{E}[\Lambda(t - T, t)] = n(t). \end{aligned}$$

Thus we obtain

$$\lim_{i\to\infty} \mathbb{E}[\Lambda(t-T_i,t)] = n(t).$$

Since the *parameter* of the Poisson random variable $\widetilde{n}_i(t)$ converges to $n(t)$ as $i \to \infty$, it it should be clear that $\widetilde{n}_i(t)$ converges in distribution to $\mathrm{Pois}(n(t))$. Thus we have the following:

$$T_i \rightsquigarrow T, \qquad \widetilde{n}_i(t) \rightsquigarrow \mathrm{Pois}(n(t)).$$

Since $T_i \rightsquigarrow T$, it follows that $\widetilde{n}_i(t) \rightsquigarrow \widetilde{n}(t)$. But we also have $\widetilde{n}_i(t) \rightsquigarrow \mathrm{Pois}(n(t))$. Thus we can conclude that $\widetilde{n}(t) \sim \mathrm{Pois}(n(t))$. □

2.3 Proving consistency of estimator for mean duration

Theorem 4. *Let X_n, Y_n be sequences of random variables, X, Y be random variables, and c be a real number. The following relationships are selected from [5]:*

1. $X_n \xrightarrow{L_1} X \implies X_n \xrightarrow{P} X$.
2. $X_n \xrightarrow{P} X \implies X_n \rightsquigarrow X$.
3. $X_n \rightsquigarrow c \implies X_n \xrightarrow{P} c$.
4. $X_n \rightsquigarrow X,\ \ |X_n - Y_n| \xrightarrow{P} 0 \implies Y_n \rightsquigarrow X$.
5. *If* $X_n \xrightarrow{P} X$ *and* $Y_n \xrightarrow{P} Y$, *then* $X_nY_n \xrightarrow{P} XY$.

Lemma 6. *Let X_n, Y_n be sequences of random variables and c be a real number. If $X_n \xrightarrow{P} c$ and $|X_n - Y_n| \xrightarrow{P} 0$, then $Y_n \xrightarrow{P} c$.*

Proof. Since $X_n \xrightarrow{P} c$, then by part 2 of Theorem 4, $X_n \rightsquigarrow c$. Thus by part 4 of Theorem 4, $Y_n \rightsquigarrow c$. Thus by part 3 of Theorem 4, $Y_n \xrightarrow{P} c$. □

Def. $\lim_{x\to\infty} f(x) = c$ if for all $\varepsilon > 0$, there exits x_0 such that for all $x > x_0$, $|f(x) - c| < \varepsilon$.

Def. Functions f and g are *asymptotically equivalent*, denoted $f(x) \sim g(x)$ as $x \to \infty$ or simply $f(x) \sim g(x)$, if

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1.$$

Lemma 7. *If $\widetilde{n}(t)$ is bounded and $\lim_{t\to\infty} \widetilde{\Lambda}(t) = \infty$, then $\widetilde{E}(t)$ is asymptotically equivalent to $\widetilde{\Lambda}(t)$.*

Proof. Since $\widetilde{n}(t)$ is bounded, there exits $M > 0$ such that for all t, $\widetilde{n}(t) \le M$.

Let $\varepsilon > 0$. Since $\widetilde{\Lambda}(t) \to \infty$ as $t \to \infty$, there exits t_0 such that $\widetilde{\Lambda}(t_0) > M/\varepsilon$. Let t_0 be any such t_0. We have that for all t,

$$\begin{aligned} 0 \le \widetilde{n}(t) &\le M \\ 0 \le \widetilde{\Lambda}(t) - \widetilde{E}(t) &\le M \\ 0 \ge \widetilde{E}(t) - \widetilde{\Lambda}(t) &\ge -M \\ \widetilde{\Lambda}(t) - M \le \widetilde{E}(t) &\le \widetilde{\Lambda}(t) \\ 1 - \frac{M}{\widetilde{\Lambda}(t)} \le \frac{\widetilde{E}(t)}{\widetilde{\Lambda}(t)} &\le 1 \end{aligned}$$

