Basic Queueing Theory

M/M/-/- Type Queues

Kendall's Notation for Queues

A/B/C/D/E

Shorthand notation where *A*, *B*, *C*, *D*, *E* describe the queue

Applicable to a large number of simple queueing scenarios

Kendall’s Notation for Queues

A/B/C/D/E

A	Inter-arrival time distribution	
B	Service time distribution	
C	Number of servers	
D	Maximum number of jobs that can be there in the system (waiting and in service)	
	Default ¥ for infinite number of waiting positions	
E	Queueing Discipline (FCFS, LCFS, SIRO etc.)	
	Default is FCFS	

(A and B:)

M exponential
D deterministic
E_k Erlangian (order k)
G general

M/M/1 or M/M/1/∞ Single server queue with Poisson arrivals, exponentially distributed service times and infinite number of waiting positions

Little’s Result

$$N = l W \tag{2.9}$$

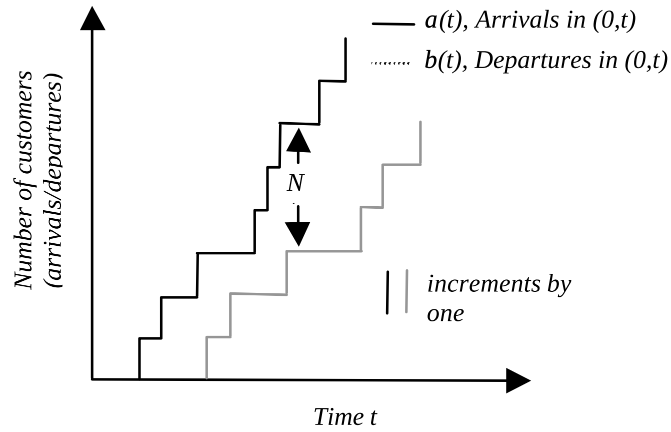
$$N_q = l W_q \tag{2.10}$$

Result holds in general for virtually all types of queueing situations where

l = Mean arrival rate of jobs that actually enter the system

Jobs blocked and refused entry into the system will not be counted in *l*

Little’s Result



Graphical Illustration/Verification of Little's Result

Little’s Result

Consider the time interval (*0*,*t*) where *t* is large, i.e. *t* ® ¥

Area(*t*) = area between *a(t)* and *b(t)* at time *t* = $\int_0^t [a(t) - b(t)]dt$

Average Time *W* spent in system = $\lim_{t\to\infty} \frac{Area(t)}{a(t)}$

Average Number *N* in system = $\lim_{t\to\infty} \frac{Area(t)}{t} = \lim_{t\to\infty} \frac{a(t)}{t} \frac{Area(t)}{a(t)}$

Since, $l = \lim_{t\to\infty} \frac{a(t)}{t}$ Therefore, $N = l W$

The PASTA Property

"Poisson Arrivals See Time Averages"

$p_k(t)$ =P{system is in state k at time t }

$q_k(t)$ =P{an arrival at time t finds the system in state k}

$N(t)$ be the actual number in the system at time t

$A(t, t+Dt)$ be the event of an arrival in the time interval $(t, t+Dt)$

Then

$$q_k(t) = \lim_{\Delta t \to 0} P\{N(t) = k \mid A(t, t+\Delta t)\}$$
$$= \lim_{\Delta t \to 0} \frac{P\{A(t, t+\Delta t \mid N(t) = k\}P\{N(t) = k\}}{P\{A(t, t+\Delta t\}} = p_k(t)$$

because P{$A(t, t+Dt)|N(t) = k$} = P{$A(t, t+Dt)$}

 7

Equilibrium Solutions for M/M/-/- Queues

Method 1: Obtain the differential-difference equations as in Section 1.2 or Section 2.2. Solve these under equilibrium conditions along with the normalization condition.

Method 2: Directly write the *flow balance equations* for proper choice of closed boundaries as illustrated in Section 2.2 and solve these along with the normalization condition.

