Online Supplement:
Appointment Scheduling under Patient No-Shows and Service Interruptions

Jianzhe Luo, Vidyadhar G. Kulkarni, Serhan Ziya
Department of Statistics and Operations Research
University of North Carolina, Chapel Hill, NC 27599
jzluo@email.unc.edu, vkulkarn@email.unc.edu, ziya@email.unc.edu

---

Appendix A: Proofs of the Analytical Results and the Complete Statement of Theorem 3

Proof of Theorem 1:

Depending on the server state and the number of scheduled patients in the system, the following events might occur during $(d_{k+1}-(t+h), d_{k+1}-t)$ (Kulkarni 1995, p. 206):

If no patient is in the system and the server is available at $d_{k+1}-(t+h)$, then at $d_{k+1}-t$ the server becomes unavailable with probability $\eta h+o(h)$, or stays available with probability $1-\eta h+o(h)$.

If there is at least one scheduled patient in the system and the server is available at $d_{k+1}-(t+h)$, then at $d_{k+1}-t$ the number of scheduled patients in the system is reduced by 1 with probability $\mu h+o(h)$, or the server becomes unavailable with probability $\eta h+o(h)$, or both the number of scheduled patients in the system and the server state remain unchanged with probability $1-\mu h-\eta h+o(h)$.

If the server is unavailable at $d_{k+1}-(t+h)$, then at $d_{k+1}-t$, the number of patients in the system does not change, and the server becomes available with probability $\theta h+o(h)$, or stays unavailable with probability $1-\theta h+o(h)$.

Then, for $k=0,1,\ldots,N$, we have

$$R^k_{0,0}(t+h)=(\eta h+o(h))R^k_{0,1}(t)+(1-\eta h+o(h))R^k_{0,0}(t),$$

$$R^k_{0,1}(t+h)=(\theta h+o(h))R^k_{0,0}(t)+(1-\theta h+o(h))R^k_{0,1}(t),$$

$$R^k_{n,0}(t+h)=-(n-1)c_w h+(\mu h+o(h))R^k_{n-1,0}(t)+(\eta h+o(h))R^k_{n,1}(t)+[1-(\eta+\mu)h+o(h)]R^k_{n,0}(t),$$

$$R^k_{n,1}(t+h)=-nc_w h+(\theta h+o(h))R^k_{n,0}(t)+(1-\theta h+o(h))R^k_{n,1}(t),\ n=1,\ldots,k.$$

Letting $h\to 0$, after some algebra, we have the following:

$$\frac{\mathrm{d}R^k_{0,0}(t)}{\mathrm{d}t}=-\eta R^k_{0,0}(t)+\eta R^k_{0,1}(t),$$

$$\frac{\mathrm{d}R^k_{0,1}(t)}{\mathrm{d}t}=\theta R^k_{0,0}(t)-\theta R^k_{0,1}(t),$$

$$\frac{\mathrm{d}R^k_{n,0}(t)}{\mathrm{d}t}=-(n-1)c_w+\mu R^k_{n-1,0}(t)+\eta R^k_{n,1}(t)-(\eta+\mu)R^k_{n,0}(t),$$

$$\frac{\mathrm{d}R^k_{n,1}(t)}{\mathrm{d}t}=-nc_w+\theta R^k_{n,0}(t)-\theta R^k_{n,1}(t),\ n=1,\ldots,k.$$

Equations (5) and (6) are then obtained by writing the above differential equations in the matrix form.

To obtain the boundary conditions, we start with the net profit that would be incurred after T. When there are $n \geq 1$ scheduled patients in the system just prior to T and the server is available, the expected amount of time the system will continue to be operated is $E(nX)$, which is also equal to the expected server overtime. And the expected total waiting time these n scheduled patients will spend in the system is

$$E\left[nX + (n-1)X + \cdots + X\right] - \frac{n}{\mu} = \frac{n(n+1)}{2}E(X) - \frac{n}{\mu}.$$

If the server is unavailable, the expected overtime and the expected waiting time for each scheduled patient in the system are increased by $1/\theta$.

