Probability in the Engineering and Informational Sciences, **21**, 2007, 67–76. Printed in the U.S.A.

NEGATIVE PROBABILITIES AT WORK IN THE $M/D/1$ QUEUE

Henk Tijms
Department of Econometrics and Operations Research
Vrije University
1081 HV Amsterdam, The Netherlands
and
Tinbergen Institute Amsterdam
1018 WB Amsterdam, The Netherlands
E-mail: tijms@feweb.vu.nl

Koen Staats
Department of Econometrics and Operations Research
Vrije University
1081 HV Amsterdam, The Netherlands

This article derives amazingly accurate approximations to the state probabilities and waiting-time probabilities in the $M/D/1$ queue using a two-phase process with negative probabilities to approximate the deterministic service time. The approximations are in the form of explicit expressions involving geometric and exponential terms. The approximations extend to the finite-capacity $M/D/1/N+1$ queue.

1. INTRODUCTION

The $M/D/1$ queue with deterministic services is a simple and generally useful model. It is an appropriate model for telecommunication applications among others (for instance, to describe the cell scale queuing problem in ATM multiplexes). In view of the wide scope of applicability of the $M/D/1$ model in modern performance evaluation, it is important to have quick and accurate approximations for the state probabilities and the waiting-time probabilities. In this article such approximation will be provided in the form of explicit expressions that are in terms of geometric distributions and exponential densities.

 0269-9648/07 $16.00

The derivation of the approximation is based on a largely unknown but extremely useful approximation idea of Nojo and Watanabe [4]. It is well known that the probability distribution of a nonnegative random variable with a squared coefficient of variation of at least $\frac{1}{2}$ can be approximated by a Coxian-2 distribution by matching the first two moments (see, e.g., Tijms [5]). The squared coefficient of variation is the ratio of the variance and the squared mean. In their article, Nojo and Watanabe [4] presented a method to approximate the probability distribution of a nonnegative random variable with a squared coefficient of variation less than $\frac{1}{2}$ by a two-phase distribution with branching probabilities that are negative or larger than one. Using this magical idea, continuous-time Markov chain analysis can be used by doing calculations with negative numbers as if they were legitimate probabilities (see also Van Hoorn and Seelen [6]). The approximation of a random lifetime X through a two-phase process is as follows. The process starts in phase 1. It stays in phase 1 for an exponentially distributed time with mean $1/\gamma$. Upon completion of the sojourn time in phase 1, the process expires with probability r_1 and moves to phase 2 with probability $1 - r_1$. The sojourn time in phase 2 is also exponentially distributed with mean $1/\gamma$. Upon completion of the sojourn time in phase 2, the process expires with probability r_2 and returns to phase 1 with probability $1 - r_2$. In phase 1, the process starts anew. The idea is to approximate the original lifetime X by the time it takes until the two-phase process expires. For a random variable X with squared coefficient of variation $c_X^2 < \frac{1}{2}$, it is, under certain side conditions, possible to fit a two-phase process matching the first *three* moments of X when a negative value or a value larger than one are allowed for the branching probabilities r_1 and r_2. These side conditions are satisfied when X has a deterministic distribution. Using the Laplace transform

$$f^*(s) = \frac{\gamma r_1 s + \gamma^2(r_1 + r_2 - r_1 r_2)}{s^2 + 2\gamma s + \gamma^2(r_1 + r_2 - r_1 r_2)} \tag{1}$$

of the density of the expiration time in the two-phase process, it is matter of algebra to derive that

$$\gamma = \frac{2}{D}, \qquad r_1 = -1, \qquad r_2 = \frac{5}{4} \tag{2}$$

when X is deterministic and equals the constant D.

The remainder of the article is organized as follows. In Section 2 we present the explicit expression for the approximations to the state probabilities and the waiting-time probabilities in the $M/D/1$ queue with infinite waiting room. Approximations for the finite-capacity $M/D/1/N+1$ queue are discussed in Section 3. We also give numerical results showing how amazingly accurate the approximations are.