Then for all $t > t_0$,

$$\widetilde{\Lambda}(t) > \frac{M}{\varepsilon}$$

$$\varepsilon > \frac{M}{\widetilde{\Lambda}(t)}$$

$$1 - \varepsilon < 1 - \frac{M}{\widetilde{\Lambda}(t)} \leq \frac{\widetilde{E}(t)}{\widetilde{\Lambda}(t)} \leq 1 < 1 + \varepsilon$$

So

$$1 - \varepsilon < \frac{\widetilde{E}(t)}{\widetilde{\Lambda}(t)} < 1 + \varepsilon$$

i.e.

$$\left| \frac{\widetilde{E}(t)}{\widetilde{\Lambda}(t)} - 1 \right| < \varepsilon.$$

□

Lemma 8. *If $f(x) \sim g(x)$ and $\lim_{x\to\infty} g(x) = \infty$, then for any $\alpha \in \mathbb{R}$, $f(x) + \alpha \sim g(x)$.*

Proof.

$$\lim_{x\to\infty} \frac{f(x)+\alpha}{g(x)} = \lim_{x\to\infty} \frac{f(x)}{g(x)} + \lim_{x\to\infty} \frac{\alpha}{g(x)} = 1 + 0 = 1.$$

□

Lemma 9. *Asymptotic equivalence is commutative, that is, if $f(x) \sim g(x)$ then $g(x) \sim f(x)$.*

Proof. Suppose $f(x) \sim g(x)$. Then

$$\begin{aligned} \lim_{x\to\infty} \frac{g(x)}{f(x)} &= \lim_{x\to\infty} \frac{1}{f(x)/g(x)} \\ &= \frac{\lim_{x\to\infty} 1}{\lim_{x\to\infty} \frac{f(x)}{g(x)}} = \frac{1}{1} = 1. \end{aligned}$$

□

Lemma 10. *If $f(x) \sim g(x)$ and $\lim_{x\to\infty} g(x) = \infty$, then for any $\alpha \in \mathbb{R}$ and $\beta \in \mathbb{R}$, $f(x) + \alpha \sim g(x) + \beta$.*

Proof. Suppose $f(x) \sim g(x)$ and $\lim_{x\to\infty} g(x) = \infty$. By Lemma 8, $f(x) + \alpha \sim g(x)$. By Lemma 9, $g(x) \sim f(x) + \alpha$. Now

$$\lim_{x\to\infty} f(x) = \lim_{x\to\infty} \frac{f(x)}{g(x)} g(x) = \left(\lim_{x\to\infty} \frac{f(x)}{g(x)} \right) \left(\lim_{x\to\infty} g(x) \right) = 1 \cdot \infty = \infty.$$

Thus

$$\lim_{x\to\infty} (f(x) + \alpha) = \alpha + \lim_{x\to\infty} f(x) = \alpha + \infty = \infty.$$

Therefore, by Lemma 8, $g(x) + \beta \sim f(x) + \alpha$. By Lemma 9, $f(x) + \alpha \sim g(x) + \beta$. □

Lemma 11. *If $\widetilde{n}(t)$ is bounded and $\lim_{t\to\infty} \widetilde{\Lambda}(t) = \infty$, then for any t_1, $\widetilde{E}(t_1, t_2) \sim \widetilde{\Lambda}(t_1, t_2)$ as $t_2 \to \infty$.*

Proof. The statement

$$\widetilde{E}(t_1, t_2) \sim \widetilde{\Lambda}(t_1, t_2) \quad \text{as } t_2 \to \infty$$

is equivalent to the statement

$$\widetilde{E}(t_2) - \widetilde{E}(t_1) \sim \widetilde{\Lambda}(t_2) - \widetilde{\Lambda}(t_1) \quad \text{as } t_2 \to \infty.$$

The result then follows from Lemma 7 and Lemma 10. □

Lemma 12. *If $a, b \in \mathbb{R}$ and $a \leq b$, then for any $x \in \mathbb{R}$,*

$$[x \in (a, b]] = [x \leq b] - [x \leq a].$$

Proof.

$$\begin{aligned}
[x \in (a, b]] = [a < x \leq b] &= [a < x \wedge x \leq b] \\
&= [x > a \wedge x \leq b] \\
&= [\neg(x \leq a) \wedge x \leq b] \\
&= (1 - [x \leq a])[x \leq b] \\
&= [x \leq b] - [x \leq a][x \leq b] \\
&= [x \leq b] - [x \leq a \wedge x \leq b] \\
&= [x \leq b] - [x \leq a]
\end{aligned}$$

