Reprinted from:

Operational Research '75

Edited by

K.B. HALEY

University of Birmingham

N·H
P C

1976

NORTH-HOLLAND PUBLISHING COMPANY
AMSTERDAM · NEW YORK · OXFORD

K.B. Haley, ed., Operational Research '75. North-Holland Publishing Company (1976)

A NEW DERIVATION OF THE KHINTCHINE–POLLACZEK FORMULA*

Nouvelle derivée de la formule Khintchine–Pollaczek

URI YECHIALI

Department of Statistics, Tel Aviv University, Ramat Aviv, Tel Aviv, Israel

Abstract. For the M/G/1 queue a new derivation of the mean number of units in system is given. The method is based on calculating the expected total waiting time incurred per transition.

Résumé. Pour la file d'attente M/G/1, une nouvelle dérivation du nombre moyen des unités du systéme est indiquée. La méthode est basée sur le calcul du temps total engendré et prévu par transition.

Consider the M/G/1 queueing system. Let λ be the rate of the Poisson input process, and let $B(\cdot)$ denote the distribution function of the service time V for which $\mathbf{E}(V) < \infty$. As usual, the system is observed at epochs of service completion, and let X_k denote the number of people the kth departing customer leaves behind him. We define X_k to be the state of the embedded Markov chain until the next person completes service.

It is well known (see, e.g., [2]) that if $\rho \equiv \lambda\mathbf{E}(V) < 1$, then the stationary distribution $\mathbf{\Pi} = (\Pi_0, \Pi_1, \Pi_2, \ldots)$ of the Markov chain $\{X_k\}$ exists and satisfies $\Pi_i > 0$, $\Sigma_{i=0}^{\infty} \Pi_i = 1$ and $\Pi_j = \Pi_0 a_j + \Sigma_{i=1}^{j+1} \Pi_i a_{j-i+1}$ $(j = 0, 1, 2, \ldots)$, where $a_j = \int_0^\infty [\exp\{-\lambda v\}(\lambda v)^j/j!]\, \mathrm{d}B(v)$. Moreover, if $\rho < 1$, $\Pi_0 = 1 - \rho$.

* This paper appears in substantially the same form in the Journal of Applied Probability, December 1975.

Our purpose in this note is to give a new derivation of the, so called, Khintchine–Pollaczek formula [3] for the mean number of units in system L. That is, we want to show that

$$L = \rho + \frac{\lambda^2 \mathbf{E}(V^2)}{2(1-\rho)}.$$

Our approach will be to obtain the mean total waiting time *per transition*. This method, we believe, may be employed in analyzing other queueing models by using results (1) and (2) in the sequel. It may also be used when dealing with Markov renewal decision processes of queueing type.

We proceed now to calculate L. To do this we let $W(X_k)$ be the total waiting time incurred at the kth step, and define its expectation $\mathbf{E}W(X_k)$ as a functional of the Markov chain. We wish to find ϕ, the long-run average expected waiting time per transition, where

$$\phi = \lim_{T\to\infty} \sum_{k=0}^{T} \mathbf{E}W(X_k)/(T+1).$$

From the theory of Markov chains it is well known [1] that, with probability 1, $\phi = \Sigma_{j=0}^{\infty} \Pi_j w_j$ where we write $w_j = \mathbf{E}W(j)$. It thus remains to obtain w_j for all j.

Theorem. *For $j \geqslant 1$, $w_j = j\mathbf{E}(V) + \lambda\mathbf{E}(V^2)/2$.*

Proof. Consider the system in state j. We write $w_j = \mathbf{E}W_1(j) + \mathbf{E}W_2(j)$, where $\mathbf{E}W_1(j)$ denotes the expected total waiting time incurred by all units currently in system, and $\mathbf{E}W_2(j)$ is the expected total waiting time of all customers who arrived *during* the sojourn time of the process in state j. It is clear that, for $j \geqslant 1$, $\mathbf{E}W_1(j) = j\mathbf{E}(V)$.