2. THE *M/D/1* QUEUE

In the single-server $M/D/1$ queue, the arrival process of customers is a Poisson process with rate λ and the service time of any customer is a constant D. There is an infinite waiting room. Customers are served in order of arrival. It is assumed that the offered load,

$$\rho = \lambda D,$$

is smaller than one. Denoting by $p_j(t)$ the probability of having j customers in the system at time t, the state probability p_j is defined by

$$p_j = \lim_{t\to\infty} p_j(t), \qquad j = 0,1,\ldots.$$

Also, letting D_n denote the delay in queue of the nth arrival, the waiting-distribution function $W_q(x)$ is defined by

$$W_q(x) = \lim_{n\to\infty} P(D_n \le x), \qquad x \ge 0.$$

Numerical methods for the exact computation of the p_j and $W_q(x)$ are available (see, e.g., Tijms [5]). However, for modern performance analysis in telecommunication applications, it is useful to have quick and accurate approximations.

APPROXIMATION RESULT 1: *For any $x \ge 0$, the complementary waiting-time distribution function $1 - W_q(x)$ can be approximated by*

$$\rho\left[\frac{\rho + 2 + c(D,\rho)}{2c(D,\rho)}\right] e^{(1/2D)[\rho-4+c(D,\rho)]x} - \rho\left[\frac{\rho + 2 - c(D,\rho)}{2c(D,\rho)}\right] e^{(1/2D)[\rho-4-c(D,\rho)]x},$$

where $c(D,\rho)$ is the real number $\sqrt{\rho^2 + 16\rho - 8}$ for $\rho \ge -8 + 6\sqrt{2}$ and $c(D,\rho)$ is the complex number $i\sqrt{8 - 16\rho - \rho^2}$ otherwise.

The derivation of this result proceeds as follows. In the general $M/G/1$ queue, the Laplace transform of the complementary waiting-time distribution $1 - W_q(x)$ is given by

$$\int_0^\infty e^{-sx}(1 - W_q(x))\,dx = \frac{\rho s - \lambda + \lambda f^*(s)}{s[s - \lambda + \lambda f^*(s)]}, \qquad \textbf{(3)}$$

where $f^*(s)$ is the Laplace transform of the service time density (see, e.g., Tijms [5]). Let us now substitute in the right-hand side of (3) the Laplace transform (1) with $\gamma = 2/D$, $r_1 = -1$, and $r_2 = \frac{5}{4}$. In other words, the deterministic service distribution is approximated by the two-phase distribution discussed in Section 1. The right-hand side of (3) then becomes, after some algebra,

$$\frac{\rho D^2 s + 3\rho D}{D^2 s^2 + (4-\rho)Ds + 6 - 6\rho}.$$

Denote the numerator of this expression by $N(s)$ and the denominator by $Q(s)$. The Laplace transform $N(s)/Q(s)$ can be analytically inverted, using partial fraction expansion. The (complex) roots of the denominator $Q(s) = D^2s^2 + (4 - \rho)Ds + 6 - 6\rho$ are

$$s_{1,2} = \frac{1}{2D}\left(\rho - 4 \pm \sqrt{\rho^2 + 16\rho - 8}\right).$$

Partial fraction expansion gives

$$\frac{N(s)}{Q(s)} = \frac{N(s_1)}{Q'(s_1)}\,\frac{1}{s - s_1} + \frac{N(s_2)}{Q'(s_2)}\,\frac{1}{s - s_2}.$$

Then inversion gives that $N(s)/Q(s)$ is the Laplace transform of

$$\frac{N(s_1)}{Q'(s_1)}\,e^{s_1 x} + \frac{N(s_2)}{Q'(s_2)}\,e^{s_2 x}, \qquad x \geq 0.$$

It is then a matter of simple algebra to get Result 1.

