Queueing Theory

There are many situations in daily life when a queue is formed. For example, machines waiting to be repaired, patients waiting in a Doctor's room, cars waiting at a traffic signal and passengers waiting to buy tickets in counters form queues. Queue is formed if the service required by the customer (machine, patient, car, etc.) is not immediately available, that is, if the current demand for a particular service exceeds the capacity to provide the service.

Queues may be decreased in size or prevented from forming by providing additional service facilities which results in a drop in the profit. On the other hand, excessively long queues may result in lost sales and lost customers. Hence the problem of interest is how to achieve a balance between the cost associated with long waiting (queues) and the cost associated with the prevention of waiting in order to maximise the profits. As queueing theory provides an answer to this problem, it has become a topic of interest. Before we consider the solutions of queueing problems, we shall consider the general framework of a queueing system.

Although there are many types of queueing systems, all of them can be classified and described according to the following characteristics:

1. The input (or arrival) pattern The input describes the manner in which the customers arrive and join the queueing system. It is not possible to observe and control the actual moment of arrival of a customer for service. Hence the number of arrivals in one time period or the interval between successive arrivals is not treated as a constant, but a random variable. So the mode of arrival of customers is expressed by means of the probability distribution of the number of arrivals per unit of time or of the inter-arrival time.

We shall mostly deal with only those queueing systems in which the number of arrivals per unit of time has a poisson distribution with mean λ. In this case, the time between consecutive arrivals has an exponential distribution with mean $\frac{1}{\lambda}$ [Refer to Property 4 of the poisson process discussed in the previous Chapter 7]

Further the input process should specify the number of queues that are permitted to form, the maximum queue length and the maximum number of customers requiring service, viz., the nature of the source (finite or infinite) from which the customers emanate.

2. *The service mechanism (or pattern)* The mode of service is represented by means of the probability distribution of the number of customers serviced per unit of time or of the inter-service time. We shall mostly deal with only those queueing systems in which the number of customers serviced per unit of time has a Poisson distribution with mean μ or equivalently the inter-service time (viz. the time to complete the service for a customer) has an exponential distribution with mean $\frac{1}{\mu}$.

Further the service process should specify the number of servers and the arrangement of servers (in parallel, in series, etc.), as the behaviour of the queueing system depends on them also. The following figures represent the framework of queueing systems in which only one queue is permitted to form:

