Queuing theory exercises

Giovanni Righini

April 1, 2022

These notes are mainly for the students of the "Logistics" course. Many exercises are taken from Hillier & Liebermann, "Introduction to Operations Research", chapters 15 and 16, and they are indicated by HL.

1 Applying Little's law

1.1 HL 15.2-2

You are given a two-server queuing system in a steady-state condition, where the number of customers in the system varies between 0 and 4. For $n = 0, \dots, 4$ the probability P_n that exactly n customers are in the system is $P_0 = \frac{1}{16}$, $P_1 = \frac{4}{16}$, $P_2 = \frac{6}{16}$, $P_3 = \frac{4}{16}$, $P_4 = \frac{1}{16}$.

- (a) Determine L, the expected number of customers in the system.
- (b) Determine L_q, the expected number of customers in the queue.
- (c) Determine the expected number of customers being served.
- (d) Given that the mean arrival rate is 2 customers per hour, determine the expected waiting time in the system, W, and the expected waiting time in queue, W_q.
- (e) Given that both servers have the same expected service time, use the results from part (d) to determine this expected service time.

Solution. The exercise requires to apply the definitions of L, L_q, W and W_q and the relations between them: $L = \lambda W$, $L_q = \lambda W_q$ and $W = W_q + 1/\mu$.

- (a) From the knowledge of P_n one can easily obtain L:

$$L = \sum_{n=0}^{4} P_n n = 0 + \frac{4}{16} + \frac{12}{16} + \frac{12}{16} + \frac{4}{16} = 2.$$

  In average there are two customers in the system.

- (b) From the probabilities P_n and from $s = 2$ one can directly obtain also L_q:

$$L_q = \sum_{n=2}^{4} P_n(n-2) = 0 + \frac{4}{16} + \frac{2}{16} = \frac{3}{8} = 0.375.$$

  In average, there are 0.375 customers waiting in queue.

- (c) From the difference $L - L_q$ we know that in average there are 1.625 customers that are receiving service; then one obtains that in average the 2 servers are busy for a fraction of time equal to $\rho = \frac{1.625}{2}$, i.e. $\rho = \frac{13}{16} = 0.8125$. The same result can be obtained directly from the values of P_n:

$$E[N.\ busy\ servers] = \sum_{n=0}^{4} P_n \min\{n, s\} = 0P_0 + 1P_1 + 2(P_2 + P_3 + P_4) = 0 + \frac{4}{16} + 2(\frac{6}{16} + \frac{4}{16} + \frac{1}{16}) = \frac{13}{8}.$$

- (d) From Little's law one obtains:

$$W = \frac{L}{\lambda} = \frac{2}{2} \frac{customers}{customers/hour} = 1\ hour.$$

  Analogously, for the queue one obtains:

$$W_q = \frac{L_q}{\lambda} = \frac{\frac{3}{8}}{2} \frac{customers}{customers/hour} = \frac{3}{16}\ hour.$$

- (e) Finally, from $W = W_q + \frac{1}{\mu}$ one obtains:

$$\mu = \frac{1}{W - W_q} = \frac{1}{1 - 3/16} = \frac{16}{13} \frac{customers}{hour}.$$

  Hence, the average service time is

$$\frac{1}{\mu} = \frac{13}{16} \frac{hour}{customer}.$$

2 Probability distributions and their properties

2.1 HL 15.4-1

Suppose that a queuing system has two servers, an exponential interarrival time distribution with a mean of 2 hours and an exponential service time distribution with a mean of 2 hours for each server. Furthermore, a customer has just arrived at 12:00 noon.

- (a) What is the probability that the next arrival will come (i) before 1:00 pm; (ii) between 1:00 and 2:00 pm; (iii) after 2:00 pm?
- (b) Suppose that no additional customers arrive before 1:00 pm. Now, what is the probability that the next arrival will come between 1:00 and 2:00 pm?
- (c) What is the probability that the number of arrivals between 1:00 and 2:00 pm will be (i) 0, (ii) 1, (iii) 2 or more?
- (d) Suppose that both servers are serving customers at 1:00 pm. What is the probability that neither customer will have service completed (i) before 2:00 pm, (ii) before 1:10 pm, (iii) before 1:01 pm?

Solution. The system is an $M/M/2$ system with $\lambda = 1/2$ and $\mu = 1/2$, both expressed in hours.

Question (a). Here T indicates the point in time of the next arrival starting from 12:00 (in hours). Using the definition of the exponential distribution,

$$P\{T \leq 1\} = \int_0^1 \frac{1}{2} e^{-\frac{1}{2}t} dt = | - e^{-\frac{1}{2}t}|_0^1 = 1 - e^{-\frac{1}{2}} \approx 0.393.$$

$$P\{1 \leq T \leq 2\} = P\{T \leq 2\} - P\{T \leq 1\} = | - e^{-\frac{1}{2}t}|_1^2 = e^{-\frac{1}{2}} - e^{-1} \approx 0.239.$$

$$P\{T \geq 2\} = 1 - P\{T \leq 2\} = 1 - (P\{T \leq 1\} + P\{1 \leq T \leq 2\}) = e^{-1} \approx 0.368.$$

Question (b). By the property of the exponential distribution of being "without memory",

$$P\{1 \leq T \leq 2 | T \geq 1\} = P\{T \leq 1\} \approx 0.393$$

as already computed in (a).

Question (c). Here N_t indicates the expected number of events between 1:00 pm and t (expressed in hours, starting from 1:00 pm). Using the definition of a Poisson process,

$$P\{N_1 = 0\} = \frac{(\frac{1}{2}1)^0 e^{\frac{1}{2}1}}{0!} = e^{-\frac{1}{2}} \approx 0.607.$$

The same result can be obtained in an easier way considering the results already computed in the previous answers:

$$P\{N_1 = 0\} = 1 - P\{T \leq 1\} = e^{-\frac{1}{2}} \approx 0.607.$$

Analogously,

$$P\{N_1 = 1\} = \frac{(\frac{1}{2}1)^1 e^{\frac{1}{2}1}}{1!} = \frac{1}{2} e^{-\frac{1}{2}} \approx 0.303.$$

$$P\{N_1 \geq 2\} = 1 - (P\{N_1 = 0\} + P\{N_1 = 1\}) = 1 - \frac{3}{2} e^{-\frac{1}{2}} \approx 0.090.$$

Question (d). Here N_t indicates the number of completions between 1:00 pm and t (expressed in hours, starting from 1:00 pm), while T indicates the time of the next completion.

$$P\{N_1 = 0\} = 1 - P\{T \leq 1\} = 1 - \int_0^1 (s\mu) e^{-(s\mu)t} dt = e^{-1}.$$

$$P\{N_{1/6} = 0\} = 1 - \int_0^{1/6} (s\mu) e^{-(s\mu)t} dt = e^{-\frac{1}{6}}.$$

$$P\{N_{1/60} = 0\} = 1 - \int_0^{1/60} (s\mu) e^{-(s\mu)t} dt = e^{-\frac{1}{60}}.$$

2.2 HL 15.4-2

The jobs to be performed on a particular machine arrive according to a Poisson input process with a mean rate of 2 per hour. Suppose that the machine breaks down and will require one hour to be repaired. What is the probability that the number of new jobs that will arrive during this time is (a) 0, (b) 2, (c) 5 or more?

Solution. Using the definition of a Poisson process,

$$P[X(t) = n] = \frac{(\lambda t)^n e^{\lambda t}}{n!}$$

with different values of n, with $t = 1$ *hour* and $\lambda = 2$ *job/hour*. Then one obtains

$$P_0 = \frac{e^{-2}}{1} = 0.1353$$

$$P_1 = \frac{2e^{-2}}{1} = 0.2707$$

$$P_2 = \frac{4e^{-2}}{2} = 0.2707$$

$$P_3 = \frac{8e^{-2}}{6} = 0.1804$$

$$P_4 = \frac{16e^{-2}}{24} = 0.0902$$

Finally

$$P(\geq 5) = 1 - \sum_{n=0}^{4} P_n = 1 - 7e^{-2} = 0.05265.$$

2.3 HL 15.4-3

The time required by a mechanic to repair a machine has an exponential distribution with a mean of 4 hours. However, a special tool would reduce this mean to 2 hours. If the mechanic repairs a machine in less than two hours he is paid \$ 100; otherwise, he is paid \$ 80. Determine the mechanic's expected increase in pay if he uses the special tool.

