NOTES

WAITING TIMES WHEN QUEUES ARE IN TANDEM

By Edgar Reich

University of Minnesota

1. We study the distribution of waiting times when customers proceed to a second (multiple-counter) queue after having been processed at a first (multiple-counter) queue[1]. For reasons of expediency we restrict ourselves to the case of unsaturated queues in "equilibrium," that is, to stationary statistics. The main results are for the case of exponential service time, where it turns out that, contrary to a-priori intuition, the situation is surprisingly simple. As shown by Theorem 6, no such simple behavior can be expected when the service time distributions are even only slightly more general. Theorem 4 was first found essentially by P. J. Burke [1], by a different method.[2]

The concept of reversibility of a Markov chain, certain aspects of which are discussed in Sec. 2, has turned out to be fruitful in connection with the analysis, and is of some independent interest.

2. A stationary stochastic process $N(t)$ is said to be reversible if $N(t)$ and $N(-t)$ have the same multivariate distributions. If $N(t)$ is a discrete or continuous parameter Markov chain with a denumerable state space, say, $0, 1, 2, \cdots$, then $N(-t)$ is a process of the same type. The necessary and sufficient condition for reversibility becomes

$$\theta_{ij}(t) = p_i P_{ij}(t) = p_j P_{ji}(t) = \theta_{ji}(t), \qquad i, j = 0, 1, 2, \cdots, \tag{1}$$

where p_i and $P_{ij}(t)$ are respectively, the stationary, and transition probabilities of $N(t)$.

Kolmogorov's criterion for reversibility of Markov chains with a finite state space ([8]; [5], p. 66) may, in a special case, be immediately generalized to the denumerable state-space case, as follows.

Theorem 1. *Let $N(k)$, $k = 0, \pm 1, \pm 2, \cdots$, be an irreducible stationary discrete-parameter Markov chain with the state space $0, 1, 2, \cdots$, the stationary probabilities u_k, and the singlestep transition probabitities π_{ij}. A necessary and sufficient condition for the reversibility of $N(k)$ is that*

$$\pi_{i_1 i_2}\pi_{i_2 i_3} \cdots \pi_{i_{n-1} i_n}\pi_{i_n i_1} = \pi_{i_1 i_n}\pi_{i_n i_{n-1}} \cdots \pi_{i_3 i_2}\pi_{i_2 i_1} \tag{2}$$

for every sequence of non-negative integers $(i_1, i_2, \cdots, i_n, i_1)$ beginning and ending with the same integer.

Received July 15, 1956; revised November 27, 1956.

[1] A part of this paper represents work done at The RAND Corporation.

[2] A special case of a part of this theorem was also treated (unpublished) by H. H. Goode and R. E. Machol. Their work is to appear in a text.

Proof. According to (1), for $t = 1$,

$$\theta_{i_1 i_2}\theta_{i_2 i_3} \cdots \theta_{i_{n-1} i_n}\theta_{i_n i_1} = \theta_{i_1 i_n}\theta_{i_n i_{n-1}} \cdots \theta_{i_3 i_2}\theta_{i_2 i_1} \tag{3}$$

is necessary. Since $u_i > 0$, we may cancel the u_i, obtaining (2). Summing both sides of (2) over $i_3, i_4, \cdots, i_n$, we find

$$\pi_{i_1 i_2}\pi_{i_2 i_1}(n-1) = \pi_{i_1 i_2}(n-1)\pi_{i_2 i_1}.$$

If τ is the period of the chain, then the lim $\sup_{n\to\infty}$ of the left and right sides are $\tau u_{i_1}\pi_{i_1 i_2}$ and $\tau u_{i_2}\pi_{i_2 i_1}$, respectively ([4], p. 331). Hence $u_{i_1}\pi_{i_1 i_2} = u_{i_2}\pi_{i_2 i_1}$. Eq. (1) follows by induction.

Let us call a finite sequence of non-negative integers, beginning and ending in the same integer a *cycle* if no proper portion begins and ends in the same integer. Evidently (2) need hold only for cycles.

Consider a continuous-parameter, time-homogeneous Markov chain $N(t)$, with $P_{ij}(h) = \delta_{ij} + hL_{ij} + o(h)$,

$$\sum_{j=0}^{\infty} L_{ij} = 0, L_{ii} < 0, 1, 2, \cdots.$$

A process of this type will be said to be of type A if, in addition,

(i) $N(t)$ has stationary probabilities p_i, $\sum_i p_i = 1$, $\sum_i p_i L_{ij} = 0$;

(ii) $N(t)$ has at most a finite number of discontinuities in every finite interval; $-\sum_{i=0}^{\infty} p_i L_{ii} < \infty$.