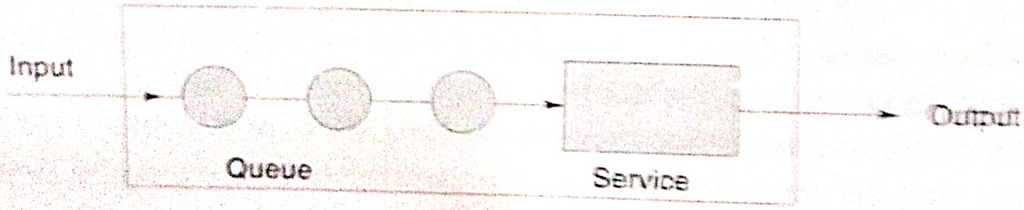

Fig. 8.1 Single server queueing system

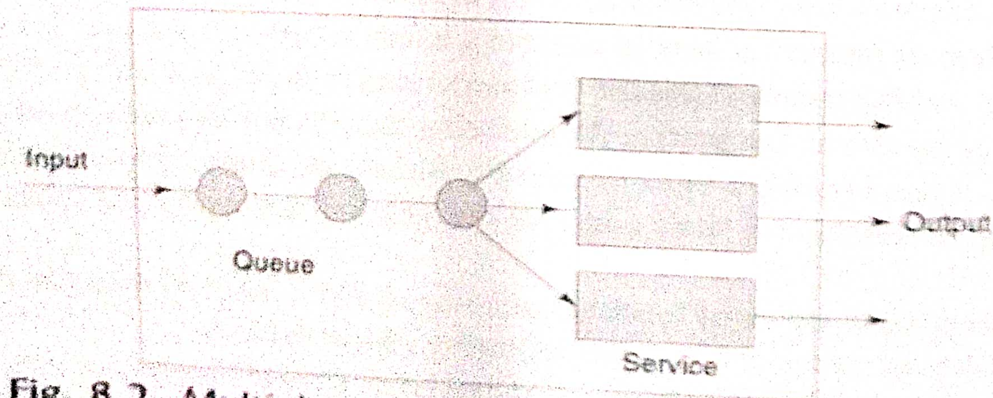

Fig. 8.2 Multiple servers (in parallel) queueing system

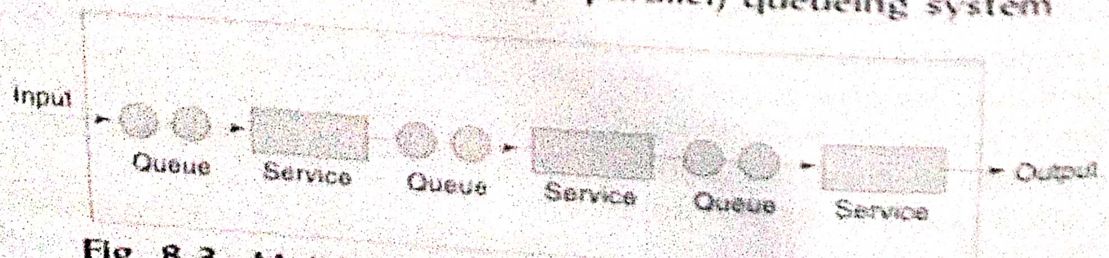

Fig. 8.3 Multiple servers (in series) queueing system

3. The queue discipline The queue discipline specifies the manner in which the customers form the queue or equivalently the manner in which they are selected for service, when a queue has been formed. The most common discipline is the *FCFS* (First Come First Served) or *FIFO* (First in First out) as per which the customers are served in the strict order of their arrival. If the last arrival in the system is served first, we have the *LCFS* or *LIFO* (last in First Out) discipline. If the service is given in random order, we have the *SIRO* discipline. In the queueing systems which we deal with, we shall assume that service is provided on the *FCFS* (First come First served) basis.

Symbolic Representation of a Queueing Model

Usually a queueing model is specified and represented symbolically in the form $(a/b/c)$:(d/e), where a denotes the type of distribution of the number of arrivals per unit time, b the type of distribution of the service time, c the number of servers, d the capacity of the system, viz., the maximum queue size and e the queue discipline.

Accordingly, the first four models which we will deal with will be denoted by the symbols $(M/M/1)$:$(\infty/FIFO)$, $(M/M/s)$: $(\infty/FIFO)$, $(M/M/1)$: $(k/FIFO)$ and $(M/M/s)$: $(k/FIFO)$.

In the above symbols the letter 'M' stands for Markov' indicating that the number of arrivals in time t and the number of completed services in time t follow Poisson process which is a continuous time Markov chain.

Difference Equations Related to Poisson Queue Systems

If the characteristics of a queueing system (such as the input and output parameters) are independent of time or equivalently if the behaviour of the system is independent of time, the system is said to be in *steady-state*. Otherwise it is said to be in *transient-state*.

Let $P_n(t)$ be the probability that there are n customers in the system at time t $(n > 0)$. Let us first derive the differential equation satisfied by $P_n(t)$ and then deduce the difference equation satisfied by P_n(probability of n customers at any time) in the steady-state.

Let λ_n be the average arrival rate when there are n customers in the system (both waiting in the queue and being served) and let μ_n be the average service rate when there are n customers in the system.

Note *The system being in steady-state does not mean that the arrival rate and service rate are independent of the number of customers in the system.*

Now $P_n(t + \Delta T)$ is the probability of n customers at time $t + \Delta T$.

The presence of n customers in the system at time $t + \Delta T$ can happen in any one of the following four mutually exclusive ways:

(i) Presence of n customers at t and no arrival or departure during Δt time.

(ii) Presence of $(n - 1)$ customers at t and one arrival and no departure during Δt time.

(iii) Presence of $(n+1)$ customers at t and no arrival and one departure during Δt time.

