

F. A. HAIGHT (Los Angeles)

THE DISCRETE BUSY PERIOD DISTRIBUTION FOR VARIOUS SINGLE SERVER QUEUES

I. Introduction and notation. For a single server queue in statistical equilibrium, consider the following probability distributions:

$p(n)$, $n = 0, 1, 2, \dots$ — the distribution of queue length, including the person being served;

$q(n, t)$, $n = 0, 1, 2, \dots$ — the probability of n persons joining the queue in a time period of length t , beginning at an arbitrary instant;

$h(x)$, $0 < x < \infty$ — the density function for the service time;

$w(x)$, $x = 0$, $0 < x < \infty$ — the (mixed discrete and continuous) distribution of waiting time;

$\pi(n, r)$, $n = r, r+1, r+2, \dots$ — the discrete busy period distribution, i. e. the probability that, beginning with r in the queue, exactly n will pass through before the queue first becomes empty.

To distinguish several types of input and service, we will use the Kendall symbol (Kendall, [12]), omitting the third element, with the following modifications: (i) In the Erlang case, it is customary to give one parameter, k , as a subscript on E . We will give both parameters, writing for example $E_{k,\lambda}$ and extend the notation to the Poisson case, writing for example M_λ , (ii) We shall consider processes more general than the renewal process, G , namely *stationary point processes*, and denote these by the letter S .

A subscript on the function h will, for S cases, have the analogous meaning to the starred superscript in the G case, namely the density function for the total length of several continuous intervals:

In a queue with S service termination points, density for length of n service periods is $h_n(x)$.

In a queue with G service termination points, density for length of n service periods is $h^{n*}(x)$.

The Laplace transformations of q and h will be denoted respectively by \bar{q} and \bar{h} .

II. The main theorems.

THEOREM 1. *For an M_λ/D_μ queue,*

$$(1) \quad \pi(n, r) = \frac{r}{n} \int_0^\infty q(n-r, t) h^{n*}(t) dt.$$

This theorem was proved for $r = 1$ by Borel ([3]), in general by Tanner ([20]). A purely combinatorial proof was given by Tanner ([21]) and Mott ([16]); this proof is quite independent of the M_λ/D_μ assumption, and in fact requires no queue-theoretic interpretation whatsoever. It merely shows that the proportion of ways $n-r$ persons can pass through a queue beginning with r in queue so that the queue does not become empty before the $(n-r)$ -th person is r/n . Exactly the same proof can be used for the following result:

THEOREM 2. *For an M_λ/G queue, equation (1) holds.*

In fact, if we replace the convolution notation by the corresponding subscript notation, we can apply Tanner's proof to the following still more general results:

THEOREM 3. *For an M_λ/S queue*

$$(2) \quad \pi(n, r) = \frac{r}{n} \int_0^\infty q(n-r, t) h_n(t) dt.$$

THEOREM 4. *For an S/M_μ queue, equation (1) holds.*

If the distribution denoted by the letter q is to have the property "beginning at an arbitrary instant", as in our definition, we must keep one of the symbols M . But if we were prepared to substitute the expression "beginning at the start of a service period", we could even allow Tanner's proof to establish equation (1) in the S/S case. Such an extension will involve much more complicated expressions for $q(n, t)$, and will not shed much light on the busy period question. For this reason we consider various special cases of M_λ/S and S/M_μ queues, most particularly the latter.

Note that none of the parameters defining the input and service processes are shown in equations (1) and (2), although of course they must enter the picture once special functions are assumed.

III. M_λ/S queues. The simplest cases have Poisson input; many of these are already well known. Let

$$(3) \quad q(n, t) = e^{-\lambda t} (\lambda t)^n / n!.$$

EXAMPLE 1. M_λ/D_μ , $\lambda/\mu = \rho$. Then

$$h(x) = \delta\left(x - \frac{1}{\mu}\right), \quad h^{n*}(x) = \delta\left(x - \frac{n}{\mu}\right),$$

leading to

$$(4) \quad \pi(n, r) = \frac{r}{n} \cdot \frac{e^{-\varrho n} \varrho^{n-r} n^{n-r-1}}{(n-r)!},$$

the Borel-Tanner distribution; see also Haight and Breuer ([10]).

EXAMPLE 2. M_λ/M_μ , $\lambda/\mu = \varrho$. Then

$$(5) \quad h(x) = \mu e^{-\mu x}, \quad h^{n*}(x) = \frac{e^{-\mu x} \mu^n x^{n-1}}{(n-1)!},$$

leading to

$$(6) \quad \pi(n, r) = \frac{r}{n} \binom{2n-r-1}{n-1} \frac{\varrho^{n-r}}{(1+\varrho)^{2n-r}}.$$

Equation (6) is a distribution discussed by Haight ([8], [9]), and it is a special case of a distribution of Narayana ([17]).