(iii) the associated discrete-parameter chain N^* defined by

$$\pi_{ij} = (\delta_{ij} - 1)L_{ij} / L_{ii}, \qquad i, j = 0, 1, 2, \cdots,$$

is irreducible. (Note: This is a chain, of period $\geqq 2$, resulting from a shift of the instants at which $N(t)$ changes state to the instants $t = \cdots -1, 0, 1, \cdots$)

THEOREM 2. *A necessary and sufficient condition for a continuous-parameter Markov process of type A to be reversible is that*

$$L_{i_1 i_2}L_{i_2 i_3} \cdots L_{i_{n-1} i_n}L_{i_n i_1} = L_{i_1 i_n}L_{i_n i_{n-1}} \cdots L_{i_3 i_2}L_{i_2 i_1} \tag{4}$$

for every cycle.

Proof. Since the matrix $P_{ij}(t)$ of a process of type A is uniquely determined by its values for infinitesimal t, condition (1) is equivalent to

$$p_i L_{ij'} = p_j L_{ji}. \tag{5}$$

Note that

$$r = \left(-\sum_{i=0}^{\infty} p_i L_{ii}\right)^{-1} > 0,$$

in view of the second part of assumption (ii), above.

$$u_i = -rp_i L_{ii}, \qquad i = 0, 1, \cdots,$$

can be assigned as stationary probabilities to N^*. Then a necessary and sufficient condition for N^* to be reversible is

$$u_i\pi_{ij} = -rp_i(\delta_{ij} - 1)L_{ij} = -rp_j(\delta_{ji} - 1)L_{ji},$$

which is equivalent to (5). The theorem follows from the fact that (4) is equivalent to (2).

Let $B(t)$ denote a stationary birth-death process with state space $0, 1, 2, \cdots$, the stationary probabilities p_i, and

$$\begin{aligned} P_{i,i+1}(h) &= \lambda_i h + o(h), \quad \lambda_i > 0, \quad i = 0, 1, \cdots, \\ P_{i,i-1}(h) &= \mu_i h + o(h), \quad \mu_i > 0, \quad i = 1, 2, \cdots, \mu_0 = 0, \\ P_{ii}(h) &= 1 - \lambda_i h - \mu_i h + o(h), \quad i = 0, 1, \cdots. \end{aligned}$$

$B(t)$ is permitted to have at most a finite number of increases (births), and decreases (deaths) in any finite time interval, $\sum p_i(\lambda_i + \mu_i) < \infty$.

THEOREM 3. *$B(t)$ is reversible.*

Proof. If (4) is to be other than of the form $0 = 0$, the cycle must be of length 3, in which case (4) still holds trivially[3].

COROLLARY. *If $\lambda_n = \lambda$, $(n = 0, 1, \cdots)$ then the death times of $B(t)$ form a Poisson process of density λ.*

Outline of proof. Since λ_n is constant the birth times are Poisson with density λ. The stochastic process $B_1(t) = B(-t)$ is statistically identical with the process $B(t)$. But if $B(t)$ is a fixed realization, and $B_1(t) = B(-t)$ then the births of $B(t)$ become the deaths of $B_1(t)$.

3. Consider an unsaturated queue of type $M/M/s$ (Poisson input, s counters, exponential service time, first come, first served), in equilibrium. If $n(t)$ is the sum of the number of customers on queue, plus those being served, then $n(t)$ is a process of type $B(t)$, in which customers' arrivals correspond to births, and departures to deaths. By considering the reversibility of $n(t)$, guaranteed by the corollary to Theorem 3, the following is now clear:

THEOREM 4. (a) *The sequence of departure times form a Poisson process.* (b) *The value of $n(t)$ is independent of all past departure times.* (c) *If t_0 is a departure time, then $n(t_0 + 0)$ is independent of all past departure times.*

Note. The above results are, of course, true for more general queue disciplines. The number of servers, instead of being fixed, may be permitted to vary as a specified function of the number of customers present. Also, instead of "first come, first served," e.g., random service, or "last come, first served," will do without effect on the results.