We also need to state the boundary conditions across appointment intervals. Specifically, at time d_k, $k = 1, 2, \ldots, N$, if the kth scheduled patient shows up, the system earns a reward r, the total number of patients in the system is increased by 1, and the number of pending appointments is decreased by 1. Otherwise, the system earns no reward, the total number of patients in the system remains unchanged, and the number of pending appointments is decreased by 1. The server status does not change in either case.

Proof of Theorem 2:

For each $k = 0, 1, \ldots, N$, take the LT of (5) and (6). We have

$$s\tilde{R}_0^k(s) - R_0^k(0^+) = ER_0^k(s),$$

$$s\tilde{R}_n^k(s) - R_n^k(0^+) = -\begin{bmatrix} n-1 & 0 \\ 0 & n \end{bmatrix} \frac{C_w}{s} + A\tilde{R}_{n-1}^k(s) + B\tilde{R}_n^k(s),\ n = 1, \ldots, k.$$

Thus,

$$\tilde{R}_0^k(s) = (sI - E)^{-1} R_0^k(0^+),$$

$$\begin{aligned}\tilde{R}_n^k(s) &= [(sI-B)^{-1}A]\tilde{R}_{n-1}^k(s) + (sI-B)^{-1}[-\begin{bmatrix} n-1 & 0 \\ 0 & n \end{bmatrix} \frac{C_w}{s} + R_n^k(0^+)] \\ &= [(sI-B)^{-1}A]^n (sI-E)^{-1}R_0^k(0^+) + \sum_{j=0}^{n-1}\{[(sI-B)^{-1}A]^j (sI-B)^{-1}[-\begin{bmatrix} n-1-j & 0 \\ 0 & n-j \end{bmatrix} \frac{C_w}{s} + R_{n-j}^k(0^+)]\}.\end{aligned}$$

Proof of Theorem 3:

For each $k = 0, 1, \ldots, N$, define $D_0^k(t) = e^{Et}R_0^k(0^+)$, and $D_n^k(t) = e^{Bt}[\int_0^t e^{-Bs}AD_{n-1}^k(s)\mathrm{d}s + R_n^k(0^+) - z_n^k]$, $n = 1, \ldots, k$. From Theorem 1, it is easy to verify that $R_0^k(t) = e^{Et}R_0^k(0^+) = D_0^k(t) + z_0^k$. Let $\bar{R}_n^k(t) = e^{-Bt}R_n^k(t)$ and take the derivative on both sides, we then have

$$\frac{\mathrm{d}\bar{R}_n^k(t)}{\mathrm{d}t} = e^{-Bt}[AR_{n-1}^k(t) - \begin{bmatrix} n-1 & 0 \\ 0 & n \end{bmatrix} C_w] = e^{-Bt}[AD_{n-1}^k(t) + Az_{n-1}^k - \begin{bmatrix} n-1 & 0 \\ 0 & n \end{bmatrix} C_w].$$

(The induction that $R_{n-1}^k(t) = D_{n-1}^k(t) + z_{n-1}^k$ is used in the above derivation.)

Integrating both sides of the above equation yields $\bar{R}_n^k(t) = \int_0^t e^{-Bs}AD_{n-1}^k(s)\mathrm{d}s + R_n^k(0^+) + (e^{-Bt} - I)B^{-1}(\begin{bmatrix} n-1 & 0 \\ 0 & n \end{bmatrix} C_w - Az_{n-1}^k)$, which implies that $R_n^k(t) = e^{Bt}[\int_0^t e^{-Bs}AD_{n-1}^k(s)\mathrm{d}s + R_n^k(0^+) - z_n^k] + z_n^k = D_n^k(t) + z_n^k$.