□

Notation The following notations for proving the convergence will be used.

$$T_i := t_{2i} - t_{1i},$$

$$J(t_1, t_2) := \sum_{i=1}^{K} [t_{1i} \in (t_1, t_2]] T_i,$$

$$\widehat{\mu}^{\lambda}_{t_1,t_2} := \frac{J(t_1, t_2)}{\widetilde{\Lambda}(t_1, t_2)},$$

$$\begin{aligned}
I(t_1, t_2) &:= \int_{t_1}^{t_2} \widetilde{n}(t)\, dt \\
&= \widetilde{n}(t_1)(t_2 - t_1) + \sum_{i=1}^{K} [t_{1i} \in (t_1, t_2]](t_2 - t_{1i}) - \sum_{i=1}^{K} [t_{2i} \in (t_1, t_2]](t_2 - t_{2i}),
\end{aligned}$$

$$\widehat{\mu}_{t_1,t_2,\alpha} := \frac{I(t_1, t_2)}{\alpha \widetilde{E}(t_1, t_2) + (1 - \alpha)\widetilde{\Lambda}(t_1, t_2)}, \quad \widehat{\mu}_{t_1,t_2,0} = \frac{I(t_1, t_2)}{\widetilde{\Lambda}(t_1, t_2)},$$

$$\mu := \mathbb{E}[T].$$

Lemma 13. *If $\widetilde{\Lambda}(t) \to \infty$ as $t \to \infty$ then*

$$\widehat{\mu}^{\lambda}_{t_1,t_2} \xrightarrow{P} \mu \quad \text{as } t_2 \to \infty.$$

Proof. Looking at the formula for $\widehat{\mu}^{\lambda}_{t_1,t_2}$, this estimator is taking the average of the durations $T_i = t_{2i} - t_{1i}$ for which $t_{1i} \in (t_1, t_2]$. Since $\widetilde{\Lambda}(t) \to \infty$ as $t \to \infty$, the number of entrances t_{1i} in the interval $(t_1, t_2]$, i.e. $\widetilde{\Lambda}(t_1, t_2)$, goes to ∞ as $t_2 \to \infty$. Therefore, the result follows from the weak law of large numbers. $\square$

Lemma 14. *Consider $\widetilde{\Lambda}$ fixed and $\widetilde{E}$ randomly generated from $\widetilde{\Lambda}$ by IID $T_1, \dots, T_K \sim T$. If $\lim_{t\to\infty} \widetilde{\Lambda}(t) = \infty$ and $\widetilde{n}(t)$ is bounded, then*

$$|\widehat{\mu}_{t_1,t_2,0} - \widehat{\mu}^{\lambda}_{t_1,t_2}| \xrightarrow{P} 0 \quad \textit{as } t_2 \to \infty.$$

Proof. First we rewrite $I(t_1, t_2)$:

$$I(t_1,t_2) = \underbrace{\widetilde{n}(t_1)(t_2 - t_1)}_{A} + \underbrace{\sum_{i=1}^{K} [t_{1i} \in (t_1, t_2]](t_2 - t_{1i})}_{B} - \underbrace{\sum_{i=1}^{K} [t_{2i} \in (t_1, t_2]](t_2 - t_{2i})}_{C}$$