Method 3: Identify the parameters of the birth-death Markov chain for the queue and directly use equations (2.7) and (2.8) as given in Section 2.2.

In the following, we have used this approach

 8

M/M/1 (or M/M/1/¥) Queue

$l_k = l \qquad \forall k$

$m_k = 0 \qquad k = 0$

$\quad = m \qquad k = 1, 2, 3, \ldots\ldots.$

For $r < 1$

$$p_k = p_0 \left(\frac{l}{m} \right)^k = p_0 r^k$$

$$p_0 = (1 - r)$$

$$N = \sum_{i=0}^{\infty} i p_i = \sum_{i=0}^{\infty} i r^i (1 - r) = \frac{r}{1 - r}$$

$$W = \frac{N}{l} = \frac{1}{m(1 - r)}$$

Using Little's Result

$$W_q = W - \frac{1}{m} = \frac{r}{m(1 - r)}$$

$$N_q = l W_q = \frac{r^2}{(1 - r)}$$

Using Little's Result

 9

M/M/1/∞ Queue with Discouraged Arrivals

$l_k = \frac{l}{k+1} \qquad \forall k$

$m_k = 0 \qquad k = 0$

$\quad = m \qquad k = 1, 2, 3, \ldots\ldots.$

For $r = l/m <$ ¥

$$p_k = p_0 \prod_{i=0}^{k-1} \frac{l}{m(i+1)} = p_0 \left(\frac{l}{m} \right)^k \frac{1}{k!} \qquad (2.14)$$

$$p_0 = \exp(-\frac{l}{m}) \qquad (2.15)$$

$$N = \sum_{k=0}^{\infty} k p_k = \frac{l}{m}$$

$$l_{eff} = \sum_{k=0}^{\infty} l_k p_k = m \left[1 - \exp(-\frac{l}{m}) \right]$$

Effective Arrival Rate

$$W = \frac{N}{l_{eff}} = \frac{l}{m^2 \left[1 - \exp(-\frac{l}{m}) \right]}$$

Little's Result

 10

M/M/1/∞ Queue with Discouraged Arrivals

In this case, *PASTA is not applicable* as the overall arrival process is not Poisson

p_r = P{arriving customer sees r in system (before joining the system)}

$P\{E_i\} = p_i = e^{-l/m}\left(\frac{l}{m}\right)^i \frac{1}{i!}$

DE be the event of an arrival in *(t, t+Dt)*

E_i is the event of the system being in state i

$P\{\Delta E \mid E_i\} = \frac{l\Delta t}{i+1}$

$$p_r = P\{E_r \mid \Delta E\} = \frac{P\{E_r\}P\{\Delta E \mid E_r\}}{P\{\Delta E\}} = \frac{P\{E_r\}P\{\Delta E \mid E_r\}}{\sum_{i=0}^{\infty} P\{E_i\}P\{\Delta E \mid E_i\}}$$

$$p_r = \left(\frac{l}{m}\right)^{r+1} \frac{1}{(r+1)!}\left(\frac{e^{-l/m}}{1-e^{-l/m}}\right) \qquad W = \sum_{k=0}^{\infty} \frac{k+1}{m} p_k = \frac{l}{m^2(1-e^{-l/m})}$$

M/M/m/∞ Queue (*m* servers, infinite number of waiting positions)

$$l_k = l \quad \forall k \qquad m_k = km \quad 0 \le k \le (m-1)$$
$$= mm \quad k \ge m$$

For $r = l/m < m$

$$p_k = p_0 \frac{r^k}{k!} \quad for \quad k \le m$$
$$= p_0 \frac{r^k}{m!m^{k-m}} \quad for \quad k > m \qquad (2.16)$$

Erlang's C-Formula

$$p_0 = \left(\sum_{k=0}^{m-1} \frac{r^k}{k!} + \frac{mr^m}{m!(m-r)}\right)^{-1} \qquad (2.17)$$

$$P\{queueing\} = \sum_{k=m}^{\infty} p_k = C(m, r) = p_0 \frac{mr^m}{m!(m-r)} \qquad (2.18)$$