To obtain the explicit recursive expression of $D_n^k(t)$, we expand e^{Bt}, e^{-Bt}, and e^{Et} in the matrix form. First, we can write matrices B and E as follows:

$$B = \frac{\theta}{b-a}\begin{bmatrix} \frac{-a+\theta}{\theta} & \frac{-b+\theta}{\theta} \\ 1 & 1 \end{bmatrix}\begin{bmatrix} -a & 0 \\ 0 & -b \end{bmatrix}\begin{bmatrix} 1 & \frac{b-\theta}{\theta} \\ -1 & \frac{-a+\theta}{\theta} \end{bmatrix}, \quad E = \frac{1}{\eta+\theta}\begin{bmatrix} 1 & \eta \\ 1 & -\theta \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 0 & -(\eta+\theta) \end{bmatrix}\begin{bmatrix} \theta & \eta \\ 1 & -1 \end{bmatrix}.$$

Hence,

$$\begin{aligned} e^{Bt} &= \frac{\theta}{b-a}\begin{bmatrix} \frac{-a+\theta}{\theta} & \frac{-b+\theta}{\theta} \\ 1 & 1 \end{bmatrix}\begin{bmatrix} e^{-at} & 0 \\ 0 & e^{-bt} \end{bmatrix}\begin{bmatrix} 1 & \frac{b-\theta}{\theta} \\ -1 & \frac{-a+\theta}{\theta} \end{bmatrix} = \frac{\theta}{b-a}(He^{-at} + Je^{-bt}), \\ e^{-Bt} &= \frac{\theta}{b-a}\begin{bmatrix} \frac{-a+\theta}{\theta} & \frac{-b+\theta}{\theta} \\ 1 & 1 \end{bmatrix}\begin{bmatrix} e^{at} & 0 \\ 0 & e^{bt} \end{bmatrix}\begin{bmatrix} 1 & \frac{b-\theta}{\theta} \\ -1 & \frac{-a+\theta}{\theta} \end{bmatrix} = \frac{\theta}{b-a}(He^{at} + Je^{bt}), \\ e^{Et} &= \frac{1}{\eta+\theta}\begin{bmatrix} 1 & \eta \\ 1 & -\theta \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 0 & e^{-(\eta+\theta)} \end{bmatrix}\begin{bmatrix} \theta & \eta \\ 1 & -1 \end{bmatrix} = \frac{1}{\eta+\theta}(L - Ee^{-(\eta+\theta)t}). \end{aligned}$$

Having the above expansions, we can further write

$$D_0^k(t) = e^{Et}R_0^k(0^+) = (L - Ee^{-(\eta+\theta)t})\frac{R_0^k(0^+)}{\eta+\theta}, \text{ and for } n = 1,2,\ldots,k,$$
$$D_n^k(t) = \frac{\theta}{b-a}(He^{-at}+Je^{-bt})(R_n^k(0^+)-z_n^k) + (\frac{\theta}{b-a})^2(He^{-at}+Je^{-bt})\int_0^t (He^{as}+Je^{bs})AD_{n-1}^k(s)\mathrm{d}s.$$