Solution. From the definition of exponential probability distribution, the probability that the repair intervention be completed in less than two hours is

$$P[\leq 2 \ hours] = \int_0^2 \mu e^{-\mu t} dt = 1 - e^{-2\mu}.$$

Hence, in the scenario with no special tool, where $\frac{1}{\mu} = 4$ *hour*,

$$P[\leq 2 \ hours] = 1 - e^{-\frac{1}{2}}$$

$$P[> 2 \ hours] = e^{-\frac{1}{2}}$$

$$E[payment] = \$[100(1 - e^{-\frac{1}{2}}) + 80(e^{-\frac{1}{2}})].$$

On the contrary, in the scenario with the special tool, where $\frac{1}{\mu} = 2$ *hours*,

$$P[\leq 2 \ hours] = 1 - e^{-1}$$

$$P[> 2 \ hours] = e^{-1}$$

$$E[payment] = \$[100(1 - e^{-1}) + 80(e^{-1})].$$

The expected difference in payment for each repair intervention is

$$\$[100(1 - e^{-1}) + 80(e^{-1})] - \$[100(1 - e^{-\frac{1}{2}}) + 80(e^{-\frac{1}{2}})] = \$4.773.$$

3 Birth-and-death processes

3.1 HL 15.5-5

A service station has one gasoline pump. Cars wanting gasoline arrive according to a Poisson process at a mean rate of 15 per hour. However, if the pump already is being used, these potential customers may balk (drive to another service station). In particular, if there are n cars already at the service station, the probability that an arriving potential customer will balk is $n/3$ for $n = 1, \ldots, 3$. The time required to service a car has an exponential distribution with a mean of 4 minutes.

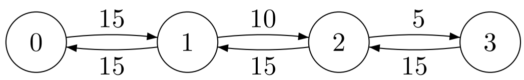
- (a) Construct the rate diagram for this queuing system.
- (b) Develop the balance equations.
- (c) Solve these equations to find the steady-state probability distribution of the number of cars at the station. Verify that this solution is the same as that given by the general solution for the birth-and-death process.
- (d) Find the expected waiting time (including service) for those cars that stay.

Solution. The system can be classified as $M/M/1$.

The value of $\lambda_0 = 15$ *customers/hour* is given. However, for different values of n, we have different values of λ_n owing to "balking" customers. Hence $\lambda_1 = 10$ *customers/hour*, $\lambda_2 = 5$ *customers/hour* and $\lambda_3 = 0$ *customers/hour*.

The state diagram has only 4 states, for $n = 0, \ldots, 3$. Transitions to states with $n > 3$ occur with zero average frequency.

The frequency of completions does not depend on the number of customers in the system, but it is always equal to $\mu = 15$ *customers/hour* in all states $n = 1, \ldots, 3$.



The balance equations for states $n = 0$, $n = 1$ and $n = 2$ are

$$\begin{cases} 15P_0 = 15P_1 \\ (10+15)P_1 = 15P_0 + 15P_2 \\ (15+5)P_2 = 10P_1 + 15P_3 \end{cases}$$

To these equations we must add the normalization constraint of the probability values:

$$P_0 + P_1 + P_2 + P_3 = 1.$$

We obtain a system of four linear equations in four variables, that can be solved expressing all probabilities as functions of P_0:

$$\begin{cases} P_1 = P_0 \\ P_2 = 2/3P_0 \\ P_3 = 2/9P_0 \end{cases}$$

from which

$$P_0 + P_0 + 2/3P_0 + 2/9P_0 = 1$$

i.e. $P_0 = 9/26$ and thus

$$\begin{cases} P_1 = 9/26 \\ P_2 = 6/26 \\ P_3 = 2/26 \end{cases}$$

Form the P values we can immediately obtain the expected number of customers in the system, indicated by L, and the average frequency of the arrivals, indicated by $\overline{\lambda}$.

$$L = \sum_{n=0}^{3} nP_n = 0 + 9/26 + 12/26 + 6/26 = 27/26 \textit{ customers}.$$

$$\overline{\lambda} = \sum_{n=0}^{3} \lambda_n P_n = 159/26 + 109/26 + 56/26 + 02/26 = 255/26 \text{ customers/hour}.$$

Then, through Little's law we can compute the required value of W:

$$W = L/\overline{\lambda} = 27/255 \text{ hour},$$

which is about 6 minutes.

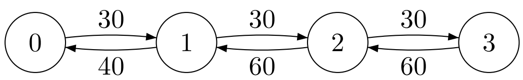
3.2 HL 15.5-9

A certain small grocery store has a single checkout stand with a full-time cashier. Customers arrive at the stand randomly (i.e. a Poisson input process) at a mean rate of 30 per hour. When there is only one customer at the stand, he is processed by the cashier alone, with an expected service time of 1.5 minutes. However, the stock boy has been given standard instructions that whenever there is more than one customer at the stand, he is to help the cashier by bagging the groceries. This help reduces the expected time required to process a customer to 1 minute. In both cases, the service-time distribution is exponential.

- (a) Construct the rate diagram for this queuing system.
- (b) What is the steady-state probability distribution of the number of customers at the checkout stand?
- (c) Derive L for this system (hint: refer to the derivation of L for the M/M/1 model at the beginning of Sec. 15.6). Use this information to determine L_q, W and W_q.

Solution. It is an $M/M/1$ system again.

The value of $\lambda = 30$ *customers/hour* is known. The completion frequency is $\mu_1 = 40$ *customers/hour* when $n = 1$ and $\mu_n = 60$ *customers/hour* when $n > 1$.



The balance equations form the following linear system:

$$\begin{cases} 30P_0 = 40P_1 \\ 70P_1 = 30P_0 + 60P_2 \\ 90P_i = 30P_{i-1} + 60P_{i+1} \quad \forall i \geq 2 \end{cases}$$

together with the normalization constraint $\sum_{n=0}^{\infty} P_n = 1$. Expressing the probability P_{i+1} as a function of P_{i-1} and P_i we get

$$\begin{cases} P_1 = 3/4P_0 \\ P_2 = \frac{70P_1 - 30P_0}{60} \rightarrow P_2 = 3/8P_0 \rightarrow P_2 = 1/2P_1 \\ P_{i+1} = \frac{90P_i - 30P_{i-1}}{60} \; \forall i \geq 2 \end{cases}$$

From the third equation we obtain $P_{i+1} = \frac{3P_i - P_{i-1}}{2}$. For $P_i = 1/2P_{i-1}$ we obtain $P_{i+1} = 1/2P_i$. Since for $i = 2$ we have $P_2 = 1/2P_1$, then $P_{i+1} = 1/2P_i \; \forall i \geq 1$. Hence, expressing now each P_i value as a function of P_0, we have

$$\begin{cases} P_1 = 3/4P_0 \\ P_2 = 3/8P_0 \\ P_3 = 3/16P_0 \\ \dots \\ P_i = 3/2^{i+1}P_0 \quad \forall i \end{cases}$$

The normalization constraint can be restated as

$$P_0(1 + \frac{3}{2}\sum_{n=1}^{\infty} \frac{1}{2^n}) = 1.$$

Since the missing term of the geometrical series (the term with $n = 0$), has valus equal to 1, we have

$$P_0(1 + \frac{3}{2}((\sum_{n=0}^{\infty} \frac{1}{2^n}) - 1)) = 1.$$

The sum of a geometrical series of ratio r is equal to $\frac{1}{1-r}$. Therefore for $r = 1/2$,

$$P_0(1 + \frac{3}{2}(\frac{1}{1 - \frac{1}{2}} - 1)) = 1,$$

$$P_0(1 + \frac{3}{2}(2 - 1)) = 1,$$

$$P_0(1 + \frac{3}{2}) = 1,$$

$$\frac{5}{2}P_0 = 1,$$

$$P_0 = \frac{2}{5}.$$

from which we can obtain al the other values of the probabilities: $P_1 = 3/10$, $P_2 = 3/20$ etc.

