Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Technical Note—On Waiting Times for a Queue in Which Customers Require Simultaneous Service from a Random Number of Servers

Andrew F. Seila,

Technical Notes

On Waiting Times for a Queue in Which Customers Require Simultaneous Service from a Random Number of Servers

ANDREW F. SEILA
University of Georgia, Athens, Georgia

(Received February 1982; accepted January 1984)

We consider a queueing system, first introduced by L. Green in 1980, in which customers from a Poisson arrival stream request simultaneous service from a random number of identical servers with exponential service times. Computational formulas for the second moment of time in queue are given, along with tables of these values for selected systems. Numerical results show that the coefficient of variation for time in queue is always greater than 1 and decreases with increasing congestion.

GREEN [1980] introduced a multiserver queueing system in which customers from a Poisson arrival stream require simultaneous service from a random number of independent and identical exponential servers with a first-come-first-served queue discipline. This model has been suggested for such systems as hospital emergency rooms, loading docks, and maintenance systems. It is also a useful system for simulation methodologists to use in testing certain estimation methods (Seila [1980]). In this note, we present an expression for the second moment of the stationary waiting time (in queue). All notation and terminology are the same as in Green. In particular, λ denotes the arrival rate, μ the common service rate for each server, s the number of servers in the system, and c_j for $j = 1, 2, \cdots, s$ the probability that an arriving customer requests exactly j servers.

1. THE SECOND MOMENT OF TIME IN QUEUE

Green gives an expression for the Laplace-Stieltjes transform $\tilde{W}(s)$ of the stationary distribution of time in queue. A straightforward, but tedious, evaluation of $\lim_{s\to 0} d^2\tilde{W}(s)/ds^2$ yields an expression for the second (noncentral) moment. Define $b_j = E(B^j)$ and $d_j = E(D^j)$ for $j = 1, 2, 3, \cdots$, where B is the interservice time random variable and D is

Subject classification: 705 simultaneous service from random numbers of servers.

Operations Research
Vol. 32, No. 5, September–October 1984
0030-364X/84/3205-1181 $01.25

the initial delay random variable. Then, if $\lambda b_1 < 1$, the second moment of time in queue is given by

$$E(W^2) = [(1 - p_q)p_d/6(1 - \lambda b_1)^3]\{2(1 - \lambda b_1)^2(3d_2 - \lambda b_3 + \lambda d_3) + 2\lambda b_3(1 - \lambda b_1)(1 - \lambda b_1 + \lambda d_1) + 3\lambda b_2(1 - \lambda b_1)(2d_1 - \lambda b_2 + \lambda d_2) + 3(\lambda b_2)^2(1 - \lambda b_1 + \lambda d_1)\}. \quad (1)$$

The third moments of B and D are

$$b_3 = (3!/\mu^3) \sum_{k=1}^{s} c_k \sum_{h=s-k+1}^{s} (1/h) \sum_{j=h}^{s} (1/j) \sum_{l=j}^{s} (1/l),$$

and

$$d_3 = (3!/\mu^3 p_d) \sum_{i=1}^{s} \bar{q}_i \sum_{k=1}^{i} c_{s-i+k} \sum_{h=i-k+1}^{s} (1/h) \sum_{j=h}^{i} (1/j) \sum_{l=j}^{i} (1/l).$$

The mean time in queue is given by

$$E(W) = (p_q/\lambda)\{1 + (\lambda/2)[(d_2/d_1) + (\lambda b_2/(1 - \lambda b_1))]\}, \quad (2)$$

where

$$b_2 = (2/\mu^2) \sum_{k=1}^{s} c_k \sum_{j=s-k+1}^{s} (1/j) \sum_{l=j}^{s} (1/l),$$

$$d_2 = (2/\mu^2 p_d) \sum_{i=1}^{s} \bar{q}_i \sum_{k=1}^{i} c_k \sum_{j=i-k+1}^{i} (1/j) \sum_{l=j}^{i} (1/l),$$

and b_1 and d_1 are given in Green. With (1) and (2), the variance and coefficient of variation of waiting times can be computed. In addition, since service times are independent of waiting times, it is straightforward to compute the first two moments of total time in system.

Tables I through III give means and standard deviations of waiting times for a selected group of systems with various numbers of servers and various levels of congestion given by λb_1. The arrival rate for all systems is normalized to 1.0. The systems in Table I give equal probabilities to the numbers of servers requested by arriving customers

$$c_j = 1/s \quad \text{for} \quad j = 1, 2, \cdots, s,$$

where c_j is the probability that j servers are requested. For the systems in Table II, the probabilities are increasing

$$c_j = 2j/s(s + 1) \quad \text{for} \quad j = 1, 2, \cdots, s,$$

and the probabilities are decreasing for the systems in Table III

$$c_j = 2(s - j + 1)/s(s + 1) \quad \text{for} \quad j = 1, 2, \cdots, s.$$

These tables will be useful to simulation methodologists who need to know the first two moments of waiting time in order to evaluate estimators of $E(W)$. In each group of numbers, the mean is on top, the standard deviation is directly below it, and the coefficient of variation is