Now define, for $k = 0,1,\ldots,N$, and $n = 1,\ldots,k$,

$$m_{0,0}^{0,k} = q_{0,0}^{0,k} = [0,0]',\ u_0^{0,k} = \frac{LR_0^k(0^+)}{\eta+\theta},\ v_0^{0,k} = \frac{-ER_0^k(0^+)}{\eta+\theta},$$
$$u_j^{n,k} = (\frac{\theta}{b-a})^2[H^2A\frac{u_j^{n-1,k}}{a+j(a-b)} + J^2A\frac{u_j^{n-1,k}}{b+j(a-b)}],\ j=-n,\ldots,n,$$
$$v_j^{n,k} = (\frac{\theta}{b-a})^2[H^2A\frac{v_j^{n-1,k}}{a+j(a-b)-(\eta+\theta)} + J^2A\frac{v_j^{n-1,k}}{b+j(a-b)-(\eta+\theta)}],\ j=-n,\ldots,n,$$
$$m_{0,0}^{n,k} = \frac{\theta}{b-a}H(R_n^k(0^+)-z_n^k) + (\frac{\theta}{b-a})^2\{H^2A[\sum_{j=1}^{n-2}\sum_{i=0}^{n-2-j}\frac{i!}{[j(a-b)]^{i+1}}m_{i,j}^{n-1,k} - \sum_{j=-n+1}^{n-1}\frac{u_j^{n-1,k}}{a+j(a-b)}$$
$$-\sum_{j=-n+1}^{n-1}\frac{v_j^{n-1,k}}{-\eta-\theta+a+j(a-b)} + \sum_{j=0}^{n-2}\sum_{i=0}^{n-2-j}\frac{i!}{[(j+1)(b-a)]^{i+1}}q_{i,j}^{n-1,k}] - J^2A\sum_{i=0}^{n-2}i!(a-b)^{-(i+1)}m_{i,0}^{n-1,k}\},$$
$$m_{i,0}^{n,k} = (\frac{\theta}{b-a})^2[H^2A\frac{m_{i-1,0}^{n-1,k}}{i} - J^2A\sum_{s=i}^{n-2}\frac{s!}{i!}(a-b)^{i-s-1}m_{s,0}^{n-1,k}],\ i=1,2,\ldots,n-1,$$
$$m_{i,j}^{n,k} = (\frac{\theta}{b-a})^2\{-H^2A\sum_{s=i}^{n-2-j}\frac{s!}{i!}[j(a-b)]^{i-s-1}m_{s,j}^{n-1,k} - J^2A\sum_{s=i}^{n-2-j}\frac{s!}{i!}[(j+1)(a-b)]^{i-s-1}m_{s,j}^{n-1,k}\},$$
$$j=1,\ldots,n-1,\ i=0,\ldots,n-1-j,$$
$$q_{0,0}^{n,k} = \frac{\theta}{b-a}J(R_n^k(0^+)-z_n^k) + (\frac{\theta}{b-a})^2\{J^2A[\sum_{j=1}^{n-2}\sum_{i=0}^{n-2-j}\frac{i!}{[j(b-a)]^{i+1}}q_{i,j}^{n-1,k} - \sum_{j=-n+1}^{n-1}\frac{u_j^{n-1,k}}{b+j(b-a)}$$
$$-\sum_{j=-n+1}^{n-1}\frac{v_j^{n-1,k}}{-\eta-\theta+b+j(b-a)} + \sum_{j=0}^{n-2}\sum_{i=0}^{n-2-j}\frac{i!}{[(j+1)(a-b)]^{i+1}}m_{i,j}^{n-1,k}] - H^2A\sum_{i=0}^{n-2}i!(b-a)^{-(i+1)}q_{i,0}^{n-1,k}\},$$
$$q_{i,0}^{n,k} = (\frac{\theta}{b-a})^2[J^2A\frac{q_{i-1,0}^{n-1,k}}{i} - H^2A\sum_{s=i}^{n-2}\frac{s!}{i!}(b-a)^{i-s-1}q_{s,0}^{n-1,k}],\ i=1,2,\ldots,n-1,$$
$$q_{i,j}^{n,k} = (\frac{\theta}{b-a})^2\{-J^2A\sum_{s=i}^{n-2-j}\frac{s!}{i!}[j(b-a)]^{i-s-1}q_{s,j}^{n-1,k} - H^2A\sum_{s=i}^{n-2-j}\frac{s!}{i!}[(j+1)(b-a)]^{i-s-1}q_{s,j}^{n-1,k}\},$$
$$j=1,\ldots,n-1,\ i=0,\ldots,n-1-j.$$