To obtain L we can proceed as follows. The definition of L is:

$$L = \sum_{n=0}^{\infty} nP_n$$

We have $P_n = \frac{3}{5}\frac{1}{2^n} \quad \forall n \geq 1$. The value of P_0 does not satisfy this relation but in the expression of L is it multiplied by 0. Then we can rewrite

$$L = \frac{3}{5}\sum_{n=0}^{\infty} n\frac{1}{2^n}.$$

To compute the sum of this series we can resort to the commutative property between the sum and the derivative. Let define $\phi = \frac{1}{2}$ and rewrite

$$L = \frac{3}{5}\sum_{n=0}^{\infty} n\phi^n \tag{1}$$

Now we have

$$\begin{array}{lll} & \sum_{n=0}^{\infty} n\phi^n & = \phi\sum_{n=0}^{\infty} n\phi^{n-1} = \\ = & \phi\sum_{n=0}^{\infty} \frac{d}{d\phi}(\phi^n) & = \phi\frac{d}{d\phi}(\sum_{n=0}^{\infty} \phi^n) = \\ = & \phi\frac{d}{d\phi}(\frac{1}{1-\phi}) & = \phi\frac{1}{(1-\phi)^2}. \end{array}$$

Remembering that $\phi = 1/2$ and replacing it in (1) we get

$$L = \frac{3}{5}\frac{\frac{1}{2}}{(1 - \frac{1}{2})^2} = \frac{3}{5}\frac{1/2}{1/4} = \frac{6}{5}.$$

From Little's law we obtain W (expressed in hours):

$$W = \frac{L}{\lambda} = \frac{\frac{6}{5}}{30} = \frac{1}{25}.$$

The length of the queue, L_q, can be computed from its definition

$$\begin{array}{ll} L_q = \sum_{n=1}^{\infty}(n-1)P_n & = \sum_{n=1}^{\infty} nP_n - \sum_{n=1}^{\infty} P_n = \\ = (\sum_{n=0}^{\infty} nP_n - 0P_0) - (\sum_{n=0}^{\infty} P_n - P_0) & = (L - 0) - (1 - P_0) = \\ = \frac{6}{5} - (1 - \frac{2}{5}) & = \frac{3}{5}. \end{array}$$

Finally, using Little's law again, we obtain W_q (expressed in hours):

$$W_q = \frac{W_q}{\lambda} = \frac{3/5}{30} = \frac{1}{50}.$$

3.3 HL 15.6-4

Jobs arrive at a particular work center according to a Poisson input process at a mean rate of 2 per day, and the operation time has an exponential distribution with a mean of 1/4 day. Enough in-process storage space is provided at the work center to accommodate 3 jobs in addition to the one being processed, whereas excess jobs excess are stored temporarily in a less convenient location. For what proportion of the time will this storage space at the work center be adequate to accommodate all waiting jobs?

Solution. It is an $M/M/1$ system. Although the description may induce to assume this is a finite capacity system, this is not the case, because the frequency of the arrivals is not zero for any given number of customers in the system.

The parameters describing the queuing system are $\lambda = 2$ and $\mu = 4$, both expressed in jobs per day. The system utilization factor is $\rho = \frac{1}{2}$. From the formulae for $M/M/1$ systems we obtain

$$P_0 = 1 - \rho = \frac{1}{2}.$$

$$P_n = (1-\rho)\rho^n = (\frac{1}{2})^{n+1}.$$

The available capacity is sufficient for a fraction of time equal to

$$\sum_{n=0}^{4} P_n = \sum_{n=0}^{4} (\frac{1}{2})^{n+1} = \frac{31}{32}.$$

3.4 HL 15.6-31

A company is designing a new production plant. A large number of identical machines have been assigned to a plant section and it is necessary to determine how many of them should be assigned to the same operator for loading, unloading, tuning, set-up and so on. For the purpose of this analysis the following data have been provided. The working time of a machine (i.e. the time from the completion of a service by the operator and the time when the machine requires his intervention again) has an exponential distribution with a mean value of 150 minutes. The service time has an exponential distribution with a mean value of 15 minutes. Each operator has its own subset of machines and he cannot give/receive help to/from other operators. For the section to reach the required level of productivity, the machines must be active for at least 89% of time in average.

- (a) What is the maximum number of machines that can be assigned to each operator?
- (b) Assuming that the number of machines per operator is the one determined in answer (a), for what fraction of time will the operators be busy?

Solution. This is an $M/M/1/N$ system with finite population. The value of N is a design variable.

Indicating with $T = \frac{5}{2}$ the average working time (expressed in hours), the constraint on productivity can be expressed as

$$T \geq \frac{89}{100}(T + W)$$

which is the same as

$$W \leq \frac{11}{89} T$$

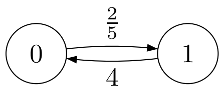
i.e.

$$W \leq 0.308.$$

The average frequency of arrivals (expressed in machines/hour) depends on the number n of idle machines and it is given by $\lambda_n = (N-n)\hat{\lambda}$, where $\hat{\lambda} = 1/T = 2/5$. The average frequency of completions (expressed in machines/hour) is $\mu = 4$.

Let examine the different cases.

$N = 1$. The state diagram is the following one.



The balance equation imposes that

$$4P_1 = 2/5P_0.$$

Since $P_0 + P_1 = 1$, we obtain

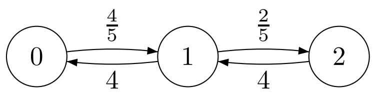
$$\begin{cases} P_0 = 10/11 \\ P_1 = 1/11 \end{cases}$$

Then we have $L = 0P_0 + 1P_1 = 1/11$. Furthermore we have $\overline{\lambda} = \frac{2}{5}P_0 + 0P_1 = \frac{2}{5}\frac{10}{11} + 0\frac{1}{11} = \frac{4}{11}$. Then, for Little's law,

$$W = \frac{L}{\overline{\lambda}} = \frac{1/11}{4/11} = \frac{1}{4} = 0.25.$$

The constraint on productivity is satisfied.

$N = 2$. The state diagram is the following one.



Balance equations impose that

$$\begin{cases} \frac{4}{5}P_0 = 4P_1 \\ (\frac{2}{5} + 4)P_1 = \frac{4}{5}P_0 + 4P_2 \end{cases}$$

from which we have

$$\begin{cases} P_1 = \frac{1}{5}P_0 \\ P_2 = \frac{1}{50}P_0 \end{cases}$$

Since $P_0 + P_1 + P_2 = 1$, we obtain $P_0(1 + \frac{1}{5} + \frac{1}{50}) = 1$ from which

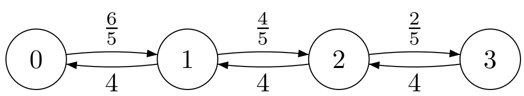
$$\begin{cases} P_0 = 50/61 \\ P_1 = 10/61 \\ P_2 = 1/61 \end{cases}$$

Then we have $L = 0P_0 + 1P_1 + 2P_2 = 12/61$. Furthermore, we have $\overline{\lambda} = \frac{4}{5}P_0 + \frac{2}{5}P_1 + 0P_2 = \frac{4}{5}\frac{50}{61} + \frac{2}{5}\frac{10}{61} = \frac{44}{61}$. Then, for Little's law,

$$W = \frac{L}{\overline{\lambda}} = \frac{12/61}{44/61} = \frac{12}{44} \approx 0.27.$$

The constraint of productivity is still satisfied.