EXAMPLE 3. $M_\lambda/E_{k,\mu}$, $\lambda/\mu = \varrho$. Then

$$h(x) = \frac{e^{-\mu x} \mu^k x^{k-1}}{(k-1)!}, \quad h^{n*}(x) = \frac{e^{-\mu x} \mu^{nk} x^{nk-1}}{(nk-1)!},$$

leading to

$$(7) \quad \pi(n, r) = \frac{r}{n} \binom{nk+n-r-1}{nk-1} \frac{\varrho^{n-r}}{(1+\varrho)^{nk+n-r}},$$

which includes examples 1 and 2.

The cases M_λ/G have been dealt with by Cox and Smith ([4]), and Takacs ([18]); the latter also gives a complete bibliography of the busy period question. In the M_λ/S case, if we assume furthermore that $h(x)$ is independent of λ , we can write

$$(8) \quad \pi(n, r) = \frac{r}{n} \cdot \frac{(-\lambda)^{n-r}}{(n-r)!} \cdot \frac{d^{n-r}}{d\lambda^{n-r}} \bar{h}_n(\lambda)$$

which reduces, for the M_λ/G case, to

$$(9) \quad \pi(n, r) = \frac{r}{n} \cdot \frac{(-\lambda)^{n-r}}{(n-r)!} \cdot \frac{d^{n-r}}{d\lambda^{n-r}} [\bar{h}(\lambda)]^n.$$

Equations (4), (6) and (7) respectively can be easily obtained from equation (9) by setting

$$\bar{h}(s) = \begin{cases} e^{-s/\mu} & \text{for equation (4),} \\ \mu(\mu+s)^{-1} & \text{for equation (6),} \\ \mu^k(\mu+s)^{-k} & \text{for equation (7).} \end{cases}$$

In general, different assumptions regarding the service time distribution are troublesome only insofar as their Laplace transformations are troublesome. Since these functions have been extensively tabulated, there is little interest in pursuing special service time distributions.

IV. S/M_μ queue, arrival rate dependent upon μ . In the present section we will assume that equations (5) hold. Then equation (1) is the derivative of a Laplace transformation (with variable μ) only if the counting distribution for joiners, $q(n, t)$ is independent of μ . We will not make that assumption for the present.

EXAMPLE 4. Queue with feedback. Takacs ([19]) discusses a queue in which a proportion P of departing customers rejoin the queue and a proportion $Q = 1 - P$ go away. With Poisson new arrivals, the queue is of type $M_{\lambda+P\mu}/M_\mu$ and the equation (1) yields

$$(10) \quad \pi(n, r) = \frac{r}{n} \binom{2n-r-1}{n-1} \frac{(P + \varrho)^{n-r}}{(Q + \varrho)^{2n-r}}.$$

EXAMPLE 5. Queue with balking. In a model proposed by Haight ([6], [7]) and studied by Finch ([5]) and Ancker and Gafarian ([1], [2]) $q(n, t)$ depends on μ in such an intricate and complicated way as to be "impossible" to write explicitly. In this case, it is desirable to find some way to relate the busy period probabilities $\pi(n, r)$ directly to the queue length probabilities $p(n)$ without needing to find first the joining probabilities $q(n, t)$.

If we consider a departing customer, who has spent time t_1 waiting and time t_2 being served, we know that the probability of a queue of n at this time is just $q(n, t_1 + t_2)$. Since the distribution of t_1 is $w(x)$ and the distribution of t_2 is $h(x)$, we can write

$$(11) \quad p(n) = \int_0^\infty q(n, x) [w(x) * h(x)] dx.$$

On the other hand, an arriving customer, if confronted with a queue of $n-1$ waiting and one being served, will wait n negative exponentially distributed service periods for service. Therefore

$$(12) \quad w(x) = p(0) \delta(x) + \sum_{j=1}^{\infty} p(j) h^{j*}(x).$$

Using the convention $h^{0*}(x) = \delta(x)$, we can write equation (12) in the form

$$(13) \quad w(x) = \sum_{j=0}^{\infty} p(j) h^{j*}(x).$$

Substituting equation (13) into equation (11), we find that the probability vector $p(n)$ is unchanged by an infinite matrix (β_{ij}) :

$$(14) \quad p(n) = \sum_{j=0}^{\infty} \beta_{nj} p(j), \quad n = 0, 1, 2, \dots$$

where

$$(15) \quad \beta_{ij} = \int_0^{\infty} q(i, x) h^{(j+1)*}(x) dx.$$

Now, the coefficients β_{ij} are connected with the discrete busy period distribution in a very simple way:

$$(16) \quad \pi(n, r) = \frac{r}{n} \beta_{n-r, n-1}, \quad r = 1, 2, \dots, \quad n = r, r+1, \dots$$

or, written the other way around,

$$(17) \quad \beta_{ij} = \frac{j+1}{j-i+1} \pi(j+1, j-i+1), \quad i, j = 0, 1, 2, \dots$$