Note that in the above recursions, $u_j^{n,k}$ and $v_j^{n,k}$, $n=0,1,\ldots,k$ exist only if $-n \le j \le n$. Otherwise, their values are defined to be 0. In the recursions for $m_{i,j}^{n,k}$ and $q_{i,j}^{n,k}$, if a sum interval does not exist, the corresponding sum is defined to be 0. Using the fact that $HJ = JH = 0$, it can be shown that (the detailed algebra is omitted for brevity)

$$D_0^k(t) = u_0^{0,k} + v_0^{0,k}e^{-(\eta+\theta)t} + m_{0,0}^{0,k}e^{-at} + q_{0,0}^{0,k}e^{-bt},$$
$$D_n^k(t) = \sum_{j=-n}^{n}u_j^{n,k}e^{j(a-b)t} + \sum_{j=-n}^{n}v_j^{n,k}e^{(-(\eta+\theta)+j(a-b))t} + \sum_{j=0}^{n-1}\sum_{i=0}^{n-1-j}m_{i,j}^{n,k}t^ie^{-(a+j(a-b))t}$$
$$+\sum_{j=0}^{n-1}\sum_{i=0}^{n-1-j}q_{i,j}^{n,k}t^ie^{-(b+j(b-a))t},\ n=1,2,\ldots,k.$$

Proof of Proposition 1:

Denote the length of an 'off' period by Y. Then Y has a phase-type distribution with Laplace-Stieltjes transform $\alpha(M-sI)^{-1}Me$ and mean $-\alpha M^{-1}e$. Conditioning on whether or not the service of a scheduled patient is interrupted for at least once, we have

$$\hat{X} = \begin{cases} \exp(\eta+\mu) & \text{w.p. } \frac{\mu}{\eta+\mu}, \\ \exp(\eta+\mu) + Y + \hat{X} & \text{w.p. } \frac{\eta}{\eta+\mu}, \end{cases}$$

which yields $E(\hat{X}) = \frac{1}{\mu}(1 - \eta\alpha M^{-1}e)$. It can be shown by induction that $\alpha M^{-1}e = -(\frac{1}{\theta} + \frac{\eta}{\theta^2} + \ldots + \frac{\eta^{m-1}}{\theta^m})$. Hence $E(\hat{X}) = \frac{1}{\mu}(1 + \frac{\eta}{\theta} + \ldots + \frac{\eta^m}{\theta^m}) = \frac{\theta(1-(\frac{\eta}{\theta})^{m+1})}{\mu(\theta-\eta)}$.

Now, let $\tilde{G}(s) = E(e^{-s\hat{X}})$ denote the Laplace-Stieltjes transform of $\hat{X}$. Then we have $\tilde{G}(s) = \frac{\mu}{\eta+\mu}\frac{\eta+\mu}{s+\eta+\mu} + \frac{\eta}{\eta+\mu}\frac{\eta+\mu}{s+\eta+\mu}\alpha(M - sI)^{-1}Me\tilde{G}(s)$. When $m = 1$, the length of each 'off' period is exponentially distributed, therefore $\alpha(M - sI)^{-1}Me$ reduces to $\frac{\theta}{s+\theta}$. In this case, $\tilde{G}(s) = \frac{\mu(s+\theta)}{s^2+(\eta+\mu+\theta)s+\theta\mu} = \frac{\mu(s+\theta)}{(s+a)(s+b)}$, where a and b are given by Equations (2) and (3) respectively. Hence Equation (1) follows by inverting $\tilde{G}(s)$.

Appendix B: Explanation of the Error in Wang (1994)

In this section, we provide an explanation as to why the analysis in Wang (1994) is incorrect. Specifically, the problem has to do with Equation (5) on page 663 of the paper. On the right hand side of the equation, the first component of the vector reads $A_i(t_n)F_{in-1}(x_n)$ which implicitly assumes that the event that the server is operational (or not) at the release time of the nth job is independent of the event that the service of the $n-1$th job is finished before the release time of the nth job (the system is empty when the nth job is released), which is not the case. Below, we provide a more detailed explanation by using an example.

