Operations Research Letters 40 (2012) 230–234

Contents lists available at SciVerse ScienceDirect

Operations Research Letters

journal homepage: www.elsevier.com/locate/orl

Extending the FCLT version of $L = \lambda W$

Ward Whitt *

Department of Industrial Engineering and Operations Research, Columbia University, United States

A R T I C L E   I N F O

Article history:
Received 29 December 2011
Accepted 15 March 2012
Available online 23 March 2012

Keywords:
Little's law
$L = \lambda W$
Functional central limit theorem
Confidence intervals
Continuous mapping theorem
Composition with centering

A B S T R A C T

The functional central limit theorem (FCLT) version of Little's law ($L = \lambda W$) established by Glynn and Whitt is extended to show that a bivariate FCLT for the number in the system and the waiting times implies the joint FCLT for all processes. It is based on a converse to the preservation of convergence by the composition map with centering on the function space containing the sample paths, exploiting monotonicity.

1. Introduction

The relation $L = \lambda W$ (Little's law [12]) states that the average number of customers (items) waiting in line (in a system), L, is equal to the arrival rate (throughput) λ multiplied by the average waiting time (time spent in the system) per customer, W. Under very general conditions, the relation is valid for both long-run averages of individual sample paths and expected values of stationary random variables in stochastic models; see [1,5,15–18,21].

1.1. The statistical approach: viewing finite averages as estimates

In applications, $L = \lambda W$ is often applied with measurements over a finite time interval, as emphasized by Buzen and Denning [3,4] and Little [13]. Given a time interval that is judged to be suitably stationary, we may exploit $L = \lambda W$ to make inferences, e.g., predictions at other times. To do so, we can assume that the system satisfies the (weak) conditions required for the relation $L = \lambda W$ to be valid, both for the limits of sample averages and for the corresponding steady-state quantities associated with stationary processes. Then we regard the sample averages based on measurements as *estimates* of the unknown parameters L, λ and W.

We consider all customers that are in the system at some time during a designated interval $[0, t]$. For customer k, let T_k be the arrival time, D_k the departure time and $W_k \equiv D_k - T_k$ the waiting time, where $-\infty < T_k < D_k < \infty$, $[0, t] \cap [T_k, D_k] \neq \emptyset$ and $\equiv$ denotes "equality by definition". Let $A(t)$ count the total number of new arrivals in the interval $[0, t]$, assuming for simplicity that $A(0) = 0$, and let $L(t)$ be the number of customers in the system at time t. Hence, $L(0)$ is the number of customers remaining in the system among those that arrived before time 0. The natural estimators of the parameters L, λ and W are the respective averages over the time interval $[0, t]$, i.e.,

$$\bar{\lambda}(t) \equiv t^{-1}A(t), \qquad \bar{L}(t) \equiv t^{-1}\int_0^t L(s)\,ds,$$
$$\bar{W}(t) \equiv \frac{\sum_{k=1}^{A(t)} W_k}{A(t)}. \tag{1}$$

If $L(0) = L(t) = 0$, then $\bar{L}(t) = \bar{\lambda}(t)\bar{W}(t)$, but more generally this finite-sample relation only holds approximately, unless the definitions are altered, in which case the relation is difficult to interpret; e.g., see Section 4.6 of [5] and [10].

Following standard statistical practice, it is appropriate to evaluate the effectiveness of these estimators by also estimating confidence intervals. As in simulation output analysis in discrete-event stochastic simulation, it is natural to use the method of batch means, e.g., see Section 3.3.1 of [2]. This approach is reviewed and illustrated with call center data in [10].

Theoretical support for estimating confidence intervals with ample data, either by independent samples or by batch means, is provided by an associated central limit theorem (CLT). Under regularity conditions, there is a joint CLT

$$(\hat{L}(t), \hat{\lambda}(t), \hat{W}(t)) \Rightarrow N(0, \Sigma) \quad \text{in } \mathbb{R}^3 \quad \text{as } t \to \infty, \tag{2}$$

* Correspondence to: Mail Code 4704, S. W. Mudd Building, 500 West 120th Street, New York, NY 10027-6699, United States.
E-mail address: ww2040@columbia.edu.

0167-6377/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.orl.2012.03.010

where $\hat{L}(t) \equiv \sqrt{t}(\bar{L}(t) - L), \hat{\lambda}(t) \equiv \sqrt{t}(\bar{\lambda}(t) - \lambda), \hat{W}(t) \equiv \sqrt{t}(\bar{W}(t) - W), N(m, \Sigma)$ denotes a trivariate normal random vector with mean vector $m \equiv (m_1, m_2, m_3)$ and 3×3 covariance matrix Σ having variances on the diagonal and covariances off the diagonal [6–8]. The associated functional CLT (FCLT; see Section 2 and [20]) establishes a limit for the entire stochastic process, which typically (with Brownian motion limits) implies that the m batches are asymptotically independent as $t \to \infty$. Because of the fundamental relation between cumulative processes underlying $L = \lambda W$, the joint CLT in (2) is essentially two-dimensional and takes a special form; see Section 2.