For several values of ρ, we give in Table 1 the approximate and exact values of $1 - W_q(x)$ for a number of x values. In all of the examples, the normalization $D = 1$ is used. The results in Table 1 show that the approximation is remarkably accurate. The accuracy also applies for small tail probabilities. However, depending on ρ, for very small tail probabilities, the approximate value of $1 - W_q(x)$ might become negative. Fortunately, for very small tail probabilities, one has available the useful asymptotic expansion

$$1 - W_q(x) \sim \frac{1-\rho}{\tau\rho - 1}\,e^{-\lambda(\tau-1)x} \quad \text{as } x \to \infty,$$

where $\tau = 1/\eta$ with η the unique root of the equation $\rho(1 - x) + x\ln(x) = 0$ on (0,1) (see, e.g., Tijms [5, p. 383]). It is interesting to observe that the solution $\eta = \eta(\rho)$ can be very well approximated by the polynomial

$$\begin{aligned}\eta_{\text{app}}(\rho) = {} & -2.4589155\rho^9 + 12.2135282\rho^8 - 25.9394588\rho^7 + 30.8300743\rho^6 \\ & - 22.5757925\rho^5 + 10.7001400\rho^4 - 3.1972952\rho^3 + 1.2583839\rho^2 \\ & + 0.1696439\rho - 0.0003216.\end{aligned}$$

The absolute error of this polynomial approximation is less than 1.1×10^{-5} for all practically relevant values of ρ (i.e., for all $\rho \geq 0.2$).

Next, we turn to approximations for the state probabilities. For the general $M/G/1$ queue, the generating function $P(z) = \sum_{j=0}^{\infty} p_j z^j$ of the state probabilities p_j is given by

Table 1. Waiting-Time Probabilities in $M/D/1$ Queue

x		$\rho = 0.2$	$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.8$	$\rho = 0.9$	$\rho = 0.95$
0.25	exa	0.1590	0.4334	0.6426	0.7557	0.8748	0.9366
	app	0.1536	0.4244	0.6347	0.7496	0.8712	0.9347
0.5	exa	0.1159	0.3580	0.5743	0.7016	0.8432	0.9196
	app	0.1066	0.3407	0.5579	0.6885	0.8353	0.9153
1.5	exa	0.0085	0.1020	0.2917	0.4553	0.6848	0.8303
	app	0.0109	0.1083	0.2965	0.4579	0.6853	0.8301
2.5	exa	5.7E−4	0.0286	0.1477	0.2951	0.5561	0.7496
	app	−8.8E−4	0.0285	0.1500	0.2975	0.5572	0.7500
3.5	exa	3.9E−5	0.0081	0.0751	0.1918	0.4520	0.6771
	app	−3.7E−4	0.0070	0.0754	0.1928	0.4529	0.6775
5.5	exa		6.6E−4	0.0195	0.0810	0.2987	0.5525
	app		3.8E−4	0.0190	0.0810	0.2991	0.5527
7.5	exa		5.3E−5	0.0050	0.0342	0.1974	0.4508
	app		1.9E−5	0.0048	0.0340	0.1975	0.4509
10.5	exa			6.6E−4	0.0094	0.1060	0.3322
	app			6.0E−4	0.0093	0.1060	0.3323
15.5	exa			2.3E−5	0.0011	0.0376	0.1998
	app			1.9E−5	0.0011	0.0376	0.1998
20.5	exa				1.3E−4	0.0134	0.1201
	app				1.2E−4	0.0133	0.1201
25.5	exa				1.5E−5	0.0047	0.0722
	app				1.4E−5	0.0047	0.0722
50.5	exa					2.7E−5	0.0057
	app					2.6E−5	0.0057

$$P(z) = \frac{(1-\rho)(1-z)f^*(\lambda(1-z))}{f^*(\lambda(1-z)) - z}, \tag{4}$$

where $f^*(s)$ is the Laplace transform of the service time density (see, e.g., Tijms [5]). Substituting (1) with $\gamma = 2/D$, $r_1 = -1$, and $r_2 = \frac{5}{4}$ in (4), we get, after some algebra, that

$$P(z) = \frac{V(z)}{W(z)},$$

where

$$V(z) = -2\rho(1-\rho)z^2 + (4\rho - 6)(1-\rho)z + (6 - 2\rho)(1-\rho), \tag{5}$$

$$W(z) = -\rho^2 z^3 + (2\rho^2 + 4\rho)z^2 - (\rho^2 + 2\rho + 6)z + 6 - 2\rho. \tag{6}$$