$$\begin{aligned}
A &= (\widetilde{\Lambda}(t_1) - \widetilde{E}(t_1))(t_2 - t_1) \\
&= \left(\sum_{i=1}^{K} [t_{1i} \le t_1] - \sum_{i=1}^{K} [t_{2i} \le t_1]\right)(t_2 - t_1) \\
&= \sum_{i=1}^{K} ([t_{1i} \le t_1] - [t_{2i} \le t_1])(t_2 - t_1) \\
&= \sum_{i=1}^{K} ([t_{1i} \le t_1]t_2 - [t_{1i} \le t_1]t_1 - [t_{2i} \le t_1]t_2 + [t_{2i} \le t_1]t_1) \\
B &= \sum_{i=1}^{K} [t_{1i} \in (t_1, t_2]](t_2 - t_{1i}) \\
&= \sum_{i=1}^{K} ([t_{1i} \le t_2] - [t_{1i} \le t_1])(t_2 - t_{1i}) \\
&= \sum_{i=1}^{K} ([t_{1i} \le t_2]t_2 - [t_{1i} \le t_2]t_{1i} - [t_{1i} \le t_1]t_2 + [t_{1i} \le t_1]t_{1i}) \\
C &= \sum_{i=1}^{K} [t_{2i} \in (t_1, t_2]](t_2 - t_{2i}) \\
&= \sum_{i=1}^{K} ([t_{2i} \le t_2] - [t_{2i} \le t_1])(t_2 - t_{2i}) \\
&= \sum_{i=1}^{K} ([t_{2i} \le t_2]t_2 - [t_{2i} \le t_2]t_{2i} - [t_{2i} \le t_1]t_2 + [t_{2i} \le t_1]t_{2i})
\end{aligned}$$

$$
\begin{aligned}
I(t_1,t_2) = A + B - C = \sum_{i=1}^{K} \Big(& [t_{1i} \le t_1]t_2 - [t_{1i} \le t_1]t_1 - [t_{2i} \le t_1]t_2 + [t_{2i} \le t_1]t_1 \\
& + [t_{1i} \le t_2]t_2 - [t_{1i} \le t_2]t_{1i} - [t_{1i} \le t_1]t_2 + [t_{1i} \le t_1]t_{1i} \\
& - [t_{2i} \le t_2]t_2 + [t_{2i} \le t_2]t_{2i} + [t_{2i} \le t_1]t_2 - [t_{2i} \le t_1]t_{2i} \Big) \\
= \sum_{i=1}^{K} \Big(& - [t_{1i} \le t_1]t_1 + [t_{2i} \le t_1]t_1 \\
& + [t_{1i} \le t_2]t_2 - [t_{1i} \le t_2]t_{1i} + [t_{1i} \le t_1]t_{1i} \\
& - [t_{2i} \le t_2]t_2 + [t_{2i} \le t_2]t_{2i} - [t_{2i} \le t_1]t_{2i} \Big)
\end{aligned}
$$

Next we rewrite $J(t_1, t_2)$:

$$
\begin{aligned}
J(t_1,t_2) &= \sum_{i=1}^{K} [t_{1i} \in (t_1, t_2]](t_{2i} - t_{1i}) \\
&= \sum_{i=1}^{K} ([t_{1i} \le t_2] - [t_{1i} \le t_1])(t_{2i} - t_{1i}) \\
&= \sum_{i=1}^{K} ([t_{1i} \le t_2]t_{2i} - [t_{1i} \le t_2]t_{1i} - [t_{1i} \le t_1]t_{2i} + [t_{1i} \le t_1]t_{1i})
\end{aligned}
$$

Next we write $I(t_1, t_2) - J(t_1, t_2)$:

$$
\begin{aligned}
I(t_1,t_2) - J(t_1,t_2) = \sum_{i=1}^{K} \Big(& -[t_{1i} \le t_1]t_1 + [t_{2i} \le t_1]t_1 \\
& + [t_{1i} \le t_2]t_2 - [t_{1i} \le t_2]t_{1i} + [t_{1i} \le t_1]t_{1i} \\
& - [t_{2i} \le t_2]t_2 + [t_{2i} \le t_2]t_{2i} - [t_{2i} \le t_1]t_{2i} \\
& - [t_{1i} \le t_2]t_{2i} + [t_{1i} \le t_2]t_{1i} + [t_{1i} \le t_1]t_{2i} - [t_{1i} \le t_1]t_{1i} \Big) \\
= \sum_{i=1}^{K} \Big(& -[t_{1i} \le t_1]t_1 + [t_{2i} \le t_1]t_1 + [t_{1i} \le t_2]t_2 \\
& - [t_{2i} \le t_2]t_2 + [t_{2i} \le t_2]t_{2i} - [t_{2i} \le t_1]t_{2i} - [t_{1i} \le t_2]t_{2i} + [t_{1i} \le t_1]t_{2i} \Big) \\
= \sum_{i=1}^{K} & \Big([t_{1i} \le t_1](t_{2i} - t_1) + [t_{2i} \le t_1](t_1 - t_{2i}) + [t_{1i} \le t_2](t_2 - t_{2i}) + [t_{2i} \le t_2](t_{2i} - t_2) \Big) \\
= \sum_{i=1}^{K} & \Big(([t_{1i} \le t_1] - [t_{2i} \le t_1])(t_{2i} - t_1) + ([t_{2i} \le t_2] - [t_{1i} \le t_2])(t_{2i} - t_2) \Big) \\
= \sum_{i=1}^{K} & \Big(([t_{1i} \le t_1] - [t_{2i} \le t_1])(t_{2i} - t_1) - ([t_{1i} \le t_2] - [t_{2i} \le t_2])(t_{2i} - t_2) \Big) \\
= \sum_{i=1}^{K} & \Big([t_1 \in [t_{1i}, t_{2i})](t_{2i} - t_1) - [t_2 \in [t_{1i}, t_{2i})](t_{2i} - t_2) \Big) \\
= \sum_{i=1}^{K} & [t_1 \in [t_{1i}, t_{2i})](t_{2i} - t_1) - \sum_{i=1}^{K} [t_2 \in [t_{1i}, t_{2i})](t_{2i} - t_2).
\end{aligned}
$$