1.2. Exploiting $L = \lambda W$ to create alternative estimators

Just as we can use the relation $L = \lambda W$ and knowledge of any two of the three quantities L, λ and W to compute the remaining one, so can we use any two of the three estimators in (1) to create a new alternative estimator for the remaining one, exploiting $L = \lambda W$:

$$\bar{L}_{\lambda,W}(t) \equiv \bar{\lambda}(t)\bar{W}(t),$$
$$\bar{\lambda}_{L,W}(t) \equiv \frac{\bar{L}(t)}{\bar{W}(t)} \quad \text{and} \quad \bar{W}_{\lambda,L}(t) \equiv \frac{\bar{L}(t)}{\bar{\lambda}(t)}. \tag{3}$$

To see that this might be useful, note that we might well have available the sample path segment $\{L(s) : 0 \le s \le t\}$, but not have access to the individual waiting times W_k. From that sample path segment, we can directly observe the arrivals (jumps up) and departures (jumps down), but in many applications we cannot determine the time each item spends in the system, because the items need not depart in the same order that they arrived. Thus, we may want to use the alternative estimator $\bar{W}_{\lambda,L}(t)$ in (3).

From the theory about the limits of sample-path averages [5], we know that all these estimators in (1) and (3) are consistent; i.e., they converge to the desired value as the sample size grows. The CLT and FCLT versions of $L = \lambda W$ provide strong support for assessing the asymptotic efficiency, e.g., estimating associated confidence intervals. Under the specified conditions, they justify asymptotic normality as in (2). Moreover, (i) there is a joint CLT for all these estimators generalizing (2), where the limit is multivariate normal, and (ii) the alternative estimators in (3) are asymptotically equivalent to the natural estimators, i.e.,

$$|(\hat{L}(t), \hat{\lambda}(t), \hat{W}(t)) - (\hat{L}_{\lambda,W}(t), \hat{\lambda}_{L,W}(t), \hat{W}_{\lambda,W}(t))| \Rightarrow 0 \quad \text{as } t \to \infty, \tag{4}$$

where $|\cdot|$ denotes the usual norm on $\mathbb{R}^3$ and the random variables are defined as in (2) in terms of the sample averages in (1) and (3). The limit (4) implies that the random variables $\hat{W}_{\lambda,W}(t)$ and $\hat{W}(t)$ not only have the same asymptotic normal distribution as t grows (i.e., the same variance constant in the CLT, so that the estimators have the same asymptotic efficiency), but that the random variables $\hat{W}_{\lambda,W}(t)$ and $\hat{W}$ take the same value, asymptotically, as well. (This is much stronger than the obvious conclusion that the estimators $\bar{W}(t)$ and $\bar{W}_{\lambda,L}(t)$ take the same value W, asymptotically.)

An *advantage* in asymptotic efficiency (lower variance) can be gained when one of the parameters L, λ or W is known in advance, rather than estimated, as often occurs in simulation [9,11]. Since the simulator directly constructs the simulation model, the arrival rate is typically known in advance. The FCLT version of $L = \lambda W$ also plays a role in determining the more efficient estimator and in quantifying the advantage.

1.3. The rest of this paper

A theoretical basis for the strong conclusions above is provided by the FCLT in [6], which we review in Section 2. It shows that a FCLT for two of the processes implies a joint FCLT for all the processes, and shows how the limit processes are related. We extend the FCLT in [6] by adding a new sufficient condition. We now show that it suffices to start with the bivariate FCLT for the number in system and waiting time processes. To achieve that new result, we establish Theorem 2 here, a new converse to the preservation of convergence for the composition map with centering, as stated in Corollary 13.3.3 of [20], which exploits the extra condition of monotonicity. That supporting preservation of convergence result should be useful for establishing new FCLT's in other contexts.

Here is how the rest of this paper is organized. We review and extend the FCLT from [6] in Section 2. We establish implications for the natural estimator $\bar{W}(t)$ in Section 3. We draw corresponding new implications about the alternative estimators exploiting $L = \lambda W$ in Section 4. We establish new converses to the preservation of convergence under the composition map in Section 5. Finally, we prove the new part of the FCLT version of $L = \lambda W$ in Section 6.

2. The joint functional central limit theorem

Consider the usual framework for stochastic process limits, as in [20], with $\mathcal{D}$ denoting the space of all right-continuous real-valued functions on the nonnegative real line, $\mathcal{D}^k$ the k-dimensional product space, and $\mathcal{C}$ the subspace of continuous functions. Given the arrival times T_n, the arrival counting process $A(t)$, the waiting times W_n and the number in system $L(t)$, construct the following four FCLT-scaled random elements of $\mathcal{D}$:

$$\hat{T}_n(t) \equiv n^{-1/2}\left(T_{\lfloor nt \rfloor} - \lambda^{-1}nt\right),$$
$$\hat{A}_n(t) \equiv n^{-1/2}\left(A(nt) - \lambda nt\right), \tag{5}$$
$$\hat{W}_n(t) \equiv n^{-1/2}\left(\sum_{k=1}^{\lfloor nt \rfloor} W_k - Wnt\right),$$
$$\hat{L}_n(t) \equiv n^{-1/2}\left(\int_0^{nt} L(s)\,ds - Lnt\right),$$

for $t \ge 0$. To make a fair comparison, we should use the same data. If we fix t, then we should look at the sum of the first $A(t)$ waiting times; if we fix k, then we should look at the time interval $[0, T_k]$. That leads to considering two additional random elements of $\mathcal{D}$:

$$\hat{W}_n^A(t) \equiv n^{-1/2}\left(\sum_{k=1}^{A(nt)} W_k - Lnt\right),$$
$$\hat{L}_n^T(t) \equiv n^{-1/2}\left(\int_0^{T_{\lfloor nt \rfloor}} L(s)\,ds - Wnt\right). \tag{6}$$

