THE TRANSIENT BEHAVIOUR OF A
SINGLE SERVER QUEUE WITH BATCH ARRIVALS

P. J. BROCKWELL

(received 25 September 1962)

Summary

A queue at which arrivals occur randomly in batches of fixed size r and for which the service times are independent negative exponential variates is considered. Expressions are obtained for the moments of the transient waiting time distribution and the distribution of the number of customers in the system just before the nth batch arrives. The distribution of the number of customers served in a busy period is also determined.

1. Introduction

Suppose that batches consisting of r customers arrive at times $t_0, t_1, \cdots, t_n, \cdots$ and define

$\eta(t) =$ the number of customers in the system at time t including the one, if any, being served,

$\eta_n = \eta(t_n - 0)$

$P_j^n = \Pr\{\eta_n = j\}$.

Suppose also that $\eta_0 = a \geq 0$.

The interarrival times, $\tau_j = t_j - t_{j-1}$, $j = 1, 2, 3, \cdots$, are independently distributed with distribution function $A(x) = 1 - e^{-\lambda x}$, $x \geq 0$, and the service times are distributed independently of each other and of the interarrival times with distribution function $B(x) = 1 - e^{-\mu x}$, $x \geq 0$.

The process $\eta(t)$ has been studied in the case $r = 1$ by Ledermann and Reuter (1) and Bailey (2). Luchak (3, 4) has found the transient distribution of $\eta(t)$ in the case where r is a random variable with distribution defined by $\Pr\{r = j\} = c_j$, $j = 1, 2, \cdots, m$.

In the present paper we consider, instead of the process $\eta(t)$ in continuous time, the process η_n which is a Markov chain. The resulting expression for the transient probability of emptiness (from which the distribution of η_n can be deduced) can be expressed in terms of binomial probabilities and is much simpler than the corresponding transient probability of emptiness

241

in continuous time. Foster (7) has determined the limiting distribution, $p_j = \lim_{n \to \infty} \Pr\{\eta_n = j\}$, in the more general case where $A(x)$ is arbitrary.

2. The distribution of η_n

Since the service times are negative exponential variates we can consider the server as generating a sequence of potential departure points constituting a Poisson process, a potential departure being a real departure if there is at least one customer in the system when it occurs.

Let $a_j = \Pr\{j \text{ potential departures occur in an interarrival period}\}$

$$= \lambda \int_0^\infty \frac{(\mu t)^j}{j!} e^{-(\lambda+\mu)t} \, dt$$

and $A_j = \sum_{n=j}^\infty a_n$.

Then if $\alpha = \lambda/\lambda + \mu$, $\beta = \mu/\lambda + \mu$, it follows that $a_j = \alpha\beta^j$, $A_j = \beta^j$. The probabilities P_j^n satisfy the equations

$$P_0^n = \sum_{k=0}^\infty P_k^{n-1} A_{k+r},$$

$$P_j^n = \sum_{k=0}^\infty P_k^{n-1} a_{k+r-j}, \qquad j = 1, 2, \cdots, r-1,$$

$$P_j^n = \sum_{k=j-r}^\infty P_k^{n-1} a_{k+r-j}, \qquad j = r, r+1, \cdots.$$

Substituting for a_k, A_k in these equations we find that for $n \geq 1$,

(1)
$$\begin{aligned}
\beta P_1^n &= \alpha P_0^n, \\
\beta P_j^n &= P_{j-1}^n, & j &= 2, 3, \cdots, r, \\
\beta P_j^n &= P_j^n - \alpha P_{j-r-1}^{n-1}, & j &= r+1, r+2, \cdots,
\end{aligned}$$

with the initial condition, $P_j^0 = \delta_{a,j} = \begin{cases} 1 \text{ if } j = a, \\ 0 \text{ otherwise.} \end{cases}$