Next we upper-bound the absolute value of this quantity, first bounding it with the triangle inequality, noting that both terms are non-negative:

$$
\begin{aligned}
|I(t_1,t_2) - J(t_1,t_2) &= \left| \sum_{i=1}^{K} [t_1 \in [t_{1i}, t_{2i})](t_{2i} - t_1) - \sum_{i=1}^{K} [t_2 \in [t_{1i}, t_{2i})](t_{2i} - t_2) \right| \\
&\le \sum_{i=1}^{K} [t_1 \in [t_{1i}, t_{2i})](t_{2i} - t_1) + \sum_{i=1}^{K} [t_2 \in [t_{1i}, t_{2i})](t_{2i} - t_2) \qquad \text{(triangle inequality)}
\end{aligned}
$$

$$
\sum_{i=1}^{K} [t_2 \in [t_{1i}, t_{2i})](t_{2i} - t_2) = \sum_{t_2 \in [t_{1i}, t_{2i})} (t_{2i} - t_2) \le \sum_{t_2 \in [t_{1i}, t_{2i})} (\underbrace{t_{2i} - t_{1i}}_{T_i})
$$

Thus we obtain

$$
\sum_{i=1}^{K} [t_2 \in [t_{1i}, t_{2i})](t_{2i} - t_2) \le \sum_{t_2 \in [t_{1i}, t_{2i})} T_i.
$$

And similarly,

$$
\sum_{i=1}^{K} [t_1 \in [t_{1i}, t_{2i})](t_{2i} - t_1) \le \sum_{t_1 \in [t_{1i}, t_{2i})} T_i.
$$

Therefore, we get

$$|I(t_1,t_2) - J(t_1,t_2)| \le \sum_{t_2 \in [t_{1i}, t_{2i})} T_i + \sum_{t_1 \in [t_{1i}, t_{2i})} T_i.$$

Since $\widetilde{n}(t)$ is bounded, there exits $M > 0$ such that for all t, $\widetilde{n}(t) \le M$. Let M be such constant. Taking expectations of both sides (something you can do with expectations), we get

$$\begin{aligned}\mathbb{E}[|I(t_1,t_2) - J(t_1,t_2)|] &\le \widetilde{n}(t_1)\mathbb{E}[T] + \widetilde{n}(t_2)\mathbb{E}[T] \\ &\le M\mathbb{E}[T] + M\mathbb{E}[T] = 2M\mathbb{E}[T].\end{aligned}$$

Now finally

$$\begin{aligned}\mathbb{E}[|\widehat{\mu}_{t_1,t_2,0} - \widehat{\mu}^{\lambda}_{t_1,t_2}|] &= \mathbb{E}\left[\left|\frac{I(t_1,t_2)}{\widetilde{\Lambda}(t_1,t_2)} - \frac{J(t_1,t_2)}{\widetilde{\Lambda}(t_1,t_2)}\right|\right] \\ &= \frac{\mathbb{E}[|I(t_1,t_2) - J(t_1,t_2)|]}{\widetilde{\Lambda}(t_1,t_2)} \\ &\le \frac{2M\mathbb{E}[T]}{\widetilde{\Lambda}(t_1,t_2)} \to 0 \quad \text{as } t_2 \to \infty.\end{aligned}$$