(iv) Presence of n customers at t and one arrival and one departure during Δt time (since more than one arrival/departure during Δt is ruled out)

$$\therefore \quad P_n(t+\Delta t) = P_n(t)\,(1-\lambda_n \Delta_t)\,(1-\mu_n \Delta t) + P_{n-1}(t)\,\lambda_{n-1}\Delta t\,(1-\mu_{n-1}\Delta t) + P_{n+1}(t)\,(1-\lambda_{n+1}\Delta t)\,\mu_{n+1}\Delta t + P_n(t)\cdot\lambda_n\Delta t\cdot\mu_n\Delta t$$

[since P(an arrival occurs during Δt time) $= \lambda\,\Delta t$ etc.]

i.e., $P_n(t+\Delta t) = P_n(t) - (\lambda_n+\mu_n)\,P_n(t)\,\Delta t + \lambda_{n-1}\,P_{n-1}(t)\,\Delta t + \mu_{n+1}\,P_{n+1}(t)$, on omitting terms containing $(\Delta t)^2$ which is negligibly small.

$$\therefore \quad \frac{P_n(t+\Delta t)-P_n(t)}{\Delta t} = \lambda_{n-1}P_{n-1}(t) - (\lambda_n+\mu_n)\,P_n(t) + \mu_{n+1}P_{n+1}(t) \qquad (1)$$

Taking limits on both sides of (1) as $\Delta t \to 0$, we have

$$P_n'(t) = \lambda_{n-1}P_{n-1}(t) - (\lambda_n+\mu_n)\,P_n(t) + \mu_{n+1}P_{n+1}(t) \qquad (2)$$

Equation (2) does not hold good for $n=0$, as $P_{n-1}(t)$ does not exist. Hence we derive the differential equation satisfied by $P_0(t)$ independently. Proceeding as before,

$$P_0(t+\Delta t) = P_0(t)\,(1-\lambda_0\Delta t) + P_1(t)\,(1-\lambda_1\Delta t)\,\mu_1\,\Delta t.$$

[by the possibilities (i) and (iii) given above and as no departure is possible when $n=0$]

$$\therefore \quad \frac{P_0(t+\Delta t)-P_0(t)}{\Delta t} = -\lambda_0 P_0(t) + \mu_1 P_1(t) \qquad (3)$$

Taking limits on both sides of (3) as $\Delta t \to 0$, we have

$$P_0'(t) = -\lambda_0 P_0(t) + \mu_1 P_1(t) \qquad (4)$$

Now in the steady-state, $P_n(t)$ and $P_0(t)$ are independent of time and hence $P_n'(t)$ and $P_0'(t)$ become zero. Hence the differential equations (2) and (4) reduce to the difference equations

$$\lambda_{n-1}P_{n-1} - (\lambda_n+\mu_n)\,P_n + \mu_{n+1}P_{n+1} = 0 \qquad (5)$$

and

$$-\lambda_0 P_0 + \mu_1 P_1 = 0 \qquad (6)$$

Values of P_0 and P_n for Poisson Queue Systems

From Equation (6) derived above, we have

$$P_1 = \frac{\lambda_0}{\mu_1}P_0 \qquad (7)$$

Putting $n=1$ in (5) and using (7), we have

$$\mu_2 P_2 = (\lambda_1+\mu_1)\,P_1 - \lambda_0 P_0$$

$$= (\lambda_1+\mu_1)\,\frac{\lambda_0}{\mu_1}P_0 - \lambda_0 P_0 = \frac{\lambda_0\lambda_1}{\mu_1}P_0$$

$$\therefore \qquad P_2 = \frac{\lambda_0 \lambda_1}{\mu_1 \mu_2} P_0 \tag{8}$$

Successively putting $n = 2, 3, \ldots$ in (5) and proceeding similarly, we can get

$$P_3 = \frac{\lambda_0 \lambda_1 \lambda_2}{\mu_1 \mu_2 \mu_3} P_0 \text{ etc.}$$

Finally
$$P_n = \frac{\lambda_0 \lambda_1 \lambda_2 \cdots \lambda_{n-1}}{\mu_1 \mu_2 \mu_3 \cdots \mu_n} . P_0, \text{ for } n = 1, 2, \ldots \tag{9}$$