Suppose the customers, after departing from a first queue of type $M/M/s$,

[3] Heuristic forms of the necessary argument date to P. and T. Ehrenfest [3]. The above proof is in the spirit of the Ehrenfests' reasoning. Simple algebraic verifications are also possible, but they leave the situation less lucid. The condition $\Sigma p_i(\lambda_i + \mu_i) < \infty$ is actually superfluous.

enter a second multiple-counter queue, where they are served first come, first served, with exponential service time. Such a combination of two tandem queues will be referred to as a σ-system. It follows, from Theorem 4b, that if $n_1(t)$, $n_2(t)$ refer, respectively, to the first and second queues of a σ-system, then $n_1(t)$ and $n_2(\tau)$ are independent, $\tau \leqq t$. This was first proved in the special case $s = 1$, $t = \tau$, by Jackson [6].

In what follows, the term *waiting time* will be used to refer to the time elapsed between a customer's arrival and departure, the service time included. Let T_1 and T_2 denote a customer's waiting time at the first and second queues of a σ-system, respectively.

THEOREM 5. *If $s = 1$, then T_1 and T_2 are independent.*

Proof. Let n_1 be the number of customers at the first queue the instant after a customer C departs, and let n_2 be the number of customers C finds at the second queue (customers being served included). As a corollary of Theorem 4c, n_1 and n_2 are independent. Let

$$A(t; k) = \Pr\{T_1 < t \mid n_2 = k\}.$$

If λ is the number of customers arriving per unit time, then n_1 is the number of Poisson events of density λ that occurred during the waiting period T_1. We have

$$\Pr\{n_1 = j \mid T_1 = t, n_2 = k\} = e^{-\lambda t}(\lambda t)^j/j!.$$

Therefore

$$E\{z^{n_1} \mid n_2 = k\} = \int_0^\infty e^{\lambda t z} e^{-\lambda t}\, dA(t; k).$$

Now the left side is independent of k. Therefore $A(t; k)$ does not depend on k. Hence n_2, and consequently also T_2, are independent of T_1.

4. We will now consider the queues of type $E_j / E_k / s$ (interarrival and service periods normalized chi-square with $2j$ and $2k$ degrees of freedom, respectively [7]), and show that Theorem 4a cannot be generalized further, in a certain direction. Note that both when $j = k = 1$, and $j = k = \infty$, the departure epochs of an $E_j / E_k / s$ queue are again E_k. (The case $j = k = \infty$ corresponds to a periodic input with constant service time.) One may therefore ask[4] whether this state of affairs holds whenever $j = k$. However, Theorem 6, below, shows this to be false.

THEOREM 6. *The departure epochs of an $E_2 / E_2 / 1$ process are not an E_2 process.*

Proof. For the case under consideration we have x = interarrival period = $x_1 + x_2$, where x_i, $i = 1, 2$, are independent, with

$$E\{e^{-sx_i}\} = \frac{\lambda}{\lambda + s}, \qquad \lambda > 0, \qquad i = 1, 2.$$

[4] This question is related to the asymptotic behavior of a large number of queues in tandem, each with E_k-type service time, k fixed.

Similarly, the service periods, y, are of the form $y = y_1 + y_2$, where y_i, $i = 1, 2$, are independent, and

$$E\{e^{-sy_i}\} = \frac{\mu}{\mu + s}, \qquad 0 < \rho = \lambda/\mu < 1, \qquad i = 1, 2.$$

At a given instant the entrance will be said to be in state 1 (2) if the system is in the portion $x_1(x_2)$ of an interarrival period. Consider instants just following a departure. Let $A_{i0} = \Pr\{$extrance is in state i, and there are 0 customers left behind$\}$, $i = 1, 2$. Let τ be the length of an interdeparture period. If the departure epochs formed an E_2 process it would follow that τ had the same marginal distribution as x, that is,

$$E\{e^{-s\tau}\} = A_{10}\left(\frac{\lambda}{\lambda + s}\right)^2 \left(\frac{\mu}{\mu + s}\right)^2 + A_{20}\left(\frac{\lambda}{\lambda + s}\right)\left(\frac{\mu}{\mu + s}\right)^2 + (1 - A_{10} - A_{20})\left(\frac{\mu}{\mu + s}\right)^2 = \left(\frac{\lambda}{\lambda + s}\right)^2.$$

Multiplying both sides by $(\lambda + s)^2(\mu + s)^2$, and equating coefficients of s^2, we have

$$P_0 = \Pr\{0 \text{ customers are left behind by a departing customer}\}$$
$$= A_{10} + A_{20} = 1 - \rho^2.$$

However this is incorrect, as it differs from Volberg's [9] formula for P_0. Thus we have a contradiction.