M/M/m/m Queue (m server loss system, no waiting)

$$l_k = l \qquad k < m$$
$$= 0 \qquad \textit{otherwise} \quad (\textit{Blocking or Loss Condition})$$

$$m_k = km \qquad 0 \le k \le m$$
$$= 0 \qquad \textit{otherwise}$$

For

$$r = \frac{l}{m} < \infty$$

$$p_k = p_0 \frac{r^k}{k!} \qquad \textit{for} \quad k \le m$$
$$= 0 \qquad \textit{otherwise} \tag{2.19}$$

$$p_0 = \frac{1}{\sum_{k=0}^{m} \frac{r^k}{k!}} \tag{2.20}$$

 13

M/M/m/m Queue (m server loss system, no waiting)

Simple model for a telephone exchange where a line is given only if one is available; otherwise the call is lost

Blocking Probability B(m,r) = P{an arrival finds all servers busy and leaves without service}

$$B(m,r) = p_0 \frac{r^m}{m!} \qquad \text{Erlang's B-Formula} \tag{2.21}$$

$$B(0,r) = 1 \qquad B(m,r) = \frac{\frac{rB(m-1,r)}{m}}{1 + \frac{rB(m-1,r)}{m}} \tag{2.22}$$

 14

M/M/1/K Queue (single server queue with K-1 waiting positions)

$$l_k = l \qquad k < K$$
$$= 0 \qquad otherwise \quad (Blocking \;\; or \; Loss \; Condition)$$
$$m_k = m \qquad k \le K$$
$$= 0 \qquad otherwise$$

For

$$r = \frac{l}{m} < \infty$$

$$\begin{cases} p_k = p_0 r^k & for \quad k \le K \qquad (2.23) \\ \quad = 0 & otherwise \\ p_0 = \dfrac{(1-r)}{(1-r^{K+1})} & \qquad (2.24) \end{cases}$$

 15

M/M/1/-/K Queue *(*single server, infinite number of waiting positions, finite customer population *K*)

$$l_k = l\,(K-k) \qquad k < K$$
$$= 0 \qquad otherwise \quad (Blocking \;\; or \; Loss \; Condition)$$
$$m_k = m \qquad k \le K$$
$$= 0 \qquad otherwise$$

For

$$p_k = p_0 r^k \frac{K!}{(K-k)!} \qquad k=1,\ldots..,K \qquad (2.25)$$

$$r = \frac{l}{m} < \infty$$

$$p_0 = \frac{1}{\sum_{k=0}^{K} r^k \dfrac{K!}{(K-k)!}} \qquad (2.26)$$

 16

Delay Analysis for a FCFS M/M/1/∞ Queue
(Section 2.6.1)

Q: Queueing Delay (not counting service time for an arrival

pdf $f_Q(t)$, cdf $F_Q(t)$, $L_Q(s) = LT(f_Q(t)\}$

W=Q+T

W: Total Delay (waiting time and service time) for an arrival

pdf $f_W(t)$, cdf $F_W(t)$, $L_W(s) = LT(f_W(t)\}$

T: Service Time

$$f_T(t) = me^{-mt} \quad F_T(t) = e^{-mt} \quad L_T(s) = \frac{m}{(s+m)}$$

Since
$Q \wedge T$

$$L_W(s) = \frac{m}{(s+m)} L_Q(s) \quad f_W(t) = f_Q(t) * [me^{-mt}] \quad (2.30)$$

Knowing the distribution of either *W* or *Q*, the distribution of the other may be found

For a particular arrival of interest -

$F_Q(t)$ = P{queueing delay ≤ t}
= P{queueing time=0} + [$\Sigma_{n^3 1}$ P{queueing time≤ t | arrival found n jobs in system}]p_n