Since the number of customers in the system can be 0 or 1 or 2 or 3 etc., which events are mutually exclusive and exhaustive, we have $\sum\limits_{n=0}^{\infty} P_n = 1$.

i.e.,
$$P_0 + \sum_{n=1}^{\infty} \left(\frac{\lambda_0 \lambda_1 \cdots \lambda_{n-1}}{\mu_1 \mu_2 \cdots \mu_n} \right) P_0 = 1$$

$$\therefore \qquad P_0 = \frac{1}{1 + \sum\limits_{n=1}^{\infty} \left(\frac{\lambda_0 \lambda_1 \cdots \lambda_{n-1}}{\mu_1 \mu_2 \cdots \mu_n} \right)} \tag{10}$$

Equations (9) and (10) will be used to derive the important characteristics of the four queueing models.

Characteristics of Infinite Capacity, Single Server Poisson Queue Model I [M/M/1): (∞/FIFO) model], when $\lambda_n = \lambda$ and $\mu_n = \mu$ $(\lambda < \mu)$

1. *Average number L_s of customers in the system:* Let N denote the number of customers in the queueing system (i.e., those in the queue and the one who is being served).

   N is a discrete random variable, which can take the values $0, 1, 2, \ldots, \infty$ such that $P(N = n) = P_n = \left(\frac{\lambda}{\mu}\right)^n P_0$, from Equation (9) of the previous discussion.

   From Equation (10), we have

$$P_0 = \frac{1}{1 + \sum\limits_{n=1}^{\infty} \left(\frac{\lambda}{\mu}\right)^n} = \frac{1}{\sum\limits_{n=0}^{\infty} \left(\frac{\lambda}{\mu}\right)^n} = 1 - \frac{\lambda}{\mu}$$

$$\therefore \qquad P_n = \left(\frac{\lambda}{\mu}\right)^n \left(1 - \frac{\lambda}{\mu}\right)$$

$$\text{Now } L_s = E(N) = \sum_{n=0}^{\infty} n \times P_n$$

$$= \left(\frac{\lambda}{\mu}\right)\left(1-\frac{\lambda}{\mu}\right)\sum_{n=1}^{\infty} n\left(\frac{\lambda}{\mu}\right)^{n-1}$$

$$= \frac{\lambda}{\mu}\left(1-\frac{\lambda}{\mu}\right)\left(1-\frac{\lambda}{\mu}\right)^{-2}, \text{ by binomial summation}$$

$$= \frac{\dfrac{\lambda}{\mu}}{1-\dfrac{\lambda}{\mu}} = \frac{\lambda}{\mu-\lambda} \qquad (1)$$

2. *Average number L_q of customers in the queue or Average length of the queue:*
If N is the number of customers in the system, then the number of customers in the queue is $(N-1)$

$$\therefore L_q = E(N-1) = \sum_{n=1}^{\infty} (n-1)P_n$$

$$= \left(1-\frac{\lambda}{\mu}\right)\sum_{n=1}^{\infty} (n-1)\left(\frac{\lambda}{\mu}\right)^{n}$$

$$= \left(\frac{\lambda}{\mu}\right)^{2}\left(1-\frac{\lambda}{\mu}\right)\sum_{n=2}^{\infty} (n-1)\left(\frac{\lambda}{\mu}\right)^{n-2}$$

$$= \left(\frac{\lambda}{\mu}\right)^{2}\left(1-\frac{\lambda}{\mu}\right)\left(1-\frac{\lambda}{\mu}\right)^{-2}$$

$$= \frac{\left(\dfrac{\lambda}{\mu}\right)^2}{1-\dfrac{\lambda}{\mu}} = \frac{\lambda^2}{\mu(\mu-\lambda)} \qquad (2)$$