To calculate $\mathbf{E}W_2(j)$ we consider the instants of arrival while the system is in state j. If the service time is v and the number of arrivals $N(v)$ during the time interval $[0, v]$ is $n \geqslant 1$, then it is known [2] that the arrival instants S_1, $S_2, \ldots, S_n$ have the distribution of the order statistics from a sample of n observations taken from the uniform distribution on $[0, v]$. That is, the p.d.f. of the S_i is given by

$$f_{S_1,S_2,\ldots,S_n}(s_1, s_2, \ldots, s_n \mid N(v) = n) = \frac{n!}{v^n}, \qquad 0 \leqslant s_1 \leqslant s_2 \leqslant \ldots \leqslant s_n \leqslant v.$$

Since a customer who arrives at instant s waits exactly $v-s$ units of time before service completion, we have

$$\mathbf{E}[W_2(j) \mid V = v, N(v) = n] = \int_0^v \int_{s_1}^v \ldots \int_{s_{n-1}}^v [(v - s_1) + (v - s_2) + \ldots + (v - s_n)] \times (n!/v^n)\, \mathrm{d}s_n \ldots \mathrm{d}s_2\, \mathrm{d}s_1 .$$

For $n = 1$ we readily have $\mathbf{E}[W_2(j) \mid V = v, N(v) = 1] = v/2$. Using an inductive argument, one obtains

$$\mathbf{E}[W_2(j) \mid V = v, N(v) = n] = n(v/2). \tag{1}$$

Better: $E[W_2(j) \mid V = v, N(v) = n] = E\left[\sum_{i=1}^{n} (v - S_i)\right] = nv - \sum_{i=1}^{n} ES_i = nv - n\frac{v}{2} = n\frac{v}{2}$

Hence, independent of j,

$$\mathbf{E}W_2(j) = \int_0^\infty \left(\sum_{n=0}^{\infty} \mathbf{E}[W_2(j) \mid V = v, N(v) = n]\, P\{N(v) = n\} \right) \mathrm{d}B(v) \tag{2}$$

$$= \int_0^\infty \left(\sum_{n=0}^{\infty} \frac{nv}{2}\, \mathrm{e}^{-\lambda v} \frac{(\lambda v)^n}{n!} \right) \mathrm{d}B(v) = \lambda \mathbf{E}(V^2)/2.$$

This completes the proof.

Result (2) may be rewritten in a slightly different way and interpreted intuitively. We write

$$\mathbf{E}W_2(j) = [\lambda \mathbf{E}(v)] \left[\frac{\mathbf{E}(v^2)}{2\mathbf{E}(v)} \right].$$

The second term in brackets is the mean value of the, so called, random modification which is the waiting time (until service completion) of an arbitrary unit that arrives during a service period. The expected number of arrivals during an average service time of length $\mathbf{E}(V)$ is, obviously, $\lambda \mathbf{E}(V)$. The product of these two terms gives the result for $\mathbf{E}W_2(j)$.

For $j = 0$, no waiting time is incurred before a unit arrives, and from that moment until its departure the expected total waiting time is $\mathbf{E}(V) + \lambda \mathbf{E}(V^2)/2$. That is, $w_0 = w_1$. Now,

$$\varphi = \sum_{j=0}^{\infty} \Pi_j w_j = \sum_{j=1}^{\infty} \Pi_j w_j + \Pi_0 w_1 = \mathbf{E}(V)L + \Pi_0 \mathbf{E}(V) + \lambda \mathbf{E}(V^2)/2,$$

where we used the relation $L = \Sigma_{j=1}^{\infty} j\, \Pi_j$.

The mean sojourn time of a transition is

$$\Pi_0[1/\lambda + \mathbf{E}(V)] + (1 - \Pi_0)\mathbf{E}(V) = 1/\lambda.$$

Thus, the expected total waiting time incurred per unit time is $\lambda\phi$, which, in turn, must be equal to L. Hence, using $\Pi_0 = 1 - \rho$, we have

$$L = \lambda\phi = \rho L + (1 - \rho)\rho + \lambda^2 \mathbf{E}(V^2)/2$$

from which it follows that

$$L = \rho + \frac{\lambda^2 \mathbf{E}(V^2)}{2(1 - \rho)}.$$

References

[1] K.I. Chung, *Markov Chains with Stationary Transition Probabilities* 2nd ed. (Springer, New York, 1967).
[2] S. Karlin, *A First Course in Stochastic Processes* (Academic Press, New York, 1968).
[3] T.L. Saaty, *Elements of Queueing Theory with Applications* (McGraw-Hill, New York, 1961).