In order to solve these equations we define the generating functions

$$P_j(w) = \sum_{n=0}^\infty P_j^n w^n, \qquad |w| < 1,$$

$$P(z, w) = \sum_{j=0}^\infty P_j(w) z^j, \qquad |z| \leq 1.$$

Then from equations (1) it is found that

$$\beta P_1(w) - \beta\delta_{a,1} = \alpha P_0(w) - \alpha\delta_{a,0},$$
$$\beta P_j(w) - \beta\delta_{a,j} = P_{j-1}(w) - \delta_{a,j-1}, \qquad\qquad j = 2, 3, \cdots, r,$$
$$\beta P_j(w) - \beta\delta_{a,j} = P_{j-1}(w) - \delta_{a,j-1} - w\alpha P_{j-r-1}(w), \qquad j \geq r+1,$$

and hence

$$(2) \qquad P(z, w) = \frac{\beta P_0(w)(1-z) - \alpha z\delta_{a,0} + z^a(\beta-z)(1-\delta_{a,0})}{\beta - z + \alpha wz^{r+1}}.$$

Once $P_0(w)$ is determined the transient probabilities P_j^n can be found by expanding (2). However, it is simpler to determine the moments of the distribution of η_n as described in § 3. An explicit expression for $P_0(w)$ is given by the following theorem.

THEOREM

$$(3) \qquad \begin{aligned} P_0(w) &= \delta_{a,0} \\ &+ \frac{1}{\beta} \sum_{n=1}^{\infty} w^n\{\psi_{n,a} + a\alpha^{-1}(1-\delta_{a,0})[\beta(\psi_{n,a}-\psi_{n,a-1}) + \psi_{n,a} - \psi_{n,a+1}]\} \end{aligned}$$

where

$$\begin{aligned} \psi_{n,a} &= \sum_{j=0}^{n}\binom{nr+n+a}{j}\alpha^j\beta^{nr+n+a-j} \\ &\quad - (r+1+an^{-1})\alpha\sum_{j=0}^{n-1}\binom{nr+n+a-1}{j}\alpha^j\beta^{nr+n+a-1-j} \\ &= \beta - r\alpha - a\alpha n^{-1} - \sum_{j=n+1}^{nr+n+a}\binom{nr+n+a}{j}\alpha^j\beta^{nr+n+a-j} \\ &\quad + (r+1+an^{-1})\alpha\sum_{j=n}^{nr+n+a-1}\binom{nr+n+a-1}{j}\alpha^j\beta^{nr+n+a-1-j}. \end{aligned}$$

PROOF. By Lagrange's theorem (5) the equation $z = \beta + \alpha wz^{r+1}$, $|w| < 1$, has only one root $z = \zeta$ in the region $|z| < 1$. Since $P(z, w)$ is convergent for $|z| \leq 1$, $|w| < 1$, ζ must also be a zero of the numerator of (2)

$$(4) \qquad\qquad \therefore P_0(w) = \frac{\delta_{a,0}\alpha\zeta + (1-\delta_{a,0})\zeta^a(\zeta-\beta)}{\beta(1-\zeta)} = f(\zeta).$$

From the Lagrange expansion of $P_0(w)$ we have

$$(5) \qquad\qquad P_0(w) = \delta_{a,0} + \sum_{n=1}^{\infty}\frac{\alpha^n w^n}{n}\{D^{n-1}[f'(z)z^{nr+n}]\}_{z=\beta}.$$

Now

$$f'(z) = \frac{1}{\beta(1-z)^2}\{\delta_{a,0}\alpha + (1-\delta_{a,0})[\alpha z^a a(1+\beta)z^a - az^{a+1} - a\beta z^{a-1}]\}$$

and

(6)

$$\frac{1}{(n-1)!}\left\{D^{n-1}\frac{z^{a+nr+n}}{(1-z)^2}\right\}_{z=\beta} = \text{Coefficient of } z^{n-1} \text{ in the expansion of } \frac{(z+\beta)^{nr+n+a}}{(\alpha-z)}$$

$$= n\psi_{n,a}.$$

The theorem is proved by substituting in the series (5) from equation (6).