3. *Average number L_w of customers in nonempty queues*

$L_w = E\{(N-1)/(N-1) > 0\}$, since the queue is non-empty

$$= \frac{E(N-1)}{P(N-1>0)} = \frac{\lambda^2}{\mu(\mu-\lambda)} \times \frac{1}{\sum_{n=2}^{\infty} P_n}$$

$$= \frac{\lambda^2}{\mu(\mu-\lambda)} \times \frac{1}{\sum_{n=2}^{\infty}\left(\frac{\lambda}{\mu}\right)^n\left(1-\frac{\lambda}{\mu}\right)}$$

$$= \frac{\lambda^2}{\mu(\mu-\lambda)} \times \frac{1}{\left(\frac{\lambda}{\mu}\right)^2\left(1-\frac{\lambda}{\mu}\right)\sum_{n=0}^{\infty}\left(\frac{\lambda}{\mu}\right)^n}$$

$$= \frac{\mu}{\mu-\lambda} \times \frac{1}{\left(1-\frac{\lambda}{\mu}\right)\left(1-\frac{\lambda}{\mu}\right)^{-1}} = \frac{\mu}{\mu-\lambda} \tag{3}$$

4. *Probability that the number of customers in the system exceeds k*

$$P(N > k) = \sum_{n=k+1}^{\infty} P_n = \sum_{n=k+1}^{\infty}\left(\frac{\lambda}{\mu}\right)^n\left(1-\frac{\lambda}{\mu}\right)$$

$$= \left(\frac{\lambda}{\mu}\right)^{k+1}\left(1-\frac{\lambda}{\mu}\right)\sum_{n=k+1}^{\infty}\left(\frac{\lambda}{\mu}\right)^{n-(k+1)}$$

$$= \left(\frac{\lambda}{\mu}\right)^{k+1}\cdot\left(1-\frac{\lambda}{\mu}\right)\sum_{n=0}^{\infty}\left(\frac{\lambda}{\mu}\right)^n$$

$$= \left(\frac{\lambda}{\mu}\right)^{k+1}\cdot\left(1-\frac{\lambda}{\mu}\right)\left(1-\frac{\lambda}{\mu}\right)^{-1} = \left(\frac{\lambda}{\mu}\right)^{k+1} \tag{4}$$

5. *Probability density function of the waiting time in the system*

Let W_s be the continuous random variable that represents the waiting time of a customer in the system, viz, the time between arrival and completion of service.

Let its pdf be $f(w)$ and let $f(w/n)$ be the density function of W_s subject to the condition that there are n customers in the queueing system when the customer arrives,

$$\text{Then } f(w) = \sum_{n=0}^{\infty} f(w/n)\, P_n \tag{5}$$

Now $f(w/n)$ = pdf of sum of $(n + 1)$ service times (one part-service time of the customer being served $+\, n$ complete service times)

= pdf of sum of $(n + 1)$ independent random variables, each of which is exponentially distributed with parameter μ

$= \frac{\mu^{n+1}}{n!}\, e^{-\mu\omega}\, w^n;\ w > 0$ which is the pdf of

Erlang distribution. [$\because$ The mgf of the exponential distribution (μ) is $\left(1-\frac{t}{\mu}\right)^{-1}$ and hence the mgf of the sum of $(n+1)$ independent exponential (μ) variables is $\left(1-\frac{t}{\mu}\right)^{n+1}$, which is the mgf of Erlang distribution with parameters μ and $(n+1)$] (refer to Erlang distribution in chapter 5)

$$\therefore \quad f(w) = \sum_{n=0}^{\infty} \frac{\mu^{n+1}}{n!} e^{-\mu w} w^n \left(\frac{\lambda}{\mu}\right)^n \left(1-\frac{\lambda}{\mu}\right), \text{ by (5)}$$

$$= \mu e^{-\mu w}\left(1-\frac{\lambda}{\mu}\right) \sum_{n=0}^{\infty} \frac{1}{n!} (\lambda w)^n$$

$$= \mu\left(1-\frac{\lambda}{\mu}\right) e^{-\mu w} e^{\lambda w}, \text{ by exponential summation}$$

$$= (\mu-\lambda)\, e^{-(\mu-\lambda) w} \quad (6),$$ which is the pdf of an exponential distribution with parameter $(\mu-\lambda)$.

6. *Average waiting time of a customer in the system:*
   W_s follows an exponential distribution with parameter $(\mu-\lambda)$.

$$\therefore \quad E(W_s) = \frac{1}{\mu-\lambda} \tag{7}$$

($\because$ the mean of an exponential distribution is the reciprocal of its parameter).