Denoting the complex roots of $W(z) = 0$ by z_1, z_2, and z_3, we then find by partial fraction expansion that

$$P(z) = \frac{a_1}{z - z_1} + \frac{a_2}{z - z_2} + \frac{a_3}{z - z_3}, \tag{7}$$

where

$$a_i = \frac{V(z_i)}{W'(z_i)} \quad \text{for } i = 1,2,3. \tag{8}$$

The right-hand side of (7) can be inverted as

$$\frac{a_1}{z - z_1} + \frac{a_2}{z - z_2} + \frac{a_3}{z - z_3} = -\sum_{j=0}^{\infty} \left(\frac{a_1}{z_1^{j+1}} + \frac{a_2}{z_2^{j+1}} + \frac{a_3}{z_3^{j+1}} \right) z^j.$$

Defining the constants p and q by

$$p = -\frac{1}{3} \frac{(2\rho + 4)^2}{\rho^2} + \frac{\rho^2 + 2\rho + 6}{\rho^2},$$

$$q = -\frac{2}{27} \frac{(2\rho + 4)^3}{\rho^3} + \frac{1}{3} \frac{2\rho + 4}{\rho} \frac{\rho^2 + 2\rho + 6}{\rho^2} + \frac{2\rho - 6}{\rho^2},$$

it is a matter of tedious algebra to verify that the roots z_1, z_2, and z_3 of the equation $W(z) = 0$ are given by

$$z_i = w_i - \frac{p}{3w_i} + \frac{2\rho + 4}{3\rho} \quad \text{for } i = 1,2,3, \tag{9}$$

where w_1, w_2, and w_3 are the complex roots of

$$w^3 = \frac{1}{2} \left\{ -q + \sqrt{q^2 + \frac{4p^3}{27}} \right\}. \tag{10}$$

Here we used the trick of reducing the equation $x^3 + px + q = 0$ to the equation $w^6 + qw^3 - (p^3/27) = 0$ by the change of variable $x = w - (p/3w)$. It is standard to solve (10). The complex roots of the equation $u^3 = 1$ are calculated as

$$u_1 = 1, \qquad u_2 = e^{i2\pi/3}, \qquad u_3 = e^{-i2\pi/3}. \tag{11}$$

Summarizing, we have the following result.

Approximation Result 2: *For any $j = 0,1,\ldots,$ the state probability p_j in the $M/D/1$ queue can be approximated by*

$$-a_1 z_1^{-(j+1)} - a_2 z_2^{-(j+1)} - a_3 z_3^{-(j+1)},$$

where z_1, z_2, and z_3 are calculated from (9)–(11) and the constants a_1, a_2, and a_3 are calculated from (8).

The calculations for the approximate p_j are simple and offer no numerical difficulties, unlike the calculations with the algebraic expression for the exact p_j in Brun and Garcia [1]. For several values of ρ, we give in Table 2 the approximate and exact values of the tail probabilities

$$Q_j = \sum_{k=j+1}^{\infty} p_k.$$

The results in Table 2 show that the approximations are remarkably accurate. The accuracy also holds for small tail probabilities. However, for very small values of Q_j, it is better to use the asymptotic expansion

$$Q_j \sim \frac{1-\rho}{\tau\rho - 1}\tau^{-j} \quad \text{as } j \to \infty,$$

TABLE 2. The State Probabilities Q_j in the $M/D/1$ Queue

j		$\rho = 0.2$	$\rho = 0.5$	$\rho = 0.7$	$\rho = 0.8$	$\rho = 0.9$	$\rho = 0.95$
0	exa	0.2000	0.5000	0.7000	0.8000	0.9000	0.9500
	app	0.2000	0.5000	0.7000	0.8000	0.9000	0.9500
1	exa	0.0229	0.1756	0.3959	0.5549	0.7540	0.8707
	app	0.0229	0.1750	0.3941	0.5527	0.7521	0.8695
2	exa	0.0020	0.0530	0.2063	0.3655	0.6164	0.7885
	app	0.0020	0.0538	0.2073	0.3660	0.6162	0.7881
3	exa	1.5E−4	0.0153	0.1053	0.2379	0.5013	0.7124
	app	1.4E−4	0.0154	0.1063	0.2390	0.5019	0.7126
5	exa		0.0012	0.0273	0.1005	0.3313	0.5813
	app		0.0011	0.0273	0.1008	0.3318	0.5815
7	exa		1.0E−4	0.0071	0.0425	0.2189	0.4743
	app		7.9E−5	0.0069	0.0424	0.2192	0.4745
10	exa			9.3E−4	0.0117	0.1176	0.3495
	app			8.9E−4	0.0116	0.1176	0.3497
15	exa			3.2E−5	0.0014	0.0417	0.2102
	app			2.9E−5	0.0013	0.0417	0.2102
20	exa				1.6E−4	0.0148	0.1264
	app				1.5E−4	0.0148	0.1264
25	exa				1.8E−5	0.0053	0.0760
	app				1.7E−5	0.0052	0.0760
50	exa					3.0E−5	0.0060
	app					2.9E−5	0.0060