Both $\hat{L}_n(t)$ and $\hat{W}_n^A(t)$ consider the data over the interval $[0, nt]$, whereas both $\hat{W}_n(t)$ and $\hat{L}_n^T(t)$ consider the data over the interval $[0, T_{\lfloor nt \rfloor}]$, so each pair is directly comparable. The two elements in each pair turn out to be asymptotically equivalent.

As a regularity condition, we need to control the asymptotic behavior of the scaled end effects. The end effects at the fixed time t or the random time T_n are harder to treat than the end effect at time 0, because they change as t or n increases, whereas the end effect at time 0 does not change. Any finite initial condition at time 0 necessarily is asymptotically negligible after scaling.

To treat the end effects at the right end of the interval, let $R(t)$ be the total work in service time in the system after time t among the first $A(t)$ arrivals up to time t and let S_k be the total work in service time in the system after time T_k among the first k arrivals. To give formulas, let $I_k(t)$ equal 1 if customer k is in the system at time t and 0 if it is not. Then

$$R(t) \equiv \sum_{j=1}^{A(t)} \int_t^{\infty} I_j(s)\,ds \quad \text{and} \quad S_k \equiv \sum_{j=1}^{k} \int_{T_k}^{\infty} I_j(s)\,ds.$$

Let the two associated scaled random elements of $\mathcal{D}$ be $\hat{R}_n(t) \equiv n^{-1/2}R(nt)$ and $\hat{S}_n(t) \equiv n^{-1/2}S_{\lfloor nt \rfloor}, t \geq 0$. Let e be the identity function in $\mathcal{D}$, i.e., $e(t) \equiv t, t \geq 0$, so that $0e$ is the zero function. A principal technical issue is developing conditions for the two remainder terms to be asymptotically negligible. That issue is addressed in [6]; we will simply assume it here.

The following theorem extends [6] by including the new sufficient condition, the convergence of the final pair $(\hat{L}_n, \hat{W}_n)$. We prove that result in Section 6, drawing on the new converse to the preservation of convergence by composition with centering in Corollary 13.3.1 in [20], stated and proved in Section 5.

Theorem 1 (*FCLT Version of $L = \lambda W$ from [6]*). *Suppose that the relation $L = \lambda W$ is valid as the limit of sample averages, where all three limits are positive and finite, and either $\hat{R}_n \Rightarrow 0e$ or $\hat{S}_n \Rightarrow 0e$ in $\mathcal{D}$. If any one of the following five limits holds in $\mathcal{D}^2$:*

$$\begin{aligned} &(\hat{W}_n, \hat{T}_n) \Rightarrow (\hat{W}, \hat{T}) \quad or \quad (\hat{L}_n, \hat{T}_n) \Rightarrow (\hat{L}, \hat{T}) \quad where \\ &\quad P(\hat{T} \in \mathcal{C}) = 1, \\ &(\hat{W}_n, \hat{A}_n) \Rightarrow (\hat{W}, \hat{A}) \quad or \quad (\hat{L}_n, \hat{A}_n) \Rightarrow (\hat{L}, \hat{A}), \quad where \\ &\quad P(\hat{A} \in \mathcal{C}) = 1, \\ &(\hat{L}_n, \hat{W}_n) \Rightarrow (\hat{L}, \hat{W}) \quad where \\ &\quad P((\hat{L}, \hat{W}) \in \mathcal{C}^2) = 1 \quad and \quad \hat{L}(0) = \hat{W}(0) = 0, \end{aligned} \tag{7}$$

then there is the joint convergence in $\mathcal{D}^8$

$$(\hat{W}_n, \hat{L}_n, \hat{T}_n, \hat{A}_n, \hat{W}_n^A, \hat{L}_n^T, \hat{R}_n, \hat{S}_n) \Rightarrow (\hat{W}, \hat{L}, \hat{T}, \hat{A}, \hat{L}, \hat{W}, 0e, 0e), \tag{8}$$

where $P((\hat{T}, \hat{A}) \in \mathcal{C}^2) = 1$ and the limit processes are related by

$$\hat{A}(t) = -\lambda \hat{T}(\lambda t) \quad and \quad \hat{L}(t) = \hat{W}(\lambda t) - L\hat{T}(\lambda t), \quad t \geq 0. \tag{9}$$

Moreover, if one of the five limits in (7) holds with zero-mean two-dimensional Brownian motion (BM) as a limit, then the limit in (8) is also a zero-mean multivariate BM. In that case the variance and covariance terms of $\hat{T}(1), \hat{W}(1)$ and $\hat{A}(1), \hat{L}(1)$ are related by

$$\begin{aligned} &\sigma_A^2 \equiv \sigma_\lambda^2 = \lambda^3 \sigma_T^2, \qquad \sigma_{A,L}^2 \equiv \sigma_{\lambda,L}^2 = \lambda^2 (L\sigma_T^2 - \sigma_{T,W}^2), \\ &\sigma_L^2 = \lambda (L^2 \sigma_T^2 - 2L\sigma_{T,W}^2 + \sigma_W^2), \\ &\sigma_W^2 = \lambda^{-1}\sigma_L^2 - 2\lambda^{-2} L \sigma_{A,L}^2 + \lambda^{-3} L^2 \sigma_L^2. \end{aligned}$$

From (8) and (9), we see that the 8-dimensional limit in (8) is essentially 2-dimensional. Sufficient conditions for the FCLT to hold based on regenerative structure are established in [7].