Because of their role in connecting the queue length distribution with the discrete busy period distribution, we will make a special study of the coefficients β_{ij} . We first note that the column sums are all unity:

$$\sum_{i=0}^{\infty} \beta_{ij} = \int_0^{\infty} h^{(j+1)*}(x) \sum_{i=0}^{\infty} q(i, x) dx = \int_0^{\infty} h^{(j+1)*}(x) dx = 1.$$

The row sums, on the other hand, are not unity:

$$(18) \quad \sum_{j=0}^{\infty} \beta_{ij} = \sum_{j=0}^{\infty} \int_0^{\infty} q(i, x) \frac{\mu^{j+1} x^j}{j!} e^{-\mu x} dx = \mu \int_0^{\infty} q(i, x) dx.$$

McFadden ([13]) has shown that for any stationary point process (in this case the joining process) the integral of the counting distribution with respect to the counting period is the covariance of gaps separated by $n-1$ gaps, multiplied by the mean count λ , that is

$$(19) \quad \sum_{j=0}^{\infty} \beta_{ij} = \lambda \mu E(X_0 X_i),$$

where X_0 is the length of an arbitrarily selected interarrival gap and X_i is the length of a gap separated from it by $i-1$ other gaps. For a G/M_μ

queue, the covariance is the product of the means, and since each mean is $1/\lambda$, the row sums reduce to

$$(20) \quad \sum_{j=0}^{\infty} \beta_{ij} = \frac{1}{\varrho}.$$

The diagonal sums of the matrix (β_{ij}) are related to mean values of the busy periods. If

$$M(r) = \sum_{n=r}^{\infty} n\pi(n, r)$$

then, using equation (16) we obtain

$$(21) \quad M(r) = \sum_{n=0}^{\infty} r\beta_{n, n+r-1}$$

so that $M(r)/r$ is the sum of the r -th superdiagonal of the matrix.

EXAMPLE 5. (Continued) In an M_λ/M_μ queue with balking,

$$(22) \quad p(n) = c_n \varrho^n p(0)$$

where $\varrho = \lambda/\mu$ is calculated for all arrivals, whether they join or not, and is therefore not a true traffic intensity, and where c_n are some constants characterizing the mentality of the customers and hence independent of the behavior of the queue. Substituting these values into equation (14), we obtain

$$(23) \quad c_n \varrho^n = \sum_{j=0}^{\infty} \beta_{nj} c_j \varrho^j.$$

If we compare equations (23) with the analogous set for an M_λ/M_μ queue without balking, namely

$$(24) \quad \varrho^n = \sum_{j=0}^{\infty} \binom{n+j}{j} \frac{\varrho^{n+j}}{(1+\varrho)^{n+j+1}}$$

we see that one set of coefficients β_{ij} satisfying equations (23) is certainly

$$(25) \quad \beta_{ij} = \frac{c_i}{c_j} \binom{i+j}{j} \frac{\varrho^i}{(1+\varrho)^{i+j+1}},$$

provided none of the c_j vanish. But before deducing $\pi(n, r)$ from equations (16) and (25) we need to know that the β_{ij} given by equations (25), are indeed of the form given in equation (15). Although they satisfy the conditions given by Mirsky ([14]) for the existence of the infinite matrix, we have not been able to prove that they are the correct coefficients for the busy period of a queue with balking.

V. S/M_μ queue, arrival rate independent of μ . If the probabilities $q(n, t)$ do not depend upon the value of μ , equation (1) is again the derivative of a Laplace transformation:

$$(26) \quad \pi(n, r) = \frac{r}{n} \cdot \frac{\mu^n (-1)^{n-1}}{(n-1)!} \cdot \frac{d^{n-1}}{d\mu^{n-1}} \bar{q}(n-r, \mu).$$