Erlang-n distribution for sum of *n* exponential r.v.s

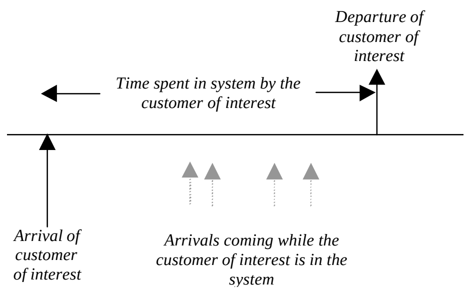
$$F_Q(t) = (1-r) + (1-r)\sum_{n=1}^{\infty} r^n \int_{x=0}^{t} \frac{m(mx)^{n-1}}{(n-1)!} e^{-mx} dx$$

$$= (1-r) + (1-r) r \int_0^t me^{-mx} \sum_{n=1}^{\infty} \frac{(mxr)^{n-1}}{(n-1)!} dx \quad (2.31)$$

$$= (1-r) + (1-r) r \int_0^t me^{-mx(1-r)} dx = (1-r) + r(1-e^{-mt(1-r)})$$

$$f_Q(t) = \frac{dF_Q(t)}{dt} = d(t)(1-r) + l(1-r)e^{-mt(1-r)} \quad (2.32)$$

$$f_W(t) = (1-r)me^{-mt} + l(1-r)m \int_0^t e^{-m(1-r)(t-x)} e^{-mx} dx = (m-l)e^{-(m-l)t}$$



• PASTA applicable to this queue

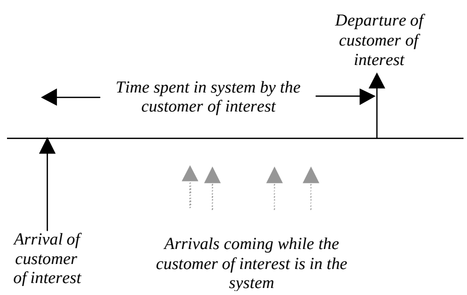
• N and N_Q seen by an arrival same as the time-averaged values

Arrival/Departure of Customer/Job of Interest from a FCFS M/M/1Queue

Let

N^* = Number in the system that a job will see left behind when it departs

$p_n^* = P\{N^*=n\}$ for $N^*=0, 1, \ldots, \infty$



For a FCFS queue, number left behind by a job will be equal to the number arriving while it is in the system.

$$G^*(z) = \sum_{n=0}^{\infty} z^n p_n^* = \sum_{n=0}^{\infty} z^n \int_{t=0}^{\infty} \frac{(\lambda t)^n}{n!} e^{-\lambda t} f_W(t)dt \qquad (2.36)$$

$$= \int_0^{\infty} e^{-\lambda t(1-z)} f_W(t)dt = L_W(\lambda - \lambda z)$$

$$E\{N^*\} = \left.\frac{dG^*(z)}{dz}\right|_{z=1} = -\lambda \left.\frac{dL_W(s)}{ds}\right|_{s=0} = \lambda W \qquad (2.37)$$

An important general observation can also be made along the lines of Eq. (2.36).

Consider the number arriving from a Poisson process with rate l in a random time interval T where $L_T(s)=LT\{f_T(t)\}$. The generating function $G(z)$ of this will be given by

$$G(z) = L_T(l - lz)$$

and the mean number will be $E\{N\} = l\,E\{T\}$

This result will be found to be useful in various places in our subsequent analysis.

Delay Analysis for the FCFS M/M/m/∝ Queue
(Section 2.6.2)

Using an approach similar to that used for the M/M/1 queue, we obtain the following

$$f_Q(t) = \left\{1 - p_0\left[\frac{m r^m}{m!(m-r)}\right]\right\} d(t) + \left[\frac{m p_0 r^m e^{-m(m-r)t}}{(m-1)!}\right] u(t) \qquad (2.34)$$

$$f_W(t) = \left\{1 - p_0\left[\frac{m r^m}{m!(m-r)}\right]\right\} m e^{-mt} - \left[\frac{m p_0 r^m [e^{-m(m-r)t} - e^{-mt}]}{(m-1)!(1-m-r)}\right] \qquad (2.35)$$

See Section 2.6.2 for the details and the intermediate steps