We cannot add convergence of the pair $(\hat{A}_n, \hat{T}_n)$ to the list of five sufficient condition in (7), because these two random elements provide alternative characterizations of the arrival process alone. Thus, from these two, we cannot extract any information about the waiting times, and we cannot extract full information about the number in system.

3. Natural estimators based on data over $[0, t]$

We now discuss the implications of Theorem 1 for the estimation. The results for $\hat{L}_n$ and $\hat{A}_n$ apply directly to the natural estimators $\bar{L}(t)$ and $\bar{\lambda}(t)$, but we have not yet considered $\hat{W}(t)$. For that purpose, let

$$\hat{W}_n^e(t) \equiv t\sqrt{n}(\bar{W}(nt) - W), \quad t \geq 0, \tag{10}$$

where $\bar{W}(t)$ is defined in (1).

We will show that, for the most part, the random elements $\hat{W}_n^e$ and $\hat{W}_n$ are asymptotically equivalent after a deterministic space and time change by the arrival rate. For the statement, let $\stackrel{d}{=}$ denote equality in distribution (as processes); let $\|\cdot\|_{t_1,t_2}$ denote the uniform norm over $[t_1, t_2]$; let $\circ$ be the composition function, i.e., $(x \circ y)(t) \equiv x(y(t))$; and let $\mathcal{D}_0$ denote the space $\mathcal{D}$ over the open interval $(0, \infty)$, with the usual topology of uniform convergence over all bounded subintervals. We work with $\mathcal{D}_0$ to avoid problems in the neighborhood of 0, because there could be division by 0, since we have divided by $A(t)$, which could be 0 for some $t > 0$.

Corollary 3.1 (*Limit for $\hat{W}_n^e$*). *Under the conditions of Theorem 1,*

$$\|\hat{W}_n^e - \lambda^{-1}(\hat{L}_n - W\hat{A}_n)\|_{t_1,t_2} \Rightarrow 0 \quad for\ all\ t_2 > t_1 > 0, \tag{11}$$

for $\hat{W}_n^e$ defined in (10), so that

$$\|\hat{W}_n^e - \lambda^{-1}(\hat{W}_n \circ \lambda e)\|_{t_1,t_2} \Rightarrow 0 \quad for\ all\ t_2 > t_1 > 0, \tag{12}$$

and

$$\hat{W}_n^e \Rightarrow \hat{W}^e \equiv \lambda^{-1}(\hat{L} - W\hat{A}) = \lambda^{-1}(\hat{W} \circ \lambda e) \quad in\ \mathcal{D}_0. \tag{13}$$

For the common case in which $\hat{W}$ is a zero-mean BM, $\hat{W}^e \stackrel{d}{=} \lambda^{-1/2}\hat{W}$.

Proof. From the definitions in (5), (6) and (10), we have

$$\hat{W}_n^e(t) = \frac{t(\hat{W}_n^A(t) - W\hat{A}_n(t))}{n^{-1}A(nt)}.$$

Since $n^{-1/2}\hat{A}_n \Rightarrow 0$ as $n \to \infty$, the relation in (11) is valid. Theorem 1 then implies (12), which in turn implies (13). The final relation in (13) follows from (9). The final relation for the Brownian case follows from the last expression in (13) and the basic scaling property of BM: $\hat{W} \circ ce \stackrel{d}{=} \sqrt{c}\hat{W}$ for any positive constant c. □

Combining Theorem 1, Corollary 3.1 and the definition in (10), we obtain the corresponding ordinary CLT by applying the continuous mapping theorem with projection at $t = 1$ and letting n run through a continuous variable.

4. Estimators exploiting Little's law

Now we turn to the use of the relation $L = \lambda W$ to estimate each of the three parameters L, λ and W in terms of the other two. In addition to the three natural estimators in (1) for estimation using data over $[0, t]$, we have the alternative estimators exploiting Little's law in (3). Paralleling (10), have the associated FCLT-scaled random elements

$$\begin{aligned} \hat{L}_n^{\lambda,W}(t) &\equiv t\sqrt{n}\left(\bar{L}_{\lambda,W}(nt) - L\right) = \hat{W}_n^A(t), \\ \hat{\lambda}_n^{L,W}(t) &\equiv t\sqrt{n}\left(\bar{\lambda}_{L,W}(nt) - \lambda\right), \\ \hat{W}_n^{\lambda,L}(t) &\equiv t\sqrt{n}\left(\bar{W}_{\lambda,L}(nt) - W\right). \end{aligned} \tag{14}$$

We next state a corollary establishing the limiting behavior of $\hat{\lambda}_n^{L,W}$ and $\hat{W}_n^{\lambda,L}$; the proof is essentially the same as for Corollary 3.1 and so is omitted.