EXAMPLE 6. For an $E_{k,\lambda}/M_\mu$ queue, $q(n, t)$ is the Morse-Jewell distribution; cf. Morse ([15]), Jewell ([11]), Whittlesey and Haight ([22]),

$$(27) \quad q(n, t) = (1+n)[\Gamma_{nk+k} - \Gamma_{nk}] + (1-n)[\Gamma_{nk} - \Gamma_{nk-k}] + \\ + (\lambda t/k)[- \Gamma_{nk+k-1} + 2\Gamma_{nk-1} - \Gamma_{nk-k-1}],$$

where

$$(28) \quad \Gamma_n = \frac{1}{\Gamma(n)} \int_0^\infty e^{-x} x^{n-1} dx.$$

For an understanding of the rather curious form of the Morse-Jewell distribution, the formula of McFadden ([13], equations 2.15) relating it to the compound gap distribution is helpful. With some attention to detail, one can obtain from equation (27) the Laplace transformation

$$(29) \quad \bar{q}(n, s) = \frac{\lambda^{nk-k+1} [(\lambda+s)^k - \lambda^k]^2}{ks^2(\lambda+s)^{nk+k}}.$$

For an S/M_μ queue, the coefficients β_{ij} are

$$(30) \quad \beta_{ij} = \frac{\mu^{j+1} (-1)^j}{j!} \cdot \frac{d^j}{d\mu^j} \bar{q}(i, \mu).$$

Differentiating equation (30) with respect to μ , we obtain the following differential difference relationship for the β_{ij} :

$$(31) \quad \frac{d}{d\mu} \beta_{ij} = \frac{j+1}{\mu} (\beta_{ij} - \beta_{i,j+1}).$$

From equation (16) we know that

$$(32) \quad \sum_{n=0}^{\infty} \frac{r}{n+r} \beta_{n,n+r-1} = 1.$$

If we use equation (31) to differentiate equation (32), we find that

$$(33) \quad \frac{M(r)}{r} = \frac{M(r+1)}{r+1}$$

which yields

$$(34) \quad M(r) = rM(1).$$

Therefore the diagonal sums of (β_{ij}) , on and above the principal diagonal are all equal to $M(1)$.

Similarly, we can differentiate the equation defining $M(r)$ with the aid of equation (31), and obtain

$$(35) \quad \frac{\mu}{r} \cdot \frac{dM(r)}{d\mu} = \frac{M_2(r)}{r} - \frac{M_2(r+1)}{r+1} + M(1),$$

where $M_2(r)$ is the second moment of the discrete busy period distribution. Reducing this further, we can write

$$(36) \quad M_2(r) = rM_2(1) - r(r-1) \left[\mu \frac{dM(1)}{d\mu} - M(1) \right]$$

for the second moment, and obtain higher moments for r in terms of the first ($r = 1$) higher moment.

Finally, if we define the generating function

$$\Pi(s, r) = \sum_{j=r}^{\infty} s^j \pi(j, r),$$

we find without too much trouble that $\Pi(s, r)$ satisfies the equation

$$(37) \quad \mu \frac{\partial \Pi(s, r)}{\partial \mu} = \frac{1}{s} \cdot \frac{\partial \Pi(s, r)}{\partial s} - \frac{r}{r+1} \cdot \frac{\partial \Pi(s, r+1)}{\partial s}.$$

References

- [1] C. J. Ancker, Jr. and A. V. Gafarian, *Some queuing problems with balking and reneging — I*, Operations Research 11 (1) (1963), pp. 88-100.
- [2] — — *Some queuing problems with balking and reneging — II*, Operations Research 11 (6) (1963), pp. 928-937.
- [3] E. Borel, *Sur l'emploi du theoreme de Bernoulli pour faciliter le calcul d'une infinite de coefficients. Application au probleme de l'attente a un guichet*, C. R. Acad. Sci. Paris 214 (1942), pp. 452-456.
- [4] D. R. Cox and W. L. Smith, *Queues*, Methuen's Monographs on Applied Probability and Statistics, London 1961.
- [5] P. D. Finch, *Balking in the Queueing System GI/M/1*, Acta Mathematica Academiae Scientiarum Hungaricae 10 (1-2) (1959), pp. 241-246.
- [6] F. A. Haight, *Queueing with balking*, Biometrika 44 (1957), pp. 360-369.
- [7] — *Queueing with balking II*, Biometrika 47 (1960), pp. 285-296.
- [8] — *A distribution analogous to the Borel-Tanner*, Biometrika 48 (1961), pp. 167-173.
- [9] — *Expected utility for queues servicing messages with exponentially decaying utility*, Annals of Math. Statist. 32 (2) (1961), pp. 587-593.
- [10] — and M. A. Breuer, *The Borel-Tanner distribution*, Biometrika 47 (1960), pp. 143-150.
- [11] W. S. Jewell, *The properties of recurrent-event processes*, Operations Research 8 (4) (1960), pp. 446-472.