7. *Probability that the waiting time of a customer in the system exceeds t*

$$P(W_s > t) = \int_t^{\infty} f(w)\, dw$$

$$= \int_t^{\infty} (\mu-\lambda)\, e^{-(\mu-\lambda) w}\, dw$$

$$= \left[-e^{-(\mu-\lambda) w}\right]_t^{\infty} = e^{-(\mu-\lambda) t} \tag{8}$$

8. *Probability density function of the waiting time Wq in the queue:*
   W_q represents the time between arrival and reach of service point.
   Let the pdf of W_q be $g(w)$ and let $g(w/n)$ be the density function of W_q subject to the condition that there are n customers in the system or there are $(n-1)$ customers in the queue apart from one customer receiving service. Now $g(w/n) =$ pdf of sum of n service times [one residual service time + $(n-1)$ full service times]

$$= \frac{\mu^n}{(n-1)!} e^{-\mu w} w^{n-1}; \; w > 0$$

$$\therefore \quad g(w) = \sum_{n=1}^{\infty} \frac{\mu^n}{(n-1)!} e^{-\mu w} w^{n-1} \left(\frac{\lambda}{\mu}\right)^n \left(1 - \frac{\lambda}{\mu}\right)$$

$$= \lambda\left(1 - \frac{\lambda}{\mu}\right) e^{-\mu w} \sum_{n=1}^{\infty} \frac{1}{(n-1)!} (\lambda w)^{n-1}$$

$$= \frac{\lambda}{\mu} (\mu - \lambda) e^{-\mu w} e^{\lambda w}$$

$$= \frac{\lambda}{\mu} (\mu - \lambda) e^{-(\mu - \lambda) w}; \ w > 0 \tag{9}$$

and $\quad g(w) = 1 - \dfrac{\lambda}{\mu}$, when $w = 0$

> **Note** *1. W_q is a continuous random variable in $w > 0$ and it takes the value 0 with a non-zero probability. 2. Wq does not follow an exponential distribution.*

9. *Average waiting time of a customer in the queue*

$$E(W_q) = \frac{\lambda}{\mu} (\mu - \lambda) \int_0^{\infty} w\, e^{-(\mu - \lambda)w}\, dw$$

$$= \frac{\lambda}{\mu} \int_0^{\infty} x\, e^{-x} \frac{dx}{\mu - \lambda}$$

$$= \frac{\lambda}{\mu(\mu - \lambda)} [x(-e^{-x}) - e^{-x}]_0^{\infty}$$

$$= \frac{\lambda}{\mu(\mu - \lambda)} \tag{10}$$

10. *Average waiting time of a customer in the queue, if he has to wait*

$$E(W_q/W_q > 0) = \frac{E(W_q)}{P(W_q > 0)}$$

$$= \frac{E(W_q)}{1 - P(W_q = 0)}$$

$$= \frac{E(W_q)}{1 - P\text{ (no customer in the queue)}}$$

$$= \frac{E(W_q)}{1 - P_0}$$

$$= \frac{\lambda}{\mu(\mu - \lambda)} \times \frac{\mu}{\lambda} \qquad \left(\because P_0 = 1 - \frac{\mu}{\lambda}\right)$$

$$= \frac{1}{\mu - \lambda} \qquad (11)$$

Relations Among $E(N_s)$, $E(N_q)$, $E(W_s)$ and $E(W_q)$

(i) $E(N_s) = \dfrac{\lambda}{\mu - \lambda} = \lambda E(W_s)$ $\quad [\because E(N_s) = L_s]$

(ii) $E(N_q) = \dfrac{\lambda^2}{\mu(\mu - \lambda)} = \lambda E(W_q)$ $\quad [\because E(N_q) = L_q]$

(iii) $E(W_s) = E(W_q) + \dfrac{1}{\mu}$

(iv) $E(N_s) = E(N_q) + \dfrac{\lambda}{\mu}$

Worked Example 8

Example 1

Arrivals at a telephone booth are considered to be Poisson with an average time of 12 min. between one arrival and the next. The length of a phone call is assumed to be distributed exponentially with mean 4 min.