$N = 3$. The state diagram is the following one.



The balance equations are

$$\begin{cases} \frac{6}{5}P_0 = 4P_1 \\ (\frac{4}{5}+4)P_1 = \frac{6}{5}P_0 + 4P_2 \\ (\frac{2}{5}+4)P_2 = \frac{4}{5}P_1 + 4P_3 \end{cases}$$

from which we have

$$\begin{cases} P_1 = \frac{3}{10}P_0 \\ P_2 = \frac{3}{50}P_0 \\ P_3 = \frac{3}{500}P_0 \end{cases}$$

Since $P_0 + P_1 + P_2 + P_3 = 1$, we have $P_0(1 + \frac{3}{10} + \frac{3}{50} + \frac{3}{500}) = 1$ from which

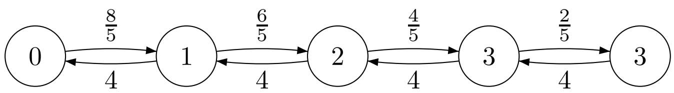
$$\begin{cases} P_0 = 500/683 \\ P_1 = 150/683 \\ P_2 = 30/683 \\ P_3 = 3/683 \end{cases}$$

Then we have $L = 0P_0 + 1P_1 + 2P_2 + 3P_3 = 0\frac{500}{683} + 1\frac{150}{683} + 2\frac{30}{683} + 3\frac{3}{683} = \frac{219}{683}$. Furthermore we have $\overline{\lambda} = \frac{6}{5}P_0 + \frac{4}{5}P_1 + \frac{2}{5}P_2 + 0P_3 = \frac{6}{5}\frac{500}{683} + \frac{4}{5}\frac{150}{683} + \frac{2}{5}\frac{30}{683} = \frac{600+120+12}{683} = \frac{732}{683}$. Then for Little's law,

$$W = \frac{L}{\overline{\lambda}} = \frac{219/683}{732/683} = \frac{219}{732} \approx 0.299.$$

The constraint of productivity is still satisfied.

$N = 4$. The state diagram is the following one.



With the same procedure above, we obtain $W \approx 0.33$ that violates the constraint.

Then, assuming $N = 3$, we can compute the system utilization factor

$$\rho = \frac{\overline{\lambda}}{\mu}.$$

With $N = 3$ we have $\overline{\lambda} = \frac{732}{683}$ and then

$$\rho = \frac{\frac{732}{683}}{4} = \frac{183}{683} \approx 26.8\%.$$

3.5 HL 15.7-4

An airline maintenance base has facilities for overhauling only one airplane engine at a time. Therefore, to return the airplanes to use as soon as possible, the policy has been to stagger the overhauling of the four engines of each airplane. In other words, only one engine is overhauled each time an airplane comes into the shop. Under this policy, airplanes have arrived according to a Poisson process at a mean rate of 1 per day. The time required for an engine overhaul (once work has begun) has an exponential distribution with a mean of $\frac{1}{2}$ day. A proposal has been made to change the policy so that all four engines are overhauled consecutively each time an airplane comes into the shop. Although this would quadruple the expected service time, each plane would need to come into the shop only one-fourth as often. Use queuing theory to compare the two alternatives on a meaningful basis.

Solution. In the first scenario (current policy) the queuing system is $M/M/1$ with parameters

$$\lambda = 1 \text{ engine/day} \quad \mu = 2 \text{ engine/day}.$$

Then we have $\rho = \frac{1}{2}$, $P_0 = \frac{1}{2}$, $P_n = P_0\rho^n = (\frac{1}{2})^{n+1}$, from which

$$\begin{array}{l} L = \frac{\lambda}{\mu-\lambda} = 1 \text{ engine} \\ W = \frac{L}{\lambda} = 1 \text{ day} \\ W_q = W - \frac{1}{\mu} = \frac{1}{2} \text{ day} \\ L_q = \lambda W_q = \frac{1}{2} \text{ engine.} \end{array}$$

In the second scenario (proposed policy) the maintenance of each aircraft is given by the sequence of four identical operations, one for each engine. Therefore the service time follows an Erlang distribution with parameter $k = 4$. The corresponding $M/E_4/1$ system has the following parameters:

$$\lambda = \frac{1}{4} \text{ aircraft/day} \quad \mu = \frac{1}{2} \text{ aircraft/day}.$$

Using the relations of $M/E_k/1$ systems we get

$$\begin{array}{l} L_q = \frac{1+k}{2k}\frac{\lambda^2}{\mu(\mu-\lambda)} = \frac{5}{8}\frac{\frac{1}{16}}{\frac{1}{2}(\frac{1}{2}-\frac{1}{4})} = \frac{5}{16} \text{ aicraft} \\ W_q = \frac{L_q}{\lambda} = \frac{5/16}{1/4} = \frac{5}{4} \text{ day} \\ W = W_q + \frac{1}{\mu} = \frac{5}{4} + \frac{1}{1/2} = \frac{13}{4} \text{ day} \\ L = \lambda W = \frac{1}{4}\frac{13}{4} = \frac{13}{16} \text{ aircraft.} \end{array}$$

In the second scenario the number of aircrafts in the system is smaller. Furthermore the time to traverse the system for each aircraft in the second scenario is also less than four times the time required for each engine in the first scenario.

4 Design and optimization of queuing systems

4.1 HL 16.4-2

A small grocery store has a single chackout stand with a full-time cashier. Customers arrive at the stand according to a Poisson process at a mean rate of 30 per hour. The service-time distribution is exponential with a mean of 1.5 minutes. This situation has resulted in occasional long lines and complaints from customers. Therefore, beacuase there is no room for a second checkout stand, the proposal has been made that another person be hired to help the cashier by bagging the groceries. This help would reduce the expected time required to process a customer to 1 minute but the distribution still would be exponential. The total compensation for the new employee would be \$ 8 per hour, which is just half that for the cashier. It is estimated that the grocery store incurs lost profit due to lost future business of \$ 0.08 for every minute that each customer has to wait (including service time). The manager wants to determine on an expected total-cost basis whether it would be worthwhile to hire the new person.

- (a) Which decision model presented in Sec. 16.4 applies to this problem? Why?
- (b) Use this model to determine whether to continue the status quo or to adopt the proposal.

Solution. The system is $M/M/1$ with $\lambda = 30$ customers/hour. In the current system we have $\mu = 40$ customers/hour; in the scenario with the additional operator we would have $\mu' = 60$ customers/hour.

The cost C due to the time required to traverse the system (expressed in \$/hour) can be evaluated as

$$C = c\lambda W = cL$$

being $c = 4.8$ \$/hour the given coefficient (equivalent to 0.08 \$/minute). Every customer is subject to a cost (expressed in \$) equal to cW (where W must be expressed in hours) and therefore the whole population of λ customers/hour pays a cost $c\lambda W$ (expressed in \$/hour). Since

$$L = \frac{\lambda}{\mu - \lambda} \quad \text{e} \quad L' = \frac{\lambda}{\mu' - \lambda},$$

we have

$$L = \frac{30}{10} = 3 \quad \text{e} \quad L' = \frac{30}{30} = 1,$$

from which

$$C = 4.8 \times 3 = 14.4 \ \ \$/\text{hour} \quad \text{and} \quad C' = 4.8 \times 1 = 4.8 \ \ \$/\text{hour}.$$

The cost reduction that can be achieved hiring the additional employee is given by $C - C' = 14.4 - 4.8 = 9.6$ \$/hour, which is larger than the cost of the employee (equal to 8 \$/hour). Therefore hiring the employee yields a net improvement of $9.6 - 8 = 1.6$ \$/hour.

4.2 HL 16.4-3

Customers arrive at a fast-food restaurant with one server according to a Poisson process at a mean rate of 30 per hour. The server has just resigned and the two candidates for the replacement are X (fast but expensive) and Y (slow but inexpensive). Both candidates would have an exponential distribution for service times, with X having a mean of 1.2 minutes and Y having a mean of 1.5 minutes. Restaurant revenue per month is given by \$ $\frac{6000}{W}$, where W is the expected waiting time (in minutes) of a customer in the system. Determine the upper bound on the difference in their monthly compensation that would justify hiring X instead of Y.