where τ is the same as in the asymptotic expansion of $1 - W_q(x)$ (see Tijms [5, p. 378]).

3. THE $M/D/1/N + 1$ QUEUE

The $M/D/1/N + 1$ queue differs only from the $M/D/1$ queue by having a finite waiting room with capacity N rather than an infinite waiting room. An arriving customer who finds, upon arrival, the server busy and all of the N additional waiting places occupied is rejected and has no further influence on the system. In the discussion below, it is again assumed that $\rho = \lambda D$ is smaller than one. Under this assumption, the stationary state probabilities $p_j^{(N)}$ in the $M/D/1/N+1$ queue can be expressed in the corresponding state probabilities p_j in the $M/D/1$ queue. The state probabilities $p_j^{(N)}$ satisfy

$$p_j^{(N)} = \begin{cases} \gamma p_j & \text{for } j = 0,1,\ldots,N, \\ 1 - \sum_{k=0}^{N} p_k^{(N)} & \text{for } j = N+1, \end{cases}$$

where the proportionality constant γ is given by

$$\gamma = \left[1 - \rho\left(1 - \sum_{j=0}^{N} p_j\right)\right]^{-1}$$

(see, e.g., Tijms [5, p. 410]). Thus, Approximation Result 2 leads directly to simple approximations for the state probabilities $p_j^{(N)}$. In turn, the state probabilities $p_j^{(N)}$ can be used to compute the stationary waiting-time distribution function $W_q(x)$, using a beautiful result of Franx [2,3]. In Franx [2], a simple and numerically stable method was given to calculate the waiting-time probabilities in the multiserver $M/D/c$ queue with infinite capacity. This exact method can be modified for the finite-capacity model when $c = 1$. For the $M/D/1/N + 1$ model, let us define $W_q(x) = \lim_{n\to\infty} P(D_n \leq x)$, with D_n denoting the delay in queue of the nth *accepted* customer. The following exact result holds for the $M/D/1/N + 1$ queue. For $(k - 1)D \leq x < kD$ and $k = 1,\ldots,N$, the waiting-time probability $W_q(x)$ is given by

$$W_q(x) = \frac{1}{1 - p_{N+1}^{(N)}} \sum_{j=0}^{k-1} Q_{k-1-j}\, e^{-\lambda(kD-x)} \frac{[\lambda(kD - x)]^j}{j!},$$

where $Q_r = \sum_{i=0}^{r+1} p_i^{(N)}$ for $r = 0,1,\ldots,N$. This exact formula applies to the case of $\rho \geq 1$ as well. For the case of $\rho < 1$, we can use the simple approximations for the $p_j^{(N)}$ in order to get an easily calculated approximation for $W_q(x)$.

In Table 3 we show that Approximation Result 2 leads to excellent approximations to the $p_j^{(N)}$. For several values of ρ and N, we give the approximate and