Thus,

$$\mathbb{E}[||\widehat{\mu}_{t_1,t_2,0} - \widehat{\mu}^{\lambda}_{t_1,t_2}| - 0|] \to 0 \quad \text{as } t_2 \to \infty,$$

or in other words

$$|\widehat{\mu}_{t_1,t_2,0} - \widehat{\mu}^{\lambda}_{t_1,t_2}| \xrightarrow{L_1} 0 \quad \text{as } t_2 \to \infty.$$

Therefore by part 1 of Theorem 4,

$$|\widehat{\mu}_{t_1,t_2,0} - \widehat{\mu}^{\lambda}_{t_1,t_2}| \xrightarrow{P} 0 \quad \text{as } t_2 \to \infty.$$

$\square$

Lemma 15. *Consider $\widetilde{\Lambda}$ fixed and $\widetilde{E}$ randomly generated from $\widetilde{\Lambda}$ by IID $T_1, \ldots, T_K \sim T$. If $\lim_{t\to\infty} \widetilde{\Lambda}(t) = \infty$ and $\widetilde{n}(t)$ is bounded, then*

$$\widehat{\mu}_{t_1,t_2,0} \xrightarrow{P} \mu \quad \text{as } t_2 \to \infty.$$

Proof. By Lemma 13,

$$\widehat{\mu}^{\lambda}_{t_1,t_2} \xrightarrow{P} \mu \quad \text{as } t_2 \to \infty.$$

By Lemma 14,

$$|\widehat{\mu}_{t_1,t_2,0} - \widehat{\mu}^{\lambda}_{t_1,t_2}| \xrightarrow{P} 0 \quad \text{as } t_2 \to \infty.$$

Thus by Lemma 6,

$$\widehat{\mu}_{t_1,t_2,0} \xrightarrow{P} \mu \quad \text{as } t_2 \to \infty.$$

$\square$

Lemma 16. *Let x_n be a (non-random) sequence and c be a real number. If $\lim_{n\to\infty} x_n = c$ then $x_n \xrightarrow{P} c$.*

Proof.

$$
\begin{aligned}
\mathrm{P}(|x_n - c| > \varepsilon) &\le \frac{\mathbb{E}(|x_n - c|)}{\varepsilon} && \text{(Markov's inequality)} \\
&= \frac{|x_n - c|}{\varepsilon} \to 0 \quad \text{as } n \to \infty.
\end{aligned}
$$

□

Theorem 5. *Consider $\widetilde{\Lambda}$ fixed and $\widetilde{E}$ randomly generated from $\widetilde{\Lambda}$ by IID $T_1, \ldots, T_K \sim T$. If $\lim_{t\to\infty} \widetilde{\Lambda}(t) = \infty$ and $\widetilde{n}(t)$ is bounded, then for any $\alpha \in \mathbb{R}$,*

$$\widehat{\mu}_{t_1,t_2,\alpha} \xrightarrow{P} \mu \quad \text{as } t_2 \to \infty.$$

Proof.

$$
\begin{aligned}
\widehat{\mu}_{t_1,t_2,\alpha} - \widehat{\mu}_{t_1,t_2,0} &= \frac{I(t_1,t_2)}{\alpha \widetilde{E}(t_1,t_2) + (1-\alpha)\widetilde{\Lambda}(t_1,t_2)} - \frac{I(t_1,t_2)}{\widetilde{\Lambda}(t_1,t_2)} \\
&= \frac{I(t_1,t_2)/\widetilde{\Lambda}(t_1,t_2)}{\alpha \frac{\widetilde{E}(t_1,t_2)}{\widetilde{\Lambda}(t_1,t_2)} + (1-\alpha)} - \frac{I(t_1,t_2)}{\widetilde{\Lambda}(t_1,t_2)} \\
&= \frac{I(t_1,t_2)}{\widetilde{\Lambda}(t_1,t_2)} \left(\frac{1}{\alpha \frac{\widetilde{\Lambda}(t_1,t_2)}{\widetilde{E}(t_1,t_2)} + 1 - \alpha} - 1 \right) \\
&= \widehat{\mu}_{t_1,t_2,0} \left(\frac{1 - \alpha \frac{\widetilde{\Lambda}(t_1,t_2)}{\widetilde{E}(t_1,t_2)} - 1 + \alpha}{\alpha \frac{\widetilde{\Lambda}(t_1,t_2)}{\widetilde{E}(t_1,t_2)} + 1 - \alpha} \right) \\
&= \widehat{\mu}_{t_1,t_2,0} \left(\frac{\alpha \left(1 - \frac{\widetilde{\Lambda}(t_1,t_2)}{\widetilde{E}(t_1,t_2)}\right)}{\alpha \left(\frac{\widetilde{\Lambda}(t_1,t_2)}{\widetilde{E}(t_1,t_2)} - 1 \right) + 1} \right).
\end{aligned}
$$