Corollary 4.1 (*Asymptotic form of $\hat{\lambda}_n^{L,W}$ and $\hat{W}_n^{\lambda,L}$*). *Under the conditions of Theorem 1,*

(a) $\|\hat{\lambda}_n^{L,W} - \hat{A}_n\|_{t_1,t_2} \Rightarrow 0$ *for all $t_2 > t_1 \geq 0$,*

for $\hat{\lambda}_n^{L,W}$ defined in (14), so that $\hat{W}_n^{\lambda,L} \Rightarrow \hat{A}$ in $\mathcal{D}$.

(b) $\|\hat{W}_n^{\lambda,L} - \lambda^{-1}(\hat{L}_n - W\hat{A}_n)\|_{t_1,t_2} \Rightarrow 0$ *for all $t_2 > t_1 > 0$,*

for $\hat{W}_n^{\lambda,L}$ defined in (14), so that

$$\|\hat{W}_n^{\lambda,L} - \hat{W}_n^e\|_{t_1,t_2} \Rightarrow 0 \quad for\ all\ t_2 > t_1 > 0 \quad and$$
$$\hat{W}_n^{\lambda,L} \Rightarrow \hat{W}^e \quad in\ \mathcal{D}_0,$$

where $\hat{W}_n^e$ and $\hat{W}^e$ are as in Corollary 3.1.

We can combine Theorem 1 and Corollaries 3.1 and 4.1 to obtain the ordinary joint CLT and establish (4) by applying the continuous mapping theorem with projection at $t = 1$ and letting n run through a continuous variable.

5. Converse for convergence preservation with composition

As a basis for completing the proof of the new part of Theorem 1 stating that convergence of $(\hat{L}_n, \hat{W}_n)$ in (7) implies the desired conclusion, we now establish a converse to the preservation of convergence by composition with centering in Corollary 13.3.1 in [20]. Let $x \circ y$ and y^{-1} denote the composition and inverse functions, respectively, as in Chapter 13 of [20]. A key assumption for this positive result is the monotonicity of the outer functions as well as the inner functions, which applies in our case. Let $\mathcal{D}_m$ be the subset of nondecreasing functions in $\mathcal{D} \equiv \mathcal{D}([0, \infty), \mathbb{R})$; let $\mathcal{D}_u$ be the subset of functions in $\mathcal{D}_m$ that are nonnegative and unbounded above.

Theorem 2 (*Converse for Composition with Centering*)**.** *Suppose that*

$$(x_n \circ y_n - c_n e, c_n(y_n - e)) \to (x, y) \quad \text{in } \mathcal{D}^2 \text{ as } n \to \infty, \tag{15}$$

where $x_n \in \mathcal{D}_m$, $y_n \in \mathcal{D}_u$ for all n, $(x, y) \in \mathcal{C}^2$, with $y(0) = 0$ and $c_n \to \infty$ as $n \to \infty$. Then

$$(x_n - c_n e) \to x - y \quad \text{in } \mathcal{D} \text{ as } n \to \infty. \tag{16}$$

Proof. By Theorem 13.7.1 of [20], the limit for the second term in (15) implies that

$$c_n(y_n^{-1} - e) \to -y \quad \text{in } \mathcal{D} \text{ as } n \to \infty. \tag{17}$$

Then, combining (17) with the first limit in (15), we can apply Corollary 13.3.1 of [20] to obtain

$$x_n \circ y_n \circ y_n^{-1} - c_n e \to x - y \quad \text{in } \mathcal{D} \text{ as } n \to \infty. \tag{18}$$

We now want to show that, under the stated conditions, (18) implies (16). That is expected since the limits in (15) and (17) imply that $y_n \to e$ and $y_n^{-1} \to e$ as $n \to \infty$, so that $y_n \circ y_n^{-1} \to e$, but it is not immediate because there is the initial factor c_n, where $c_n \to \infty$.

Let $\|\cdot\|_t$ be the uniform norm over the interval $[0, t]$. For a function $x \in \mathcal{D}$, let ω be the modulus of continuity

$$w_{x,T}(\delta) \equiv \sup_{0 \le t \le u \le t+\delta \le T} \{|x(u) - x(t)|\}.$$

In the remaining proof, we will use bounds on the uniform distance over bounded subintervals. We are thus using the fact that convergence to continuous limit functions in $\mathcal{D}$ with domain $[0, \infty)$ is equivalent uniform convergence of the restrictions to a bounded interval $[0, T]$ for all T. To be sure that a bound for $x_n \circ y_n \circ y_n^{-1}$ has direct implications for x_n, we work with the larger interval $[0, T + 1]$ in intermediate steps. Since the limits x and y are assumed to be continuous, there are no genuine difficulties.