(a) Find the average number of persons waiting in the system.

(b) What is the probability that a person arriving at the booth will have to wait in the queue?

(c) What is the probability that it will take him more than 10 min. altogether to wait for the phone and complete his call?

(d) Estimate the fraction of the day when the phone will be in use.

(e) The telephone department will install a second booth, when convinced that an arrival has to wait on the average for at least 3 min. for phone. By how much the flow of arrivals should increase in order to justify a second booth?

(f) What is the average length of the queue that forms from time to time?

$$\text{Mean inter-arrival time} = \frac{1}{\lambda} = 12 \text{ min.}$$

$$\text{Therefore mean arrival rate} = \lambda = \frac{1}{12} \text{ per minute.}$$

$$\text{Mean service time} = \frac{1}{\mu} = 4 \text{ min.}$$

$$\text{Therefore, mean service rate} = \mu = \frac{1}{4} \text{ per minute.}$$

(a) $E(N) = \dfrac{\lambda}{\mu - \lambda}$, (by formula (1) of model I)

$$= \frac{\frac{1}{12}}{\frac{1}{4} - \frac{1}{12}} = 0.5 \text{ customer}$$

(b) $P(W > 0) = 1 - P(W = 0)$

$= 1 - P$ (no customer in the system)

$= 1 - P_0$

$$= 1 - \left(1 - \frac{\lambda}{\mu}\right) \text{ (by the formula for } P_0 \text{ of model I)}$$

$$= \frac{\lambda}{\mu} = \frac{1/12}{1/4} = \frac{1}{3}$$

(c) $P(W > 10) = e^{-(\mu - \lambda) \times 10}$ [by formula (8) of model I]

$$= e^{-\left(\frac{1}{4} - \frac{1}{12}\right) \times 10}$$

$$= e^{-\frac{5}{3}} = 0.1889$$

(d) $P(\text{the phone will be idle}) = P(N = 0) = P_0$

$$= 1 - \frac{\lambda}{\mu} = \frac{2}{3}$$

$\therefore P$ (the phone will be in use) $= 1 - \frac{2}{3} = \frac{1}{3}$

or the fraction of the day when the phone will be in use $= \frac{1}{3}$.

(e) The second phone will be installed, if $E(W_q) > 3$.

i.e., if $\frac{\lambda}{\mu(\mu - \lambda)} > 3$ [by formula (10) of model I]

i.e., if $\frac{\lambda_R}{\frac{1}{4}\left(\frac{1}{4} - \lambda_R\right)} > 3$,

where λ_R is the required arrival rate.

i.e., if $\lambda_R > \frac{3}{4}\left(\frac{1}{4} - \lambda_R\right)$

i.e., if $\lambda_R > \frac{3}{28}$

Hence the arrival rate should increase by $\frac{3}{28} - \frac{1}{12} = \frac{1}{42}$ per minute, to justify a second phone.

(f) $E(N_q/\text{the queue is always available})$

$$= E(N_q/N_q > 0)$$

$$= E(N_q/N > 1)$$

$$= \frac{E(N_q)}{P(N > 1)} = \frac{E(N_q)}{1 - P_0 - P_1} = \frac{\lambda^2}{\mu(\mu - \lambda)} \times \frac{1}{1 - \left(1 + \frac{\lambda}{\mu}\right) P_0},$$

[by formula (2) of model I]

$$= \frac{\lambda^2}{\mu(\mu-\lambda)} \cdot \frac{1}{1-\left(1+\frac{\lambda}{\mu}\right)\left(1-\frac{\lambda}{\mu}\right)}$$

$$= \frac{\lambda^2}{\mu(\mu-\lambda)} \cdot \frac{\mu^2}{\lambda^2} = \frac{\mu}{\mu-\lambda} = \frac{1/4}{1/4-1/12} = 1.5 \text{ persons.}$$

Example 2

Customers arrive at a one-man barber shop according to a Poisson process with a mean interarrival time of 12 min. Customers spend an average of 10 min in the barber's chair.

(a) What is the expected number of customers in the barber shop and in the queue?