Solution. The system is $M/M/1$ with $\lambda = \frac{1}{2}$ customers/minute. With X we have $\mu_X = \frac{5}{6}$ customers/minute; with Y we have $\mu_Y = \frac{2}{3}$ customers/minute. For $M/M/1$ systems we have $W = \frac{1}{\mu - \lambda}$. Then

$$W^X = \frac{1}{\frac{5}{6} - \frac{1}{2}} = 3 \ \ \text{minutes}$$

$$W^Y = \frac{1}{\frac{2}{3} - \frac{1}{2}} = 6 \ \ \text{minutes}$$

from which, being $W = W_q + \frac{1}{\mu}$,

$$W_q^X = 3 - \frac{1}{\frac{5}{6}} = \frac{9}{5} \ \ \text{minutes}$$

$$W_q^Y = 6 - \frac{1}{\frac{2}{3}} = \frac{9}{2} \ \ \text{minutes}$$

Then we have

$$R^X = \frac{6000}{\frac{9}{5}} = 3333.3 \ \ \$/\text{month}$$

$$R^Y = \frac{6000}{\frac{9}{2}} = 1333.3 \ \ \$/\text{month}.$$

The maximum difference in salary that justifies hiring X instead of Y is $R^X - R^Y = 2000$ \$/month.

4.3 HL 16.4-10

A single crew is provided for unloading and7or loading each truck that arrives at the loading dock of a warehouse. These trucks arrive according to a Poisson input process at a mean rate of one per hour. The time required by a crew to unload and/or load a truck has an exponential distribution (regardless of the crew size). The expected time required by a one-person crew is 1 hour. The cost of providing each additional member is \$ 20 per hour. The cost that is attributable to having a truck not in use (i.e. a truck standing at the loading dock) is estimated to be \$ 30 per hour.

- (a) Assume that the mean service rate of the crew is proportional to its size. What should be the size to minimize the expected total cost per hour?
- (b) Assume that the mean service rate of the crew is proportional to the square root of its size. What should be the size to minimize the expected total cost per hour?

Solution. This is an $M/M/1$ system: the server is the team of workers, which is always one although it can be composed by several workers. The input average frequency is $\lambda = 1$ customers/hour.

Indicating with s the number of workers in the team, the cost for providing the service is $C^{provider} = c_s(s-1)$, being $c_s = 20$ \$/hour the given coefficient.

The cost for each vehicle is given by $c_c \lambda W$ (expressed in \$), being $c_c = 30$ \$/hour the given coefficient, where the overall service time W is expressed in hours. The total cost for the whole population of customers is $C^{customers} = c_c \lambda W$ and it is expressed in \$/hour. In turn, in a $M/M/1$ system one has $W = \frac{1}{\mu-\lambda}$. Then, under assumption (a) one has $\mu = s\mu_1$ with $\mu_1 = 1$ customer/hour, while under assumption (b) one has $\mu = \sqrt{s}\mu_1$.

Then the total cost in the system is

$$C = C^{provider} + C^{customers} = c_s(s-1) + c_c \frac{\lambda}{s\mu_1 - \lambda}$$

under assumption (a) and

$$C = C^{provider} + C^{customers} = c_s(s-1) + c_c \frac{\lambda}{\sqrt{s}\mu_1 - \lambda}$$

under assumption (b).

The minimum value of the cost can be found by trials, assuming different values of s, but in order to better orient our trials it is advisable to first solve the continuous relaxation of the problem analytically (or with a non-linear programming solver), computing the derivative with respect to the independent variable s. In scenario (a) we have:

$$\frac{d\ C}{d\ s} = c_s + c_c \lambda \frac{-\mu_1}{(s\mu_1 - \lambda)^2}$$

The first order condition is then

$$c_s + c_c \lambda \frac{-\mu_1}{(s\mu_1 - \lambda)^2} = 0$$

from which we get

$$s = \frac{\lambda + \sqrt{\frac{c_c \lambda \mu_1}{c_s}}}{\mu_1} = \frac{1 + \sqrt{\frac{30\ 1\ 1}{20}}}{1} = 1 + \sqrt{\frac{3}{2}} \approx 2.22$$

Then, it makes sense to compare the cost for $s = 2$ and $s = 3$.

For $s = 2$ we have

$$C = c_s(2-1) + c_c \frac{\lambda}{2\mu_1 - \lambda} = 20 + 30 \frac{1}{2-1} = 50 \text{ \$/hour.}$$

For $s = 3$ we have

$$C = c_s(3-1) + c_c \frac{\lambda}{3\mu_1 - \lambda} = 40 + 30 \frac{1}{3-1} = 55 \text{ \$/hour.}$$

Therefore, under assumption (a) the optimal size of the team is $s = 2$.

In scenario (b) per $s = 2$ we have

$$C = c_s(2-1) + c_c \frac{\lambda}{\sqrt{2}\mu_1 - \lambda} = 20 + 30 \frac{1}{\sqrt{2}-1} \approx 92.43.$$

For $s = 3$ we have

$$C = c_s(3-1) + c_c \frac{\lambda}{\sqrt{3}\mu_1 - \lambda} = 40 + 30 \frac{1}{\sqrt{3}-1} \approx 80.98.$$

For $s = 4$ we have

$$C = c_s(4-1) + c_c \frac{\lambda}{\sqrt{4}\mu_1 - \lambda} = 60 + 30 \frac{1}{\sqrt{4}-1} = 90.00.$$

Them under assumption (b) the optimal size of the team is $s = 3$.

5 Additional exercises on queuing systems

5.1 First aid

In a small hospital two physicians are available for first aid. Patients are classified as "urgent" and "non-urgent". Once the treatment of a patient has started, it cannot be interrupted. However, it is undesirable that both physicians be simultaneously busy with one or two "non-urgent" patients and thus nobody be available in case a new urgent patient arrives. Therefore the policy is the following: when both physicians are idle, any incoming patient is immediately treated; when one of the two physicians is busy the next patient is treated only if he is urgent; otherwise he is kept in a waiting room. To avoid excessively long queues, the hospitals immediately alerts the ambulances so that no further patients are carried there. In particular no urgent patients are accepted when both physicians are busy and no additional patients are accepted if there is one already in the waiting room. Arrivals are allowed again as soon as the conditions for blocking them is no longer verified. When arrivals are allowed, non-urgent patients arrive with an average period of 15 minutes; urgent patients arrive with an average period of 45 minutes. All treatments last 20 minutes in average.

The hospital manager wants to know:

- For what fraction of time the first aid system is blocked for urgent patients?
- For what fraction of time the first aid system is blocked for non-urgent patients?
- What is the workforce utilization factor?
- What is the probability that a non-urgent patient must wait?
- How would these indicators change if the physicians treated all incoming patients immediately?
- How would these indicators change with three physicians, with the current policy of keeping the last physician idle if there no urgent patients to treat?

Solution. The exercise requires to represent the system with a state-transition graph, where each node corresponds to a possible state and each arc to a possible state transition. Then, we can compute the probability P_i associated with each state i from the knowledge of the average frequency f_{ij} of each transition from each state i to each state j. In this example the system can be represented with the graph in Figure 1.