For given $\epsilon > 0$ and $T > 0$, choose n_0 such that, for all $n \ge n_0$,

(i) $\|c_n(y_n^{-1} - e) + y\|_{T+1} \le \epsilon$,
(ii) $\|c_n(y_n - e) - y\|_{T+1} \le \epsilon$,
(iii) $\|x_n \circ y_n \circ y_n^{-1} - c_n e - (x - y)\|_{T+1} \le \epsilon$,
(iv) $\omega_{x-y}(2\epsilon/c_n, T + 1) \le \epsilon$. (19)

First, (i), (ii) and (iii) are possible by (15), (17) and (18), respectively. Then (iv) is possible because $x - y \in \mathcal{C}$; e.g., see Theorem 11.6.2 of [20].

We now establish a string of inequalities at different times t. Let time arguments falling outside the designated interval be interpreted as the nearest endpoint; i.e., replace any t outside the interval by the nearest boundary point, e.g., for interval $[0, T + 1]$, use $t \vee 0 \wedge (T + 1)$. With that convention, we apply (ii) and then (i) to conclude that

$$\begin{aligned} y_n \circ y_n^{-1}\left(t + \frac{\epsilon}{c_n}\right) &\ge y_n^{-1}\left(t + \frac{\epsilon}{c_n}\right) - \left(\frac{\epsilon + y\left(t + \frac{\epsilon}{c_n}\right)}{c_n}\right) \\ &\ge \left(t + \frac{\epsilon}{c_n}\right) - \left(\frac{\epsilon - y\left(t + \frac{\epsilon}{c_n}\right)}{c_n}\right) - \left(\frac{\epsilon + y\left(t + \frac{\epsilon}{c_n}\right)}{c_n}\right) \\ &= \left(t - \frac{\epsilon}{c_n}\right). \end{aligned} \tag{20}$$

Reasoning the same way in the other direction, we obtain the relation

$$\left(y_n \circ y_n^{-1}\right)\left(t - \frac{\epsilon}{c_n}\right) - \frac{\epsilon}{c_n} \le t \le \left(y_n \circ y_n^{-1}\right)\left(t + \frac{\epsilon}{c_n}\right) + \frac{\epsilon}{c_n} \tag{21}$$

for all t with $0 \le t \le T + 1$. We now substitute $t - (\epsilon/c_n)$ for t on the left and $t + (\epsilon/c_n)$ for t on the right in (21), exploit the assumed monotonicity of x_n for each n and subtract $c_n e$, to conclude that

$$\begin{aligned} &\left(x_n \circ y_n \circ y_n^{-1}\right)\left(t - \frac{2\epsilon}{c_n}\right) - c_n e \\ &\le x_n(t) - c_n e \le \left(x_n \circ y_n \circ y_n^{-1}\right)\left(t + \frac{2\epsilon}{c_n}\right) - c_n e \end{aligned} \tag{22}$$

for all t, $0 \le t \le T + 1$. Using (iii) and (iv) in (19) with (22), we get

$$(x - y)(t) - 2\epsilon \le x_n(t) - c_n e \le (x - y)(t) + 2\epsilon \tag{23}$$

for $n \ge n_0$. That proves (16). □

For the sake of completeness, even though we do not need it here, we now show that there is a similar result without centering, which is easier to prove.

Theorem 3 (*Converse for Composition without Centering*)**.** *Suppose that*

$$(x_n \circ y_n, y_n) \to (x, y) \quad \text{in } D^2 \text{ as } n \to \infty, \tag{24}$$

where $x_n \in \mathcal{D}_m$, $y_n \in \mathcal{D}_u$ for all n, $(x, y) \in \mathcal{C}^2$, y is strictly increasing, with $y(0) = y^{-1}(0) = 0$. Then

$$x_n \to x \circ y^{-1} \quad \text{in } \mathcal{D} \text{ as } n \to \infty. \tag{25}$$

Proof. First, we apply Theorems 13.6.1 and 13.2.2 of [20] to get the convergence $y_n^{-1} \to y^{-1}$, $y_n \circ y_n^{-1} \to e$ and $x_n \circ y_n \circ y_n^{-1} \to x \circ y^{-1}$ from condition (24). We want to show that these limits plus the other conditions imply (25).

Analogous to (19), for given $\eta > 0$ and $T > 0$, choose $\epsilon < \eta \wedge 1$ so that

$$\omega_{x \circ y^{-1}}(2\epsilon, T + 1) < \eta. \tag{26}$$

Then for that given ϵ and T, first choose $T_1 = y^{-1}(T)$ and $\delta > 0$ such that $\delta < \epsilon$ and $\omega_y(\delta, T + 1) < \epsilon$. Then choose n_0 such that, for all $n \ge n_0$,

(i) $\|y_n^{-1} - y^{-1}\|_{T+1} \le \delta$,
(ii) $\|y_n - y\|_{T_1+1} \le \epsilon$,
(iii) $\|x_n \circ y_n \circ y_n^{-1} - x\|_{T+1} \le \epsilon$. (27)

As in the previous proof, we now establish a string of inequalities at different times t. As before, let time arguments falling outside the designated interval be interpreted as the nearest endpoint. Assume that $n \ge n_0$. Paralleling (20), we can apply (ii), (i) and then (26) to conclude that

$$y_n \circ y_n^{-1}(t) \ge y(y_n^{-1}(t)) - \epsilon \ge y(y^{-1}(t) - \delta) - \epsilon \ge t - 2\epsilon. \tag{28}$$