Note 1. The terms $\psi_{n,a}$ appearing in P_0^n can easily be determined from tables of the binomial distribution (which are readily available for values of $nr + n + a$ up to 150). This reduces the determination of P_0^n to a simple arithmetical operation. For larger values of n the binomial probabilities can be calculated using the method of continued fractions described by Bahadur (6) or by using the recurrence relation

$$\sum_{r=c}^{n+1}\binom{n+1}{r}p^r q^{n+1-r} = \sum_{r=c}^{n}\binom{n}{r}p^r q^{n-r} + p\binom{n}{c-1}p^{c-1}q^{n-c+1}.$$

2. Inspecting the expressions for $\psi_{n,a}$ we see that

$$\lim_{n\to\infty}\psi_{n,a} = 0 \qquad \text{if} \quad n < (nr+n)\alpha,$$

$$\text{i.e.} \quad \text{if} \quad r\alpha > \beta$$

and

$$\lim_{n\to\infty}\psi_{n,a} = \beta - r\alpha \quad \text{if} \quad r\alpha < \beta.$$

The deviation of $\psi_{n,a}$ from its limiting value in the first case involves two lower tail binomial probabilities and in the second case two upper tail binomial probabilities.

It follows from the limiting behaviour of $\psi_{n,a}$ that

$$\lim_{n\to\infty} P_0^n = \max\left(0, 1-\frac{r\alpha}{\beta}\right).$$

3. The moments of the distribution of η_n

From the derivatives of $P(z, w)$ with respect to z at $z = 1$ we can derive expressions for the moments about the origin of the distribution of η_n. Each moment is found in terms of the transient probabilities P_0^n and moments of lower order. For example if $M_k^n = \sum_{j=0}^{\infty} j^k P_j^n$, $P_0 = 1 - (r\alpha)/\beta$ and $M_1 = r\alpha(1 + r)/2(\beta - \alpha r)$ then

(7)

$$M_1^n = a + \frac{\beta}{\alpha}\sum_{j=1}^{n}(P_0^j - P_0)$$

$$M_2^n = a^2 + (2r+1)(a - M_1^n) + \frac{2}{\alpha}(\beta - r\alpha)\sum_{j=1}^{n}(M_1 - M_1^j)$$

and higher moments can be obtained in the same way. If the limiting distribution of η_n exists, i.e. if $r\alpha < \beta$, then P_0 and M_1 are the limits of P_0^n and M_1^n respectively.

To determine M_1^n and M_2^n, P_0^n is first found from (3) and then equations (7) are used. The summations in (7) can be carried out either directly or by approximating the discreet sets of values by continuous curves and integrating numerically.

4. The distribution of waiting time

If w_n is the time which elapses between the arrival of the $(n+1)$th batch and its commencement of service and $f_n(x)$ is the probability density function of w_n, defined for $x > 0$, we have

$$(8) \qquad f_n(x) = \sum_{k=1}^{\infty} P_k^n \frac{\mu^k x^{k-1}}{(k-1)!} e^{-\mu x}, \qquad x > 0$$

and $\Pr\{w_n = 0\} = P_0^n$.

If W_q^n is the qth moment about the origin of the distribution of w_n then

$$(9) \qquad W_q^n = E\{w_n^q\} = \frac{1}{\mu^q} \sum_{k=1}^{\infty} \frac{(k+q-1)!}{(k-1)!} P_k^n$$

and in particular

$$
W_1^n = \frac{1}{\mu} M_1^n
$$

(10)

$$
W_2^n = \frac{1}{\mu^2} (M_2^n + M_1^n).
$$

5. The distribution of the number of batches served in a busy period

Let q_j be the probability that exactly j batches are served in a busy period and let p_{00}^n be the value of P_0^n given that $\eta_0 = 0$.