We have that, since $\lim_{t_2\to\infty} \frac{\widetilde{\Lambda}(t_1,t_2)}{\widetilde{E}(t_1,t_2)} = 1$ by Lemma 11,

$$\lim_{t_2\to\infty} \frac{\alpha \left(1 - \frac{\widetilde{\Lambda}(t_1,t_2)}{\widetilde{E}(t_1,t_2)}\right)}{\alpha \left(\frac{\widetilde{\Lambda}(t_1,t_2)}{\widetilde{E}(t_1,t_2)} - 1 \right) + 1} = 0.$$

Thus by Lemma 16,

$$\frac{\alpha \left(1 - \frac{\widetilde{\Lambda}(t_1,t_2)}{\widetilde{E}(t_1,t_2)}\right)}{\alpha \left(\frac{\widetilde{\Lambda}(t_1,t_2)}{\widetilde{E}(t_1,t_2)} - 1 \right) + 1} \xrightarrow{P} 0 \quad \text{as } t_2 \to \infty.$$

By the continuous mapping theorem [4],

$$|\widehat{\mu}_{t_1,t_2,0}| \xrightarrow{P} \mu, \quad \left| \frac{\alpha \left(1 - \frac{\widetilde{\Lambda}(t_1,t_2)}{\widetilde{E}(t_1,t_2)}\right)}{\alpha \left(\frac{\widetilde{\Lambda}(t_1,t_2)}{\widetilde{E}(t_1,t_2)} - 1 \right) + 1} \right| \xrightarrow{P} 0 \qquad \text{as } t_2 \to \infty.$$

Therefore, by part 5 of Theorem 4,

$$|\widehat{\mu}_{t_1,t_2,\alpha} - \widehat{\mu}_{t_1,t_2,0}| = |\widehat{\mu}_{t_1,t_2,0}| \left| \frac{\alpha\Big(1 - \frac{\widetilde{\Lambda}(t_1,t_2)}{\widetilde{E}(t_1,t_2)}\Big)}{\alpha\Big(\frac{\widetilde{\Lambda}(t_1,t_2)}{\widetilde{E}(t_1,t_2)} - 1\Big) + 1} \right| \xrightarrow{P} \mu \cdot 0 = 0 \quad \text{as } t_2 \to \infty.$$

And since by Lemma 15,

$$\widehat{\mu}_{t_1,t_2,0} \xrightarrow{P} \mu \quad \text{as } t_2 \to \infty,$$

by Lemma 6,

$$\widehat{\mu}_{t_1,t_2,\alpha} \xrightarrow{P} \mu \quad \text{as } t_2 \to \infty.$$

□

References

[1] Stephen G Eick, William A Massey, and Ward Whitt. The physics of the $M_t/G/\infty$ queue. *Operations Research*, 41(4):731–742, 1993.

[2] Song-Hee Kim and Ward Whitt. Estimating waiting times with the time-varying little's law. *Probability in the Engineering and Informational Sciences*, 27(4):471–506, 2013.

[3] Song-Hee Kim and Ward Whitt. Statistical analysis with little's law. *Operations Research*, 61(4):1030–1045, 2013.

[4] Wikipedia contributors. Continuous mapping theorem — Wikipedia, the free encyclopedia, 2022. [Online; accessed 26-November-2022].

[5] Wikipedia contributors. Convergence of random variables — Wikipedia, the free encyclopedia, 2022. [Online; accessed 26-November-2022].