Reasoning the way on the other side and looking at times $t \pm \epsilon$, paralleling (22), we get

$$y_n \circ y_n^{-1}(t - 2\epsilon) \le t \le y_n \circ y_n^{-1}(t + 2\epsilon) \quad (29)$$

for all $t \in [0, T + 1]$. Since $x_n \in \mathcal{D}_m$, from (29) we obtain

$$(x_n \circ y_n \circ y_n^{-1})(t - 2\epsilon) \le x_n(t) \le (x_n \circ y_n \circ y_n^{-1})(t + 2\epsilon).$$

Applying (iii) in (27), we then get

$$(x \circ y^{-1})(t - 2\epsilon) - \epsilon \le x_n(t) \le (x \circ y^{-1})(t + 2\epsilon) + \epsilon.$$

Finally, from (26), we get

$$(x \circ y^{-1})(t) - 2\eta \le x_n(t) \le (x \circ y^{-1})(t) + 2\eta$$

for $0 \le t \le T + 1$ and $n \ge n_0$. That proves the desired limit in (25). □

We now give a counterexample showing that the monotonicity condition for the outer function x_n is necessary.

Example 5.1 (*Counterexample without Monotonicity in* Theorem 3). Let the interior function be defined by $y_n(t) = t$ for $0 \le t < 1/2$ and for $(1/2) + (1/n) \le t < \infty$, but let $y_n(t) = (1/2) + (1/n)$ for $(1/2) \le t < (1/2) + (1/n)$. It is easy to see that $y_n \in \mathcal{D}_u$ and $y_n \to e$, but y_n has a discontinuity. Let $x_n(t) = t$ for $0 \le t < (1/2)$ and for $(1/2) + (1/n) \le t < \infty$, but let $x_n((1/2) + (1/2n)) = n$ and let x_n be defined by linear interpolation in the two subintervals $[(1/2), (1/2)+(1/2n)]$ and $[(1/2)+(1/2n), (1/2)+(1/n)]$, so that x_n is a continuous function for all n. Clearly, $x_n((1/2) + (1/2n)) = n \to \infty$ as $n \to \infty$, so the sequence $\{x_n : n \ge 1\}$ does not converge pointwise. However, $x_n \circ y_n = y_n$, so that $x_n \circ y_n \to e$ in $\mathcal{D}$. Hence, we have all the conditions of Theorem 3 satisfied except that x_n is not monotone, but we fail to have convergence of x_n. The same example works for Theorem 2 in the common case that $c_n/n \to 0$.

6. Proof of Theorem 1 for the new sufficient condition

Given that the scaled remainder terms $\hat{R}_n$ and $\hat{S}_n$ are asymptotically negligible by assumption, $\|\hat{L}_n - \hat{W}_n^A\|_T \Rightarrow 0$ as $n \to \infty$ for $\hat{L}_n$ in (5) and $\hat{W}_n^A$ in (6). Hence, we can replace the new condition with $(\hat{W}_n^A, \hat{W}_n) \Rightarrow (\hat{L}, \hat{W})$ in $\mathcal{D}^2$. We now want to apply Theorem 2, but we see that it is not in the right form, because here the inner process is A_n, for which we have no given limit. To put this in the setting of Theorem 2, we invert these processes, using the fact that both are in $\mathcal{D}_u$.

In particular, let $\bar{A}_n(t) \equiv n^{-1}A(nt)$, $\bar{W}_n(t) \equiv n^{-1}\sum_{k=1}^{\lfloor nt \rfloor}$ and $\bar{W}_n^A(t) \equiv n^{-1}\sum_{k=1}^{\lfloor A(nt) \rfloor}$. Then let $\bar{B}_n \equiv \bar{A}_n^{-1}$, $\bar{V}_n \equiv \bar{W}_n^{-1}$ and $\bar{V}_n^A \equiv (\bar{W}^A)_n^{-1}$. ($\bar{B}_n$ is intimately related to the average of the interarrival times $\bar{T}_n$; see Section 13.88 of [20].) Since we assume that the relation $L = \lambda W$ is valid, there are SLLN's, so that $\bar{A}_n$ and $\bar{W}_n$ are elements of $\mathcal{D}_u$. Hence,

$$\bar{V}_n^A \equiv (\bar{W}^A)_n^{-1} = (\bar{W}_n \circ \bar{A}_n)^{-1} = \bar{B}_n \circ \bar{V}_n. \quad (30)$$

In (30), $\bar{B}_n$, the inverse of $\bar{A}_n$, appears as the outer process, putting us in the setting of Theorem 2.

Since $\hat{W}_n \equiv \sqrt{n}(\bar{W}_n - We) \Rightarrow \hat{W}$, Corollary 13.7.3 of [20] implies that $\sqrt{n}(\bar{V}_n - W^{-1}e) \Rightarrow -W^{-1}(\hat{W} \circ W^{-1}e)$. Multiplying through by W yields

$$\sqrt{n}(W\bar{V}_n - e) \Rightarrow -\hat{W} \circ W^{-1}e. \quad (31)$$

Next, exploiting (30), since $\hat{W}_n^A \equiv \sqrt{n}(\bar{W}_n \circ \bar{A}_n - Le) \Rightarrow \hat{L}$, Corollary 13.7.3 of [20] implies that

$$\sqrt{n}(\bar{B}_n \circ \bar{V}_n - L^{-1}e) \Rightarrow -L^{-1}(\hat{L} \circ L^{-1}e).$$