- [12] D. G. Kendall, *Stochastic processes occurring in the theory of queues and their analysis by the method of the imbedded Markov chain*, Annals of Mathematical Statistics 24 (5) (1953), pp. 338-354.
- [13] J. A. McFadden, *On the lengths of intervals in a stationary point process*, J. Royal Statist. Soc. (Ser. B) 24 (2) (1962), pp. 364-382.
- [14] L. Mirsky, *An existence theorem for infinite matrices*, Am. Math. Monthly 68 (5) (1961), pp. 465-469.
- [15] P. M. Morse, *Queues, Inventories and Maintenance*, Publications in Operations Research No. 1, New York 1958.
- [16] J. L. Mott, *The distribution of the time to emptiness of a discrete dam under steady demand*, J. Royal Statist. Soc. (Ser. B) 25 (1) (1963), pp. 137-139.
- [17] T. V. Narayana, J. Indian Soc. of Agricultural Statistics 6 (1954), pp. 139-148.
- [18] L. Takacs, *The stochastic law of the busy period for a single server queue with Poisson input*, J. of Math. Analysis and Applications 6 (1) (1963), pp. 33-42.
- [19] — *A single-server queue with feedback*, Bell System Technical J. 42 (2) (1963), pp. 505-519.
- [20] J. C. Tanner, *A problem of interference between two queues*, Biometrika 40 (1953), pp. 58-69.
- [21] — *A derivation of the Borel distribution*, Biometrika 48 (1 and 2) (1961), pp. 222-224.
- [22] J. R. B. Whittlesey, and F. A. Haight, *Counting distributions for an Erlang process*, Annals Inst. Statist. Math. 13 (2) (1961), pp. 91-103.

INSTITUTE OF TRANSPORTATION AND TRAFFIC ENGINEERING
UNIVERSITY OF CALIFORNIA, LOS ANGELES

Received on 15. 3. 1964

F. A. HAIGHT (Los Angeles)

DYSKRETNY ROZKŁAD OKRESÓW PRACY RÓŻNYCH JEDNOKANAŁOWYCH SYSTEMÓW OBSŁUGI MASOWEJ

STRESZCZENIE

Rozważmy jednokanałowy system obsługi masowej pracujący w warunkach równowagi statystycznej. Niech $\pi(n, r)$ oznacza prawdopodobieństwo tego, że w okresie od dowolnego momentu, w którym było r osób w systemie, do najbliższego momentu, w którym kanał będzie pusty, dokładnie n osób przejdzie przez system. Litera S w symbolice Kendalla będzie oznaczała stacjonarny proces punktowy. Autor wykazuje, że podana przez Tannera kombinatoryczna metoda wyznaczania $\pi(n, r)$ dla systemu $M/D/1$ może być również zastosowana w przypadkach $M/S/1$ i $S/M/1$. Pozwoliło to autorowi powiązać prawdopodobieństwa $\pi(n, r)$ z prawdopodobieństwami długości kolejki za pomocą nieskończonej macierzy. Macierz ta została zbadana dla różnych szczególnych przypadków systemów obsługi masowej, wśród nich także dla takich, w których proces wejścia zależy od średniej intensywności obsługi. Wyprowadzono cząstkowe równanie różniczkowo-różnicowe na funkcję tworzącą prawdopodobieństw $\pi(n, r)$.

Ф. А. ХАЙТ (Лос Анджелос)

**РАСПРЕДЕЛЕНИЯ ДИСКРЕТНЫХ ПЕРИОДОВ НЕПРЕРЫВНОЙ РАБОТЫ
РАЗНЫХ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ С ОДНИМ КАНАЛОМ
ОБСЛУЖИВАНИЯ**

РЕЗЮМЕ

Рассматривается система с одним каналом обслуживания работающая в условиях статистического равновесия. Пусть $\pi(n, r)$ вероятность того, что в интервале времени, начиная с любого момента, в котором в системе было r лиц и к ближайшему моменту, в котором система впервые становится пустой, точно n лиц перейдет через систему. Буква S в символике Кендалла будет обозначать стационарный точечный процесс. Автор показывает, что комбинаторный метод Таннера определения вероятностей $\pi(n, r)$ в системе $M/S/1$ применим тоже к системам $M/S/1$ и $S/M/1$. Используя этот метод автор связывает вероятности $\pi(n, r)$ с распределением длинны очереди посредством бесконечной матрицы. Эта матрица исследована автором для различных систем массового обслуживания, среди них для систем, в которых входный процесс зависит от средней интенсивности обслуживания. Выведем разностно-дифференциальное уравнение в частных производных для образующей функции вероятностей $\pi(n, r)$.