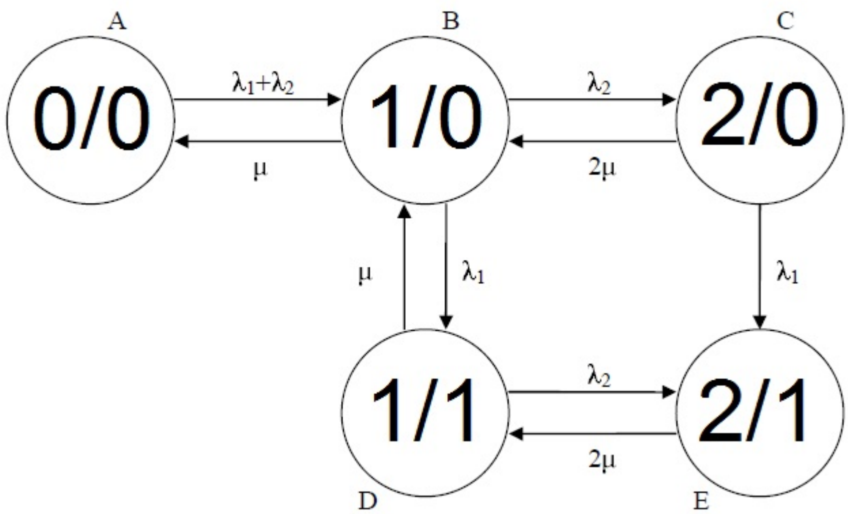


Figure 1: The state-transition diagram for questions 1 to 4.

Each state is indicated by two numbers: the first one is the number of busy servers (physicians), while the second one is the number of patients in the queue (waiting room). Therefore the sum of the two numbers is the total number of patients in the system. The frequencies on the arcs are indicated as functions of

- the mean arrival frequency of urgent patients: $\lambda_1 = \frac{4}{3}$ patients/hour;
- the mean arrival frequency of non-urgent patients: $\lambda_2 = 4$ patients/hour;
- the mean frequency of completions for each physician: $\mu = 3$ patients/hour.

f	A	B	C	D	E
A	-	$\frac{16}{3}$	0	0	0
B	3	-	$\frac{4}{3}$	4	0
C	0	6	-	0	4
D	0	3	0	-	$\frac{4}{3}$
E	0	0	0	0	-

f	A	B	C	E
A	-	$\frac{16}{3}$	0	0
B	3	-	$\frac{16}{3}$	0
C	0	6	-	4
E	0	0	6	-

The same frequencies can be reported in a square matrix with as many rows and columns as the number of states.

In order to compute the probabilities associated with each state, we have to formulate a system of balance equations with five unknowns and five equations (but one of the equations is redundant).

$$\begin{cases} (\lambda_1 + \lambda_2)P_A = \mu P_B \\ (\lambda_1 + \lambda_2 + \mu)P_B = (\lambda_1 + \lambda_2)P_A + 2\mu P_C + \mu P_D \\ (\lambda_1 + 2\mu)P_C = \lambda_2 P_B \\ (\lambda_2 + \mu)P_D = \lambda_1 P_B + 2\mu P_E \\ 2\mu P_E = \lambda_1 P_C + \lambda_2 P_D \end{cases}$$

Since the system is ill-conditioned, we must replace one of the equations with the normalization constraint:

$$P_A + P_B + P_C + P_D + P_E = 1.$$

Solving the linear system (with a spreadsheet, for instance), we obtain:

$$\begin{cases} P_A \approx 0.15 \\ P_B \approx 0.28 \\ P_C \approx 0.04 \\ P_D \approx 0.42 \\ P_E \approx 0.12 \end{cases}$$

Now we can answer some questions.

Question 1. The system does not accept further urgent patients when it is in states C and E. This occurs with probability $P_C + P_E =\approx 0.15$, i.e. for about 15% of the time.

Question 2. The system does not accept further non-urgent patients when it is in states D and E. This occurs with probability $P_D + P_E =\approx 0.53$, i.e. for about 53% of the time.

Question 3. The workforce is not used when the system is in state A, half ot it is used in states B and D and it is fully used in states C and E. Therefore the utilization factor is

$$\rho = 0.0P_A + 0.5(P_B + P_D) + 1.0(P_C + P_E) \approx 0.50.$$

Question 4. A non-urgent patient is kept in queue when the arrival occurs with the system in states B or C and this occurs with probability $P_B + P_C \approx 0.31$.

Question 5. To answer question 5 we must slightly modify the description of the system and then to repeat the same procedure. With the alternative policy of immediately serving all incoming patients, state D is never reached. The new state-transition diagram is indicated in Figure 2.

The matrix with the transition frequencies is now the following.

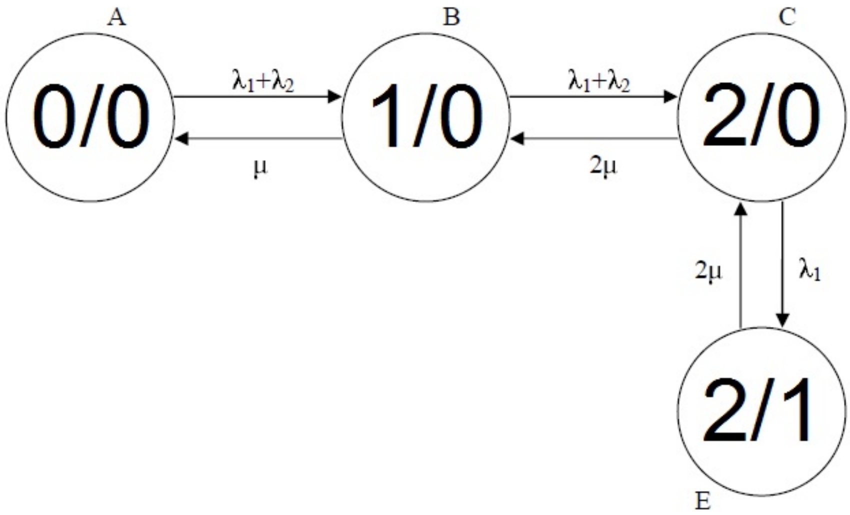


Figure 2: The state-transition diagram for question 5.

The solution is:

$$\begin{cases} P_A \approx 0.18 \\ P_B \approx 0.33 \\ P_C \approx 0.29 \\ P_E \approx 0.19 \end{cases}$$

Therefore the four answers to the four previous questions are now:

- Question 1: $P_C + P_E \approx 48.7\%$
- Question 2: $P_D + P_E \approx 19.5\%$
- Question 3: $\rho \approx 65.1\%$
- Question 4: $P_C \approx 29\%$

Question 6. With three physicians instead of two we have the state-transitions diagram outlined in Figure 3.

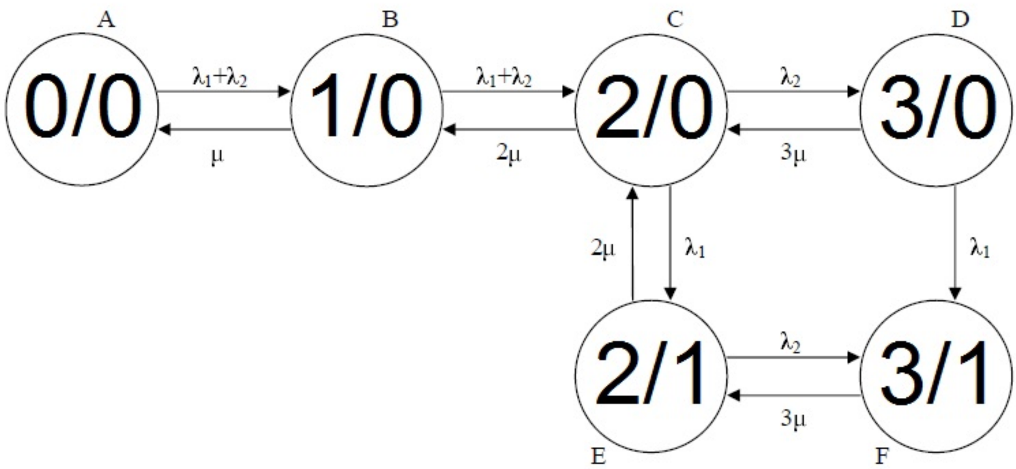


Figure 3: The state-transition diagram for question 6.