(b) Calculate the percentage of time an arrival can walk straight into the barber's chair without having to wait.

(c) How much time can a customer expect to spend in the barber's shop?

(d) Management will provide another chair and hire another barber, when a customer's waiting time in the shop exceeds 1.25 h How much must the average rate of arrivals increase to warrant a second barber?

(e) What is the average time customers spend in the queue?

(f) What is the probability that the waiting time in the system is greater than 30 min?

(g) Calculate the percentage of customers who have to wait prior to getting into the barber's chair.

(h) What is the probability that more than 3 customers are in the system?

$$\frac{1}{\lambda} = 12 \qquad \therefore \lambda = \frac{1}{12} \text{ per minute}$$

$$\frac{1}{\mu} = 10 \qquad \therefore \mu = \frac{1}{10} \text{ per minute}$$

(a) $E(N_s) = \dfrac{\lambda}{\mu-\lambda} = \dfrac{1/12}{1/10-1/12} = 5$ customers [by formula (1) of model I]

$E(N_q) = \dfrac{\lambda^2}{\mu(\mu-\lambda)}$ [by formula (2) of model I]

$$= \frac{\frac{1}{144}}{\frac{1}{10}\left(\frac{1}{10}-\frac{1}{12}\right)} = 4.17 \text{ customers}$$

(b) P(a customer straight goes to the barber's chair)

$= P$ (No customer in the system)

$$= P_0 = 1 - \frac{\lambda}{\mu} = 1 - \frac{\frac{1}{12}}{\frac{1}{10}} = \frac{1}{6}$$

Therefore, percentage of time an arrival need not wait = 16.7.

(c) $E(W) = \dfrac{1}{\mu - \lambda}$ [by formula (7) of model I]

$$= \frac{1}{\frac{1}{10} - \frac{1}{12}} = 60 \text{ min or } 1 \text{ h}$$

(d) $E(W) > 75$, if $\dfrac{1}{\mu - \lambda_r} > 75$

i.e., $\quad$ if $\lambda_r > \mu - \dfrac{1}{75}$

i.e., $\quad$ if $\lambda_r > \dfrac{1}{10} - \dfrac{1}{75}$

i.e., $\quad$ if $\lambda_r > \dfrac{13}{150}$

Hence, to warrant a second barber, the average arrival rate must increase by $\dfrac{13}{150} - \dfrac{1}{12} = \dfrac{1}{300}$ per minute.

(e) $E(W_q) = \dfrac{\lambda}{\mu(\mu - \lambda)}$

[by formula (10) of model I]

$$= \frac{\frac{1}{12}}{\frac{1}{10}\left(\frac{1}{10} - \frac{1}{12}\right)} = 50 \text{ min}$$

(f) $P(W > t) = e^{-(\mu - \lambda)t}$, [by formula (8) of model I]

$$\therefore \quad P(W > 30) = e^{-\left(\frac{1}{10} - \frac{1}{12}\right) \times 30}$$

$$= e^{-0.5} = 0.6065$$

(g) P(a customer has to wait) $= P(W > 0)$

$$= 1 - P(W = 0) = 1 - P(N = 0) = 1 - P_0$$

$$= \frac{\lambda}{\mu} = \frac{1/12}{1/10} = \frac{5}{6}$$

$\therefore$ Percentage of customers who have to wait

$$= \frac{5}{6} \times 100 = 83.33$$

(h) $P(N > 3) = P_4 + P_5 + P_6 + \ldots$

$$= 1 - \{P_0 + P_1 + P_2 + P_3\}$$

$$= 1 - \left(1 - \frac{\lambda}{\mu}\right)\left\{1 + \frac{\lambda}{\mu} + \left(\frac{\lambda}{\mu}\right)^2 + \left(\frac{\lambda}{\mu}\right)^3\right\}$$

[since $P_n = \left(\frac{\lambda}{\mu}\right)^n \left(1 - \frac{\lambda}{\mu}\right)$, for $n \geq 0$, for model I]

$$= \left(\frac{\lambda}{\mu}\right)^4 = \left(\frac{5}{6}\right)^4 = 0.4823$$