We then can rewrite this limit as

$$(\sqrt{n}L\bar{B}_n) \circ W^{-1}e \circ W\bar{V}_n - \sqrt{n}e \Rightarrow -\hat{L} \circ L^{-1}e. \quad (32)$$

We can now apply Theorem 2 with (31) and (32) to obtain

$$(\sqrt{n}L\bar{B}_n) \circ W^{-1}e - \sqrt{n}e \Rightarrow -\hat{L} \circ L^{-1}e + \hat{W} \circ W^{-1}e.$$

Transforming time by W, dividing by L and exploiting $L = \lambda W$ yields

$$\hat{B}_n \equiv \sqrt{n}(\bar{B}_n - \lambda^{-1}e) \Rightarrow \hat{B} \equiv -L^{-1}(\hat{L} \circ \lambda^{-1}e) + L^{-1}\hat{W}. \quad (33)$$

Applying corollary Theorem 13.7.3 of [20] once again, now to (33), yields

$$\hat{A}_n \equiv \sqrt{n}(\bar{A}_n - \lambda e) \Rightarrow \hat{A} \equiv \lambda L^{-1}\hat{L} - \lambda L^{-1}(\hat{W} \circ \lambda e) = W^{-1}\hat{L} - W^{-1}(\hat{W} \circ \lambda e).$$

From this limit, we see that $\hat{L} = \hat{W} \circ \lambda e + W\hat{A}$, which agrees with (9). □

7. Extensions

The references show that there are other conservation laws closely related to $L = \lambda W$, notably $H = \lambda G$. The seemingly minor extension to $H = \lambda G$ is surprisingly far-reaching; see Remark 6.6 of [18] and [19] and references therein. Under appropriate regularity conditions, both the FCLT and the statistical analysis extend to these related settings. In particular, Theorem 1 extends to the formulation of $H = \lambda G$ in Theorem 6.1 of [18], because the instantaneous cost rate of customer k at time t, $f_k(t)$, is assumed to be nonnegative. That monotonicity is an important condition in the theorems in Section 5. In view of Remark 6.6 of [18] and [19], the paper [14] is relevant.

Acknowledgment

This research was supported by NSF grant CMMI 1066372.

References

[1] F. Baccelli, P. Bremaud, Elements of Queueing Theory: Palm-Martingale Calculus and Stochastic Recurrences, second ed., Springer, New York, 2003.
[2] P. Bratley, B.L. Fox, L.E. Schrage, A Guide to Simulation, second ed., Springer, New York, 1987.
[3] J.P. Buzen, Fundamental operational laws of computer system performance, Acta Inform. 7 (1976) 167–182.
[4] P.J. Denning, J.P. Buzen, The operational analysis of queueing network models, Comput. Surv. 10 (1978) 225–261.
[5] M. El-Taha, S. Stidham Jr., Sample-Path Analysis of Queueing Systems, Kluwer, Boston, 1999.
[6] P.W. Glynn, W. Whitt, A central-limit-theorem version of $L = \lambda W$, Queueing Syst. 1 (1986) 191–215.
[7] P.W. Glynn, W. Whitt, Sufficient conditions for functional-limit-theorem versions of $L = \lambda W$, Queueing Syst. 1 (1987) 279–287.
[8] P.W. Glynn, W. Whitt, Ordinary CLT and WLLN versions of $L = \lambda W$, Math. Oper. Res. 13 (1988) 674–692.
[9] P.W. Glynn, W. Whitt, Indirect estimation via $L = \lambda W$, Oper. Res. 37 (1989) 82–103.
[10] S. Kim, W. Whitt, Statistical analysis with Little's law, Columbia University, 2012. Available at: http://www.columbia.edu/~ww2040/allpapers.html.
[11] A.M. Law, Efficient estimators for simulated queueing systems, Manage. Sci. 22 (1975) 30–41.
[12] J.D.C. Little, A proof of the queueing formula: $L = \lambda W$, Oper. Res. 9 (1961) 383–387.
[13] J.D.C. Little, Little's law as viewed on its 50th anniversary, Oper. Res. 59 (2011) 536–539.
[14] G. Nieuwenhuis, Equivalence of functional limit theorems for stationary point processes and their Palm distributions, Probab. Theory Related Fields 81 (1989) 593–608.
[15] R. Serfozo, Introduction to Stochastic Networks, Springer, New York, 1999.
[16] K. Sigman, Stationary Marked Point Processes, An Intuitive Approach, Chapman and Hall, New York, 1995.
[17] S. Stidham Jr., A last word on $L = \lambda W$, Oper. Res. 22 (1974) 417–421.
[18] W. Whitt, A review of $L = \lambda W$ and extensions, Queueing Syst. 9 (1991) 235–268.
[19] W. Whitt, $H = \lambda G$ and the Palm transformation, Adv. in Appl. Probab. 24 (1992) 755–758.
[20] W. Whitt, Stochastic-Process Limits, Springer, New York, 2002.
[21] R.W. Wolff, Little's law and related results, in: J.J. Cochran (Ed.), Wiley Encyclopedia of Operations Research and Management Science, Wiley, New York, 2011.