TABLE 3. Various Performance Measures for the $M/D/1/N+1$ Queue

ρ		$N=2$				$N=5$			
		P_{rej}	P_{delay}	$E(L_q)$	$\sigma(L_q)$	P_{rej}	P_{delay}	$E(L_q)$	$\sigma(L_q)$
0.5	exa	0.0272	0.4720	0.1804	0.4498	0.0006	0.4994	0.2466	0.6136
	app	0.0276	0.4716	0.1798	0.4503	0.0006	0.4994	0.2469	0.6153
0.8	exa	0.1033	0.6848	0.4743	0.6752	0.0219	0.7777	1.0998	1.3618
	app	0.1035	0.6845	0.4710	0.6754	0.0219	0.7776	1.0993	1.3652
0.9	exa	0.1384	0.7393	0.5860	0.7208	0.0472	0.8505	1.5955	1.5580
	app	0.1383	0.7395	0.5819	0.7211	0.0473	0.8503	1.5928	1.5610
0.95	exa	0.1571	0.7636	0.6419	0.7377	0.0649	0.8806	1.8656	1.6169
	app	0.1568	0.7640	0.6374	0.7381	0.0650	0.8805	1.8617	1.6199
0.99	exa	0.1725	0.7816	0.6860	0.7486	0.0813	0.9015	2.0839	1.6425
	app	0.1720	0.7823	0.6814	0.7491	0.0813	0.9015	2.0793	1.6455

exact values of the rejection probability $P_{\text{rej}} = p_{N+1}^{(N)}$, the delay probability $P_{\text{delay}} = \sum_{j=1}^{N} p_j^{(N)}/(1 - p_{N+1}^{(N)})$, $E(L_q)$, and $\sigma(L_q)$, where $E(L_q)$ and $\sigma(L_q)$ denote the expected value and the standard deviation of the stationary queue size L_q.

For several values of ρ and N, we give in Table 4 the approximate and exact values of $1 - W_q(x)$ for a number of x values. The results in Table 4 show an excellent performance of the approximation.

Summarizing, we can conclude for the $M/D/1/N+1$ queue that approximating the deterministic service time distribution by a two-phase process with a negative

TABLE 4. Results for $1 - W_q(x)$ in the $M/D/1/N+1$ Queue

ρ		$N=2$				$N=5$			
		$x = 0.5$	0.75	1.5	1.9	$x = 0.5$	1.75	3.9	4.5
0.5	exa	0.3220	0.2318	0.0517	0.0074	0.3572	0.0722	0.0037	0.0011
	app	0.3210	0.2306	0.0514	0.0073	0.3568	0.0729	0.0037	0.0011
0.8	exa	0.5298	0.4257	0.1416	0.0251	0.6683	0.3414	0.0677	0.0269
	app	0.5271	0.4224	0.1405	0.0248	0.6666	0.3413	0.0683	0.0270
0.9	exa	0.5912	0.4880	0.1784	0.0333	0.7655	0.4759	0.1268	0.0541
	app	0.5882	0.4843	0.1771	0.0330	0.7635	0.4747	0.1271	0.0542
0.95	exa	0.6198	0.5179	0.1975	0.0378	0.8080	0.5439	0.1644	0.0724
	app	0.6169	0.5142	0.1961	0.0374	0.8060	0.5423	0.1644	0.0724
0.99	exa	0.6417	0.5411	0.2131	0.0415	0.8384	0.5966	0.1976	0.0893
	app	0.6388	0.5374	0.2116	0.0411	0.8365	0.5948	0.1975	0.0892

branching probability leads to remarkably accurate approximations that are easy to use in practical applications. The approximation approach works only for the steady-state behavior of the system, not for the transient behavior. It is our conjecture that the approximation approach in conjunction with numerical Laplace inversion will lead to practically useful approximations for the stationary waiting-time probabilities in the notoriously difficult multiserver $M/D/c/N + c$ queue. This will be the subject of further research.

References

1. Brun, O. & Garcia, J. (2000). Analytical solution of finite capacity $M/D/1$ queues. *Journal of Applied Probability* 37: 1092–1098.
2. Franx, G.J. (2001). A simple solution for the $M/D/c$ waiting-time distribution. *Operations Research Letters* 29: 221–229.
3. Franx, G.J. (2004). A full probabilistic approach to waiting-time distributions. Doctoral dissertation, Vrije University, Amsterdam.
4. Nojo, S. & Watanabe, H. (1987). A new stage method getting arbitrary coefficient of variation through two stages. *Transactions of the IECIE* 70: 33–36.
5. Tijms, H.C. (2003). *A first course in stochastic models*. Chichester: Wiley.
6. Van Hoorn, M.H. & Seelen, L.P. (1986). Approximations for the $G/G/c$ queue. *Journal of Applied Probability* 23: 484–494.