The matrix with the transition frequencies is now the following.

f	A	B	C	D	E	F
A	-	$\frac{16}{3}$	0	0	0	0
B	3	-	$\frac{16}{3}$	0	0	0
C	0	6	-	$\frac{4}{3}$	4	0
D	0	0	9	-	0	4
E	0	0	6	0	-	$\frac{4}{3}$
F	0	0	0	0	9	-

The solution is:

$$\begin{cases} P_A \approx 0.17 \\ P_B \approx 0.30 \\ P_C \approx 0.27 \\ P_D \approx 0.03 \\ P_E \approx 0.20 \\ P_F \approx 0.04 \end{cases}$$

Therefore the four answers to the four previous questions are now:

- Question 1: $P_D + P_F \approx 7\%$
- Question 2: $P_E + P_F \approx 23.7\%$
- Question 3: $\rho \approx 47.7\%$
- Question 4: $P_C + P_D \approx 29\%$

5.2 The buffer

Jobs arrive at a working center as Poisson events with an average frequency of 10 jobs per minute. They are put in a buffer, managed according to a FIFO policy, from which they are taken to be processed on one of the two available machines. The two machines are identical and they work in parallel independently of each other. Each one requires a variable time to process each job: the distribution of the service times is exponential with average value equal to 10 seconds per job.

1. What is the utilization factor of the queuing system?
2. What is the average number of jobs in the buffer?
3. What is the average number of jobs in the system (buffer+machines)?
4. What is the average time during which a job remains in the buffer?
5. What is the average time that a jobs spends in the system?

Furthermore one wants to know how to design the capacity of the buffer so that the probability that it si not enough to accommodate the waiting jobs is less than 3%.

Finally one wants to evaluate the profitability of replacing the two machiens with a single new machine that processes all jobs in a fixed time equal to 5 seconds. If the cost of having jobs in the queue is evaluated to be 0.2 €/minute and the cost for upgrading the production line is 2000 €, how long does it take to amortize the investment?

Solution. In the current scenario we must analyze a $M/M/2$ system with parameters $\lambda = 10$ jobs/minute and $\mu = 6$ jobs/minute. Therefore we have:
a) $\rho = \lambda/s\mu = 5/6 = 0.8333 = 83.33\%$.
b) $P_0 = 1/11$ and then $L_q = 125/33 = 3.788$ jobs.
c) $L = 60/11 = 5.4545$ jobs
d) $W_q = 25/66 = 0.3788$ minutes
e) $W = 6/11 = 0.5454$ minutes

From the probabilities P_i we can see that the probability of having more than K jobs in the system (buffer + machines) is 2.84% for $K = 19$ while it would be $2.84 + 0.57 = 3.41\%$ for $K = 18$. Then the probability is less than 3% for $K = 19$, which corresponds to a buffer capacity of 17 (2 jobs are in the machines, not in the buffer).

In the second scenario the model is $M/D/1$ with parameters $\lambda = 10$ jobs/minute and $\mu' = 12$ jobs/minute. Therefore we have:
a) $\rho' = \lambda/\mu' = 10/12 = 0.8333 = 83.33\%$.
b) $L'_q = 2.0833$ jobs
c) $L' = 2.9167$ jobs
d) $W'_q = 0.2083$ minutes
e) $W' = 0.2917$ minutes

The cost C due to the queue is given by the average number of jobs in the queue, L_q, times the cost for each job, which is given. In the first scenario we have $C = 3.0788 \times 0.2 \approx 0.75$ €/minute. In the second scenario we have $C' = 2.0833 \times 0.2 \approx 0.4167$ €/minute. The difference is $C - C' = 0.33$ €/minute. Therefore, to amortize the upgrade cost of 2000 €, 6000 minutes are needed, i.e. 100 hours.

5.3 The railway station

In a railway station customers arrive at the counter to buy tickets. The inter-arrival time is a random variable with an exponential distribution of probability with a mean of 4 customers every 10 seconds. There are 6 operators at the counter and each one takes in average 10 seconds to serve a customer. The probability distribution of the service time is exponential.

Evaluate

1. what is the utilization factor
2. what is the average number of customers in the system
3. what is the average number of customers in queue
4. what is the average time for each customer to traverse the system
5. what is the average time spent in queue for each customer.

Evaluate what would happen by replacing the 6 operators with 5 machines. Their service time would be constant and equal to the average service time of operators, i.e. 10 seconds per customer.

Solution. In the first scenario we have an $M/M/6$ system with parameters $\lambda = 0.4$ customers/second and $\mu = 0.1$ customers/second. Hence, we have:

1. $\rho = \lambda/s\mu = 2/3 = 0.6667 = 66.67\%$.
2. $L = 4.5695$ customers.
3. $L_q = 0.5695$ customers
4. $W = 11.4238$ seconds
5. $W_q = 1.4238$ seconds

In the second scenario the model is $M/D/5$ with parameters $\lambda = 0.4$ customers/second and $\mu = 0.1$ customers/second. Hence, we have:

1. $\rho' = \lambda'/s'\mu' = 0.8 = 80\%$.
2. $L' = 5.6220$ customers.
3. $L'_q = 1.6220$ customers.
4. $W' = 14.0550$ seconds.
5. $W'_q = 4.0550$ seconds.

5.4 Drilling machines

In a mechanical factory there is a machine with two drills that work in parallel, picking up the mechanical parts from a "first-in-first-out" buffer fed by a conveyor belt. When the buffer is full, the conveyor stops and it restarts as soon as there is room in the buffer again. When the conveyor works, the mechanical parts arrive with random inter-arrival times, according to an exponential distribution of probability with a mean of 12 seconds. The drilling time for each part is a random variable with an exponential distribution and a mean of 20 seconds. Compute what the buffer size should be so that the probability that the conveyor should stop be less than 4%.

Evaluate the proposal of replacing the current machine with a new one, equipped with a single drill working three times faster than the current drills. i.e. able to process three mechanical parts in 20 seconds in average. The drilling time would still have an exponential distribution.

Solution. The system is an $M/M/2/K$ with finite capacity K, because arrivals do not occur when the buffer is full. The parameters are $\lambda = 5$ parts/minute and $\mu = 3$ parts/minute. We need room for 10 parts in the system, i.e. 8 in the buffer and 2 in the machine. With this choice the probability that the buffer be full, i.e. P_{10} is 3.34%. With smaller capacity, the requirement is violated: with $K = 9$ we have $P_9 = 4.28\%$.

In the alternative scenario the system is $M/M/1/K$ with $\lambda = 5$ parts/minute and $\mu = 9$ parts/minute. We need room for 4 parts in the buffer besides 1 part in the machine. With $K = 5$, we have $P_5 = 2.42\%$, which is acceptable, while with $K = 4$ we have $P_4 = 4.47\%$ which is not acceptable.

5.5 Truck unloading

Trucks provide coils to a steelworks warehouse. They arrive with random inter-arrival time, distributed according an exponential probability distribution with a mean of 20 minutes. The time needed by a team of workers to unload a truck is variable and it has an exponential probability distribution with a mean of half an hour. Evaluate how many teams must work in parallel to reduce the average time of each truck in the system to less than one hour and the average number of trucks at the steelworks to less than three.

Solution. The system is an $M/M/s$ system, where $\lambda = 3$ trucks/hour, $\mu = 2$ trucks/hour and s is a design variable. For $s = 1$ the system does not reach a steady-state condition, because the utilization factor is larger than 1. For $s = 2$ we have $W = 1.14$ hours and $L = 3.43$ trucks, which does not satisfy the constraints. For $s = 3$ we have $W = 0.58$ hours and $L = 1.74$ trucks, which satisfies both constraints.

5.6 E-mail orders

Orders arrive from commercial agents by e-mail. The messages arrive with random inter-arrival times, with an exponential probability distribution with a mean of 15 minutes. They are read and processed by a secretary, who takes a variable time, with an exponential probability distribution and a mean of 12 minutes. Evaluate the percentage variation in the number of messages in the queue if the secretaries were two, working in parallel with the same speed of current one.