A related question is that of the possibility of *imbedding* $n(t)$ in a reversible Markov process, e.g., for $s = 1$. To this end we define the "pseudostate" $\tilde{n}(t)$ of an $E_j / E_k / 1$ queue. We shall say that $a(t) = r$, $r = 0, 1, 2, \cdots, j - 1$, if the $(r + 1)$st stage of the interarrival period is in progress. Similarly, put $b(t) = r$, $r = 0, 1, \cdots, k - 1$, if the $(r + 1)$st stage of the service period is in progress; if the counter is empty, $b(t) = 0$. Define

$$\tilde{n}(t) = n(t) + \frac{a(t)}{j} - \frac{b(t)}{k}. \tag{6}$$

The realizations of the process $\tilde{n}(t)$ are constant except for jumps of height $1/j$, upward, and jumps of height $1/k$, downward. We make the following observation.

THEOREM 7[5]. *If j and k are relatively prime, $\tilde{n}(t)$ is a Markov process.*

Proof. If the hypothesis is satisfied, $n(t_0)$, $a(t_0)$, $b(t_0)$ can be recovered from a knowledge of $\tilde{n}(t_0)$.

We conclude that if j and k are relatively prime, and $j = k$, then $\tilde{n}(t)$ is reversible. Since $j = k = 1$ is the only admissible possibility, the special nature of the $E_1 / E_1 / 1$ queue is seen in a new light.

A straightforward computation shows that the following partial *converse* of Theorem 4a holds.

[5] This fact enables one to study the transient behavior of $n(t)$ for $E_j/E_k/1$, j, k relatively prime. We shall not explore this further at this time, however.

THEOREM 8. *If the arrival and departure epochs of a single-counter queue are both Poisson, then the service time distribution is exponential, or a step function at* 0.

The author has had valuable discussions with A. W. Marshall and T. E. Harris in connection with this work.

REFERENCES

[1] P. J. BURKE, "The Output of a Queuing system," *Operations Research*, Vol. 4 (1956), pp. 699–704.
[2] J. L. DOOB, *Stochastic Processes*, John Wiley and Sons, New York, 1953.
[3] P. EHRENFEST AND T. EHRENFEST, "Über zwei bekannte Einwände gegen das Boltzmannsche H-Theorem," *Physikalische Zeitschrift*, Vol. 8 (1907), pp. 311–314.
[4] W. FELLER, *An Introduction to Probability Theory and its Applications*, Vol. 1, John Wiley and Sons, New York, 1950.
[5] M. FRÉCHET, *Méthode des fonctions arbitraires. Théorie des événements en chaîne dans le cas d'un nombre fini d'etats possibles*, Gauthier-Villars, Paris, 1938.
[6] R. R. P. JACKSON, "Queueing systems with phase type service," *Operational Research Quarterly*, Vol. 5 (1954), pp. 109–120.
[7] D. G. KENDALL, "Stochastic processes occurring in the theory of queues and their analysis by the method of the imbedded Markov chain," *Ann. Math. Stat.*, Vol. 24 (1953), pp. 338–354.
[8] A. KOLMOGOROFF, "Zur Theorie der Markoffschen Ketten," *Math. Ann.*, Vol. 112 (1936), pp. 155–160.
[9] O. A. VOLBERG, "Problème de la queue stationnaire et nonstationnaire," *Doklady Akad. Nauk SSSR*, N.S., Vol. 24 (1939), pp. 657–661.
[10] R. R. P. JACKSON, "Random queueing processes with phase-type service," *J. Roy. Stat. Soc. B*, Vol. 18 (1956), pp. 129–132.

ON THE POWER OF OPTIMUM TOLERANCE REGIONS WHEN SAMPLING FROM NORMAL DISTRIBUTIONS[1]

BY IRWIN GUTTMAN

University of Alberta

1. Introduction and Summary. In [1], optimum β-expectation tolerance regions were found by reducing the problem to that of solving an equivalent hypothesis testing problem. The regions produced when sampling from a k-variate normal distribution were found to be of similar β-expectation and optimum in the sense of minimax and most stringency. It is the purpose of this paper to discuss the "Power" or "Merit" of such regions, when sampling from the k-variate normal distribution.

Let $X = (X_1, \cdots, X_n)$ be a random sample point in n dimensions, where each X_i is an independent observation, distributed by $N(\mu, \sigma^2)$. It is often desirable to estimate on the basis of such a sample point a region which contains a given fraction β of the parent distribution. We usually seek to estimate the center $100\beta\%$ of the parent distribution and/or the $100\beta\%$ left-hand tail of the parent distribution.

Received April 20, 1956; revised December 17, 1956.
[1] Research supported by the University of Alberta General Research Fund.