Then if

$$
q(w) = \sum_{j=1}^{\infty} q_j w^j, \qquad |w| \leqq 1,
$$

$$
p_{00}(w) = \sum_{j=0}^{\infty} p_{00}^j w^j, \qquad |w| < 1,
$$

it follows from the relation

$$
p_{00}^n = q_1 p_{00}^{n-1} + q_2 p_{00}^{n-2} + \cdots + q_n p_{00}^0, \qquad n \geqq 1,
$$

that

(11) $$q(w) = 1 - p_{00}(w)^{-1}, \qquad |w| < 1.$$

Hence from equation (4)

$$q(w) = \frac{\zeta - \beta}{\alpha \zeta} = g(\zeta)$$

where ζ is the only root of the equation $z = \beta + \alpha w z^{r+1}$, $|w| < 1$, inside the region $|z| < 1$.

Expanding $q(w)$ by Lagrange's theorem we find that

$$q(w) = \sum_{n=1}^{\infty} w^n \frac{\alpha^{n-1} \beta^{nr}}{n} \binom{nr + n - 2}{n - 1}$$

and hence

(12) $$q_n = \frac{\beta}{n} \binom{nr + n - 2}{n - 1} \alpha^{n-1} \beta^{nr-1}.$$

q_n can be evaluated from tables of the binomial distribution or, for large values of n, using Stirling's approximation to evaluate the binomial coefficient.

6. The relation between the batch arrivals model and $M/E_r/1$

Consider the system $M/E_r/1$ with $k + 1$ customers in the system at time $t_0 - 0$, the elapsed service time of the customer in service at time t_0 being equal to u. For this system let $\pi_j^n = \Pr\{$the $(n + 1)$th arrival finds j customers in the system$\}$.

Now for the batch arrivals model suppose that $\eta_0 = kr + \theta$ where θ is a random variable such that

$$\Pr\{\theta = j\} = \begin{cases} v_j, & j = 1, 2, \cdots, r, \\ 0 & \text{otherwise} \end{cases}$$

where

$$v_j = \frac{(\lambda u)^{r-j}/(r-j)!}{\sum_{n=0}^{r-1} (\lambda u)^n/n!}.$$

Then if $Q_s^n = \sum_{j=1}^r v_j \Pr\{\eta_n = s | \eta_0 = kr + j\}$, the probabilities π_s^n can be written

(13)
$$\pi_0^n = Q_0^n$$
$$\pi_j^n = \sum_{s=jr-r+1}^{jr} Q_s^n, \qquad j = 1, 2, \cdots$$

The moments about the origin of the waiting time distribution of the $(n + 1)$th arrival at $M/E_r/1$ are given by
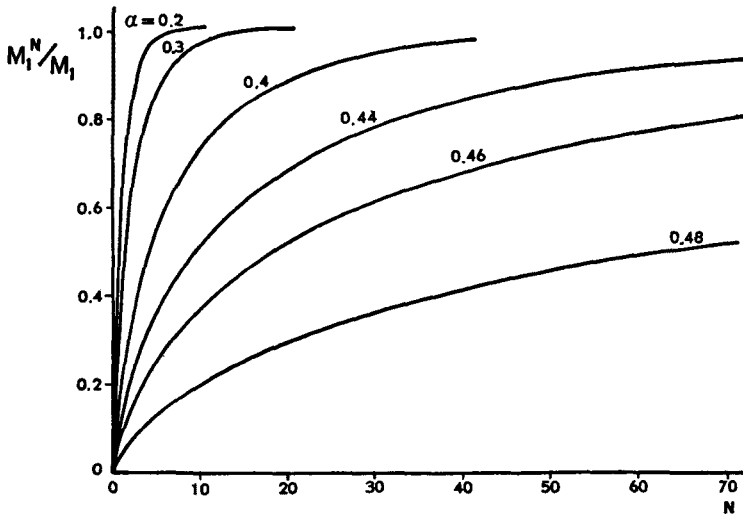
Figure 1

