May 1982 LIDS-P-1207

DETERMINISM MINIMIZES WAITING TIME IN QUEUES*

by

Pierre A. Humblet**

---

*The author is with the Dept. of Electrical Engineering and Computer Science and the Laboratory for Information and Decision Systems at the Mass. Institute of Technology, Cambridge, Massachusetts 02139.

**This research was conducted at the M.I.T. Laboratory for Information and Decision Systems with partial support provided by the National Science Foundation under Contract NSF-ECS 79-19880.

This paper gives an elementary proof of a "folk theorem" of queueing theory, namely that for a given mean interarrival (or service) time, the average waiting time (and other related quantities) is minimum when the interarrival (service) times are constant. This is well known [1], [2] when the interarrival and service times are mutually independent. Hajeck [3] proved a similar result for general arrival processes (non necessarity renewal or stationary) when the service times are independent and exponentially distributed. We give a very general and elementary proof relying essentially on the convexity of the Max (•) function and on the use of Jensen's inequality. One method was inspired by Friedman [4]. In light of their simplicity, these results may be known to many.

Definitions: Let t_i and s_i, $i = 0,1,2,\ldots$ be sequences of non negative random variables, possibly dependent and non stationary.

Let $\mathcal{T}_n$ and $\mathcal{S}_n$, $n = 1,2,\ldots$ be the smallest σ-fields generated by $(t_o, t_1 \ldots t_{n-1})$ and $(s_o, s_1 \ldots s_{n-1})$ respectively.

Let $T_{n,k} = \sum_{i=n-k}^{n-1} t_n$ $k = 1,2,\ldots n$; $T_{n,0} = 0$ $n = 1,2\ldots$

$S_{n,k} = \sum_{i=n-k}^{n-1} s_i$ $k = 1,2,\ldots n$; $S_{n,0} = 0$ $n = 1,2,\ldots$

It is well known that

$$w_n = \underset{k=0,1,\ldots n}{\text{Max}} (S_{n,k} - T_{n,k})$$

can be interpreted as the waiting time of the n^{th} customer in a first come, first served queue, where the service time of the i^{th} customer is s_i, the interarrival time between the i^{th} and $(i+1)^{th}$ customer is t_i, and customer 0 finds the queue empty.

We will show here that

a) if (1) $E(t_i | S_n) = E\, t_i \quad n \geq 0,\ i \geq 0$

(2) $(s_i)_{i \geq 0}$ is stationary

(3) $\limsup_{n\to\infty} \frac{E\, T_{n,n}}{n} \leq t$

then $\liminf_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} E\, w_i \geq \lim_{n\to\infty} E(\max_{k=0,1,\ldots n} (S_{n,k} - kt))$.

The right hand side (which always exist if ∞ is allowed) can be interpreted as the limit of the expected wait of the n^{th} customer in a queue with same service process as the original queue, but interarrival times constant and equal to t. It is also equal to the expected "stationary waiting time" in that new queue.

b) (1) $E(s_i | T_n) = E\, s_i \quad n \geq 0,\ i \geq 0$

(2) $(t_i)_{i \geq 0}$ is stationary

(3) $\liminf_{n\to\infty} \frac{E\, S_{n,n}}{n} \geq s$

then $\liminf_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} E\, w_i \geq \lim_{n\to\infty} E\, (\max_{k=0,1,\ldots n} (k\, s - T_{n,k}))$.

The interpretation of this result is dual to the one in a), with services replacing interarrival times.

c) The previous results hold, mutatis mutandis, for convex and monotone functions of (s_i) and (t_i).

Proof:

$$E\,w_n = E(E[\max_{k=0,1,\dots n}(S_{n,k}-T_{n,k}) \mid S_n])$$

$$\geq E(\max_{k=0,1,\dots n}(E[S_{n,k}-T_{n,k} \mid S_n]))$$

$$= E(\max_{k=0,1,\dots n}(S_{n,k}-E\,T_{n,k}))$$

The inequality above results from the convexity of Max and Jensen's inequality. The last equality results from hypothesis (1).

At this point we have shown that $E\,w_n$ is minimized by a deterministic arrival process. It remains to be proved that it is best to have constant interarrival times.

Let $m \in Z^+$ and consider

$$\frac{1}{m+n}\sum_{i=1}^{m+n} E\,w_i \geq \frac{n}{m+n}\,\frac{1}{n}\sum_{i=m+1}^{m+n} E\max_{k=0,1,\dots m}(S_{i,k}-E\,T_{i,k})$$

Note that by stationarity (2) of $(s_i)_{i>0}$, for $(x_0,x_1\dots x_m) = \underline{x}$ given, $E(\max_{k=0,1,\dots m}(S_{i,k}-x_k))$ is a function $F(\underline{x})$ independent of i $(i \geq m)$. Moveover, F is convex in $\underline{x}$ and nonincreasing for each x_i. By Jensen's inequality

$$\frac{1}{m+n}\sum_{i=0}^{m+n} Ew_i \geq \frac{n}{m+n}\,F\left(\frac{\sum_{i=m+1}^{m+n} E\,T_{i,0}}{n}, \frac{\sum_{i=m+1}^{m+n} E\,T_{i,1}}{n}, \dots, \frac{\sum_{i=m+1}^{m+n} E\,T_{i,m}}{n}\right)$$

$$= \frac{n}{m+n}\,F\left(\frac{0}{n}, \frac{E\,T_{m+n,n}}{n}, \dots \frac{E\,T_{m+n,n} + \dots + E\,T_{n+1,n}}{n}\right)$$

The equality results from the fact that, going back to the definition of $T_{i,k}$, $\sum_{i=m+1}^{m+n} T_{i,k} = \sum_{j=0}^{k-1} T_{m+n-j,n}$.

Keeping m fixed and letting n grow yields

$$\liminf_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n} E(w_i) \geq F(0,t,2t,\ldots,mt) = E \max_{k=0,1\ldots m} (S_{m,k}-kt)$$

where we used hypothesis (3) and the nonincreasing nature of $F(\cdot)$.

The right hand side can be interpreted as the waiting time of the m^{th} customer in the queue with constant interarrival times. Letting m grow yields the desired result.

One can further note that, because of the stationarity of $(s_i)_{i>0}$, the probability distribution function of $\max_{k=0,\ldots,m} (S_{m,k}-kt)$ converges from above to some (possibly defective) distribution function that one may call the "stationary waiting time" distribution function.

By Lebesque Monotone Convergence Theorem, the expected value of the "stationary waiting time" is equal to the lower bound derived above.

The proofs of b) and the generalization to moments of other functions convex and monotone in (s_i) or (t_i) proceed along similar lines.

References

[1] Borovkov, A. A., Stochastic Processes in Queueing Theory, Springer-Verlag, New York, 1976.

[2] Stoyan, Eng. Cyb., Vol. 12, pp. 79-81, Nov. 1974.

[3] Hajek, B., "The Proof of a Folk Theorem on Queueing Delay with Applications to Routing in Networks", University of Illinois, April 1981.

[4] Friedman, D. U., "Queueing Analyais of a Shared Voice/Data Link", Ph.D. Thesis, Dept. of Electrical Engineering and Computer Science, Mass. Institute of Technology, November 1981; also Laboratory for Information and Decision Systems Report LIDS-TH-1161.