Suppose a server is operational at time 0 and two jobs are to be released and served by that server. The first job is released at time 0 and the second job is released at time t. The job processing time, the server operational time ('on' period), and the server repair time ('off' period) are independent and exponentially distributed with rates μ, η, and θ, respectively. First, we would like to compute $p_j(t)$, the probability that the service of the first job finishes before t, the release time of the second job, and the server is in state j at t, where $j = 0$ indicates that the server is operational and $j = 1$ indicates that the server is in repair.

Denote $r_{ij}(k,t)$, $k = 0, 1, \ldots$, to be the probability that the services of exactly k jobs finish during $(0,t]$ and the server is in state j at time t given that the server is in state i at time 0. Then $p_j(t) = \sum_{k=1}^{\infty} r_{0j}(k,t)$. Denote $r(k,t) = \begin{bmatrix} r_{0,0}(k,t) & r_{0,1}(k,t) \\ r_{1,0}(k,t) & r_{1,1}(k,t) \end{bmatrix}$. Similar to the differential equations derived in Theorem 1, we can show the following:

$$\frac{\mathrm{d}r(0,t)}{\mathrm{d}t} = \begin{bmatrix} -(\eta+\mu) & \eta \\ \theta & -\theta \end{bmatrix} r(0,t), \quad \frac{\mathrm{d}r(k,t)}{\mathrm{d}t} = \begin{bmatrix} \mu & 0 \\ 0 & 0 \end{bmatrix} r(k-1,t) + \begin{bmatrix} -(\eta+\mu) & \eta \\ \theta & -\theta \end{bmatrix} r(k,t), \; k \geq 1.$$

Recall that $A = \begin{bmatrix} \mu & 0 \\ 0 & 0 \end{bmatrix}$, and $B = \begin{bmatrix} -(\eta+\mu) & \eta \\ \theta & -\theta \end{bmatrix}$. Taking the LT on both sides of the above differential equations, we have $sr^*(0,s) - I = Br^*(0,s)$, and $sr^*(k,s) = Ar^*(k-1,s) + Br^*(k,s)$, $k \geq 1$. Hence, $r^*(0,s) = (sI - B)^{-1}$, and $r^*(k,s) = [(sI - B)^{-1}A]^k(sI - B)^{-1}$, $k \geq 1$. Thus $\sum_{k=1}^{\infty} r^*(k,s) = \{[I - (sI - B)^{-1}A]^{-1} - I\}(sI - B)^{-1}$. Inverting it yields a 2×2 matrix, where the $(j+1)$th component in the first row is the desired $p_j(t)$, $j = 0, 1$.

On the other hand, from Equation (1), we know that the probability that the service for the first job finishes before t is $G(t) = \mu[\frac{\theta-a}{a(b-a)}(1 - e^{-at}) + \frac{b-\theta}{b(b-a)}(1 - e^{-bt})]$. Also, since the state of the server at time t follows a continuous-time Markov chain with state space $\{0, 1\}$, the probability that the server is operational at t is given by $\frac{\eta}{\eta+\theta}e^{-(\eta+\theta)t} + \frac{\theta}{\eta+\theta}$ (Kulkarni 1995, p. 260). Given any set of model parameters, one can easily check to see that $p_0(t) \neq G(t)[\frac{\eta}{\eta+\theta}e^{-(\eta+\theta)t} + \frac{\theta}{\eta+\theta}]$. For example, when $\mu = 1$, $\eta = 0.1$, $\theta = 0.5$, and $t = 1$, $p_0(t) = 0.5816$ while $G(t)[\frac{\eta}{\eta+\theta}e^{-(\eta+\theta)t} + \frac{\theta}{\eta+\theta}] = 0.5649$.

Thus, the event that the server is operational at the release time of the second job and the event that the service of the first job is finished before the release time of the second job (the system is empty when the second job is released) are not independent.

References

Kulkarni, V. G. 1995. *Modeling and analysis of stochastic systems*. Chapman & Hall, London, U.K..

Wang, P. P. 1994. Releasing N jobs to an unreliable machine. *Computers and Industrial Engineering* **26**(4) 661–671.