TABLE I
MEANS AND STANDARD DEVIATIONS OF WAITING TIMES: CONSTANT PROBABILITIES

b_1	Servers					
	2	3	4	5	7	9
0.3	0.108	0.010	0.095	0.093	0.090	0.088
	0.284	0.275	0.270	0.266	0.262	0.260
	2.64	2.76	2.83	2.87	2.93	2.96
0.5	0.444	0.421	0.409	0.401	0.391	0.386
	0.832	0.816	0.806	0.801	0.793	0.789
	1.87	1.94	1.97	2.00	2.03	2.05
0.7	1.529	1.483	1.457	1.440	1.419	1.407
	2.190	2.173	2.162	2.155	2.147	2.142
	1.43	1.47	1.48	1.50	1.51	1.52
0.8	3.067	3.006	2.971	2.948	2.920	2.904
	3.890	3.875	3.865	3.859	3.852	3.847
	1.27	1.29	1.30	1.31	1.32	1.33
0.9	7.935	7.857	7.812	7.782	7.746	7.724
	8.937	8.927	8.921	8.917	8.912	8.909
	1.13	1.14	1.14	1.15	1.15	1.15
0.95	17.868	17.781	17.730	17.697	17.655	17.631
	18.966	18.961	18.958	18.955	18.952	18.951
	1.06	1.07	1.07	1.07	1.07	1.07
0.99	97.814	97.719	97.664	97.627	97.581	97.554
	98.993	98.992	98.991	98.991	98.990	98.989
	1.01	1.01	1.01	1.01	1.01	1.01

TABLE II
MEANS AND STANDARD DEVIATIONS OF WAITING TIMES: INCREASING PROBABILITIES

b_1	Servers					
	2	3	4	5	7	9
0.3	0.107	0.096	0.089	0.084	0.077	0.072
	0.303	0.298	0.293	0.288	0.278	0.269
	2.84	3.11	3.30	3.43	3.62	3.74
0.5	0.457	0.431	0.411	0.396	0.373	0.357
	0.898	0.904	0.901	0.893	0.873	0.855
	1.97	2.10	2.19	2.25	2.34	2.40
0.7	1.611	1.580	1.547	1.517	1.465	1.423
	2.378	2.432	2.446	2.442	2.416	2.383
	1.48	1.54	1.58	1.61	1.65	1.67
0.8	3.262	3.255	3.224	3.185	3.109	3.044
	4.224	4.340	4.379	4.383	4.352	4.306
	1.29	1.33	1.35	1.38	1.40	1.41
0.9	8.510	8.627	8.632	8.592	8.473	8.352
	9.688	9.977	10.084	10.108	10.060	9.974
	1.14	1.16	1.17	1.18	1.19	1.19
0.95	19.232	19.644	19.752	19.730	19.550	19.333
	20.528	21.140	21.368	21.423	21.332	21.160
	1.07	1.08	1.08	1.09	1.09	1.09
0.99	105.57	108.46	109.46	109.64	109.05	108.12
	106.96	110.08	111.22	111.48	111.00	110.12
	1.01	1.01	1.02	1.02	1.02	1.02

on the bottom. Pascal routines for computing these quantities are available from the author.

It is interesting to note that, for this system, the coefficient of variation (c.v.) of waiting times is always greater than 1, and it increases as the congestion level decreases. For example, in Table III, the system with 5 servers and congestion level (λb_1) 0.3 has a mean waiting time of 0.090

TABLE III

MEANS AND STANDARD DEVIATIONS OF WAITING TIMES: DECREASING PROBABILITIES

b_1	Servers					
	2	3	4	5	7	9
0.3	0.106	0.097	0.093	0.090	0.087	0.085
	0.264	0.247	0.238	0.232	0.225	0.221
	2.50	2.54	2.57	2.58	2.60	2.61
0.5	0.426	0.396	0.380	0.370	0.358	0.352
	0.765	0.772	0.698	0.683	0.665	0.654
	1.80	1.95	1.84	1.85	1.86	1.86
0.7	1.436	1.352	1.306	1.277	1.242	1.222
	2.006	1.907	1.852	1.817	1.774	1.749
	1.40	1.41	1.42	1.42	1.43	1.43
0.8	2.857	2.706	2.622	2.568	2.504	2.467
	3.560	3.396	3.303	3.243	3.171	3.129
	1.25	1.26	1.26	1.26	1.27	1.27
0.9	7.336	6.990	6.794	6.668	6.516	6.428
	8.186	7.830	7.626	7.495	7.336	7.243
	1.12	1.12	1.12	1.12	1.13	1.13
0.95	16.463	15.730	15.313	15.045	14.719	14.529
	17.392	16.651	16.228	15.954	15.622	15.428
	1.06	1.06	1.06	1.06	1.06	1.06
0.99	89.887	86.083	83.903	82.495	80.785	79.785
	90.881	87.071	84.887	83.474	81.757	80.754
	1.01	1.01	1.01	1.01	1.01	1.01

and a c.v. of 2.58. However, if the congestion level is 0.8, the mean increases to 2.568, but the c.v. decreases to 1.26. Therefore, although lower congestion levels mean smaller mean waiting times, they also produce greater relative variation.

REFERENCES

GREEN, L. 1980. A Queueing System in Which Customers Require a Random Number of Servers. *Opns. Res.* **28,** 1335–1345.

SEILA, A. F. 1980. Stratified Estimation in Regenerative Simulations. Working Paper, Department of Quantitative Business Analysis, University of Georgia, Athens.