Mean number of customers at the $(n+1)$th arrival for $M/M/1$ starting from emptiness (expressed as a fraction of the asymptotic value).
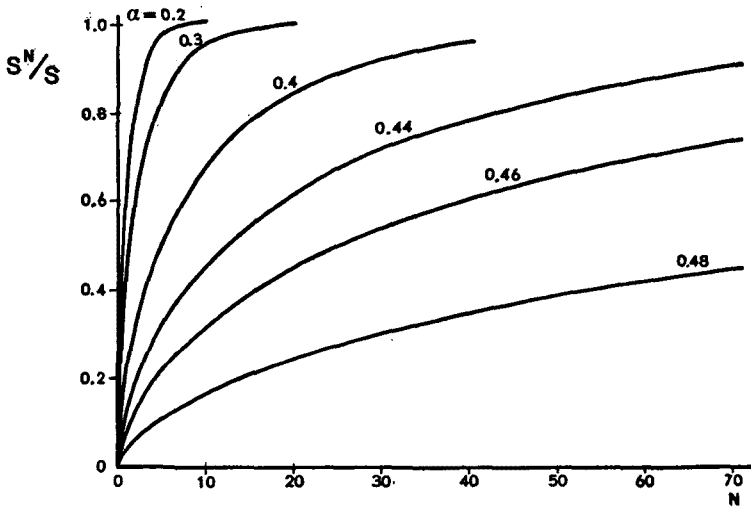

Figure 2

Standard deviation of queue size expressed as a fraction of the asymptotic value.

$$(14) \qquad \overline{W}_q^n = \sum_{j=1}^{r} v_j W_{q,j}^n$$

where $W_{q,j}^n$ is the value of W_q^n obtained from equation (9) when $\eta_0 = kr + j$.

The distribution of the number of customers served in a busy period for $M/E_r/1$ is given by equation (12).

7. Numerical calculations

The probabilities P_0^n, $n = 1, 2, 3, \cdots$, can be found as indicated in § 2. Once P_0^j is known for $j \leqq n$ the transient moments M_1^n and M_2^n can be determined from equations (7). This method was used to derive the transient mean and standard deviation of queue size at an arrival in the particular case $r = 1$, $\eta_0 = 0$. The results are shown in figures 1 and 2 where the discrete sets of points have been plotted as continuous curves. Calculations were carried out only for traffic intensity $\alpha/\beta < 1$ but it is to be noted that the transient moments of the queue size distribution can be obtained in the same way even if $\alpha \geqq \beta$.

I am indebted to Mr. P. D. Finch for his valuable suggestions during the course of this work and to the referee for a number of helpful comments.

References

[1] Ledermann, W. and Reuter, G. E. H., Spectral Theory for the Differential Equations of Simple Birth and Death Processes, *Phil. Trans. Roy. Soc. London*, (A), **246**, 326—369 (1954).
[2] Bailey, N. T. J., A Continuous Time Treatment of a Simple Queue using Generating Functions, *J.R.S.S.* (B), **16**, 288—291 (1954).
[3] Luchak, G., The Solution of the Single Channel Queueing Equation Characterized by a Time-dependent Poisson-distributed Arrival Rate and a General Class of Holding Times, *Operations Research*, **4**, 711—732 (1956).
[4] Luchak, G., The Continuous Time Solution of the Equations of the Single Channel Queue with a General Class of Service-time Distributions by the Method of Generating Functions, *J.R.S.S.* (B), **20**, 176—181 (1958).
[5] Whittaker, E. T. and Watson, G. N., *A Course of Modern Analysis* (C.U.P.), 4th ed., 1952.
[6] Bahadur, R. R., Some Approximations to the Binomial Distribution Function *A.M.S.*, **31**(1), 43—54 (1960).
[7] Foster, F. G., Queues with Batch Arrivals I, *Acta Math. Hungarica*, **12**, 1—10 (1961).

University of Melbourne.