Solution. This is an $M/M/1$ system, to be compared with an $M/M/2$ system, in which $\lambda = 4$ msg/hour and $\mu = 5$ msg/hour. With $s = 1$ we have $L_q = 3.2$ msg. With $s = 2$ we would have $L'_q = 0.15$ msg, equal to 4.76% of L_q.

5.7 Maintenance

In a production plant a certain number of identical machines will be installed. These machines require maintenance after 5 days, i.e. 120 hours, in average. This working time is a random variable with an exponential probability distribution. Some operators will be in charge for the maintenance. The time needed by each of them for each machine is a random variable with exponential probability distribution and a mean of 2 hours. Owing to a suitable rotation of work-shifts a team of operators will be available 24h a day.

According to the design the number of machines should be 20 and the number of operators in the maintenance team should be 3. What is the fraction of time during which each operator would be busy in average? How long will a machine stop last, in average?

The general manager of the new plant wants every worker to be active for maintenance of machines for no less than 20% and no more than 33% of its working time. Within these constraints he wants to decide the number of machines in the plant and the number of operators in the maintenance team so that the duration of the machine stop be minimum. Owing to economies of scale, machines must be bought in groups of five and there is no room for more than 30 machines.

Solution. This is an $M/M/s/K$ system with finite population. When $s = 3$ and $K = 20$ we have $\rho = 10.93\%$ and $W = 0.0834$ hours. Therefore the operators are under-utilized. Considering $K = 20, 25, 30$ and $s = 1, 2, 3$, we obtain $20\% \leq \rho \leq 33\%$ only in three cases: (a) $K = 20$ and $s = 1$, (b) $K = 25$ and $s = 2$, (c) $K = 30$ and $s = 2$. The minimum machine stop time is in case (b) and it is equal to 0.0865 days, i.e. 2.076 hours for each maintenance intervention.

5.8 The highway

Vehicles arrive at a highway toll booth as Poisson events, with an average frequency of one vehicle every 10 seconds. The three automatic card readers have become unavailable owing to a black-out. Their service time was constant and equal to 20 seconds per vehicle. They must be replaced by operators working in parallel. Each operator has a service time described by a random variable with an exponential probability distribution with a mean of 30 seconds. How many operators must be used to reach a steady-state condition? What would the expected waiting time in queue for each vehicle be in that case? How many operators must be used so that the expected time spent in queue by each vehicle be not larger than with the three parallel card readers?

Solution. In the scenario with the three parallel card readers, the system is $M/D/3$, with $\lambda = 6$ vehicles/minute and $\mu = 3$ vehicles/minute. So, we have $W_q \approx 0.1$ minutes. In the new scenario, the system is $M/M/s$ with $\lambda = 6$ vehicles/minute and $\mu = 2$ vehicles/minute. The system needs at least 4 operators to reach a steady-state condition. In this case $W_q \approx 0.25$ minutes. With $s = 5$ we have $W_q = 0.06$ minutes, which is not worse than in the case with the card readers.

5.9 Bar-code and RFID

Goods received by a large supermarket are unloaded in a suitable structure equipped with 4 unloading bays, usable in parallel, each one with its own team of workers. Trucks arrive as Poisson events with an average frequency of 5 trucks per hour. The time needed to unload them in 30 minutes in average and it has an exponential distribution of probability. All unloaded packages are immediately identified by a bar-code to be sent to the internal warehouse and possibly to the shelves. You are required to provide a quantitative evaluation of the impact of replacing bar-code with RFID: it is estimated that this would allow reducing the average unloading time from 30 to 25 minutes, while keeping the same probability distribution.

Solution. In the current $M/M/4$ system, with the bar-code, we have $\lambda = 5$ trucks/hour, $\mu = 2$ trucks/hour and $s = 4$ parallel servers. The performance indicators are: $\rho = 62.5\%$, $L = 3.03$ trucks, $L_q = 0.53$ trucks, $W \approx 0.6$ hours (i.e. about 36 minutes), $W_q \approx 0.1$ hours (i.e. about 6 minutes).

In the alternative $M/M/4$ system, with RFID, we have $\lambda' = 5$ trucks/hour, $\mu' = 2.4$ trucks/hour and $s = 4$ parallel servers. The performance indicators are: $\rho' \approx 52.1\%$, $L' \approx 2.3$ trucks, $L'_q \approx 0.2$ trucks, $W' \approx 0.46$ hours (i.e. about 27.5 minutes), $W'_q \approx 0.04$ hours (i.e. about 2.4 minutes).

5.10 Bays

At a warehouse trucks arrive as Poisson events at an average frequency of one truck every 15 minutes. The duration of the unloading operations has an exponential probability distribution with a mean of 45 minutes. When a truck must wait because all bays are busy, the warehouse must pay a penalty of 1 €/minute. The hourly wage for the operators at each unloading bay is 75 €/hour. What is the optimal number of bays and what are the corresponding costs? A new contract with the transportation companies compels the warehouse to keep the average system traversal time (waiting + unloading) for each truck within one hour. Solve the problem with this additional constraint.

Solution. This is an $M/M/s$ system. The minimum number of servers (bays) needed to reach steady-state condition is 4; otherwise the requirement $\rho < 1$ is not satisfied. With 4 bays we have $W = 1.13$ hours, $W_q = 0.38$ hours, $L_q = 1.53$ vehicles, waiting cost equal to 91.70 €/h, cost for the bays equal to 300 €/h and total cost equal to 391.70 €/h. This is the solution of minimum cost. However it would become infeasible with the new contract, because W is larger than 1 hour. With 5 bays we have $W = 0.84$ hours (50.3 minutes), $W_q = 0.09$ hours (5.3 minutes), $L_q = 0.35$ vehicles, waiting cost equal to 21.25 €/h (0.35 €/min), service cost equal to 375 €/h (6.25 €/min) and total cost equal to 396.25 €/h (6.60 €/min). This solution is feasible with the new contract and it is optimal in the new scenario.

5.11 Quality test

Mechanical parts output from a high precision process are checked by a quality test operator. He does three measures on each mechanical part. The three operations are identical but their duration is a random variable with exponential distribution and a mean of 1 minute. Mechanical parts arrive as Poisson events at an average rate of two parts every three minutes. For each given number of operators s, the policy is to split the incoming flow of parts into s identical flows, so that operator k checks parts, k, $s+k$, $2s+k$, ... What is the minimum number of operators needed? With that number of operators, how long does it take in average for a part to traverse the quality check phase? What is the average number of parts waiting for being checked? If one wants to cut both the above indicators by 50%, what is the minimum number of operators needed and what is their utilization factor?

Solution. The way the input flow of parts is partitioned a priori into identical flows, independently of the state of the queues at each point in time, makes this system a set of s identical single-server queuing systems, each one serving a fraction $1/s$ of the demand.

Since service consists of a sequence of three identical operations, the service time has an Erlang distribution with parameter $k = 3$. The mean time per each measure is 1 minute and three measures are required; therefore the mean time for checking a part is 3 minutes. Hence, the mean completion rate for each operator is $\mu = 1/3$ part per minute.

The input process is Poisson and when it is partitioned into s flows, all of them are Poisson. The arrival rate for each queue is a fraction $1/s$ of the total arrival rate, which is 2/3 parts per minute. Hence, the arrival rate for each queue is $\lambda = 2/(3s)$ parts per minute.

The utilization factor of each $M/E_3/1$ system is $\rho = \lambda/\mu = 2/s$. Hence, to reach a steady-state condition (which implies $\rho < 1$) at least $s = 3$ operators are needed.

From the analysis of the $M/E_3/1$ systems with $\lambda = 2/9$ and $\mu = 1/3$, we obtain $W = 7$ minutes and $L_q = 0.\overline{8}$ parts.

To cut these two indicators by 50%, we require $W \leq 3.5$ minutes and $L_q \leq 0.\overline{4}$ parts. Four operators are enough to satisfy the latter condition, but ten operators are required to satisfy the former one.