Exercises
Stochastic Performance Modelling

Hamilton Institute, Summer 2010

Exercises Stochastic Performance Modelling
Instruction 1

Exercise 1 Let X be a non-negative random variable with $\mathbb{E}[X^2] < \infty$, having probability density function $f(\cdot)$.

Use partial integration to show that

$$\mathbb{E}[X] = \int_{y=0}^{\infty} \mathbb{P}(X > y)\mathrm{d}y,$$

and

$$\mathbb{E}[X^2] = \int_{y=0}^{\infty} 2y\mathbb{P}(X > y)\mathrm{d}y.$$

Exercise 2 Let X_i be an exponentially distributed random variable with parameter λ_i, $i = 1, 2$.

(a) Determine $\mathbb{P}(X_1 > s + t \mid X_1 > s)$.
(b) Determine $\mathbb{P}(X_1 \leq X_2)$.
(c) Determine $\mathbb{P}(\min\{X_1, X_2\} < t)$.

Exercise 3 Consider two parallel processors, 1 and 2. Job A_i is being processed by processor $i = 1, 2$. The processing times of jobs at processor i are exponentially distributed with parameter λ_i, $i = 1, 2$. Job A_3 is waiting in line and will be processed by the processor that completes its current job first.

(a) Let $\lambda_1 = \lambda_2 = \lambda$. Without calculations, determine the probability that job A_3 is the last of the three jobs to be completed.
(b) Determine the probability that job A_3 is the last of the three jobs to be completed for arbitrary values of λ_1 and λ_2.

Exercise 4 (The Erlang distribution)
Let $X_1, X_2, \ldots$ be a sequence of independent, exponentially distributed random variables with common parameter λ.

(a) Determine the density of $X_1 + X_2$.
Hint: Use the following property: If X and Y are independent, non-negative random variables with densities $f(\cdot)$ and $g(\cdot)$, respectively, then the density of $X + Y$ equals $\int_{u=0}^{t} f(u)g(t-u)\mathrm{d}u$.
(b) Use induction on n to show that the density of $S_n = \sum_{i=1}^{n} X_i$ is given by

$$f_{S_n}(t) := \lambda \mathrm{e}^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!}.$$

Remark: This is the density of the *Erlang* distribution with parameters n and λ.
(c) Verify that the above density is the derivative with respect to t of the Erlang distribution given in Section 2.2.4 of the lecture notes.

Exercise 5 (The hyper-exponential distribution)
Let X_i be an exponentially distributed random variable with parameter λ_i, $i = 1, 2$. Suppose B is a random variable with $\mathbb{P}(B = 1) = p$ and $\mathbb{P}(B = 2) = 1 - p$, with $0 < p < 1$. Let Y be defined as follows: $Y = X_1$ if $B = 1$ and $Y = X_2$ if $B = 2$.

Calculate $\mathbb{P}(Y > x)$ by conditioning on the values of B. Subsequently, determine the density of Y.

Exercise 6 The discrete random variable N is said to have a *geometric* distribution if $\mathbb{P}(N = n) = (1-p)p^{n-1}$, with $p \in (0,1)$ for all $n = 1, 2, 3, \ldots$. Note that this is indeed a probability distribution because $\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$, if $|x| < 1$.

(a) Use the above to show that $\mathbb{P}(N > m) = p^m$, for all $m = 0, 1, 2, \ldots$.
(b) Show that the expectation of N is given by $\sum_{m=0}^{\infty} \mathbb{P}(N > m)$, which, according to the above, equals $\frac{1}{1-p}$.
Hint: Compare with the expression for the mean in Exercise 1.

Let $X_1, X_2, X_3, \ldots$ be a sequence of independent, exponentially distributed random variables with common parameter λ. All X_i's are independent of N. Let $S_N := X_1 + X_2 + \ldots + X_N$, that is, if $N = n$ then S_N is the sum of the first n terms of the sequence $X_1, X_2, X_3, \ldots$.

(c) Show that S_N has an exponential distribution by arguing that the density of S_N equals

$$f_{S_N}(t) := \sum_{n=1}^{\infty} \mathbb{P}(N = n) f_{S_n}(t),$$

where $f_{S_n}(t)$ is given in Exercise 4.
Hint: Use that $\sum_{k=0}^{\infty} \frac{x^k}{k!} = \mathrm{e}^x$ for any real number x.

Exercises Stochastic Performance Modelling
Instruction 2

Exercise 7 Consider a Markov chain with state space $\{1, 2, 3\}$ and transition probabilities $p_{12} = 1, p_{21} = 1/2, p_{23} = 1/2, p_{32} = 1/2$ and $p_{33} = 1/2$.

Determine the steady-state distribution of this Markov chain.

Exercise 8 Consider a Markov chain with state space $\{1, 2, 3, 4\}$ and transition probability matrix

$$P = \begin{bmatrix} 0 & \frac{1}{4} & 0 & \frac{3}{4} \\ 0 & 0 & 1 & 0 \\ 0 & \frac{2}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Let X_n be the state of the Markov chain at time $n = 0, 1, 2, \ldots$. Calculate $\lim_{n\to\infty} \mathbb{P}(X_n = k|X_0 = 1)$, $k = 1, 2, 3, 4$.

Exercise 9 A processor is inspected weekly in order to determine its condition. The condition of the processor can either be *perfect*, *good*, *reasonable*, or *bad*. A new processor is still perfect after one week with probability 0.7, with probability 0.2 the state is good, and with probability 0.1 it is reasonable. A processor in good condition is still good after one week with probability 0.6, reasonable with probability 0.2, and bad with probability 0.2. A processor in reasonable condition is still reasonable after one week with probability 0.5 and bad with probability 0.5. A bad processor must be repaired. The repair takes one week, after which the processor is again in perfect condition.

(a) Formulate a Markov chain that describes the state of the machine, and draw the corresponding transition diagram.
(b) Determine the steady-state distribution of the Markov chain.

Exercise 10 The number of orders being processed at a factory can be described by a Markov chain with state space $\{0, 1, 2, 3, \ldots\}$. For a given positive integer N, the transition probabilities are $P_{0,0} = 0.5$, $P_{i,i+1} = 0.5$ for $i = 0, \ldots, N$, $P_{i,i+1} = p$ for $i = N + 1, N + 2, \ldots$, $P_{i,i-1} = 0.5$ for $i = 1, \ldots, N$ and $P_{i,i-1} = 1 - p$ for $i = N + 1, N + 2, \ldots$, with $p < 0.5$.

(a) Why is this Markov chain irreducible, aperiodic and positive recurrent?
(b) Determine the steady-state distribution of the Markov chain.
(c) Let $N = 3$. Calculate the expected number of transitions needed to reach state 3, starting from state 0.

Exercise 11 Consider a Markov chain with state space $\{1, 2, 3\}$ and transition probabilities $p_{11} = p_{12} = 1/4$, $p_{13} = 1/2$, $p_{21} = 1/4$, $p_{23} = 3/4$, $p_{31} = p_{33} = 1/2$.

(a) Determine the steady-state distribution of this Markov chain.
(b) Calculate the expected number of transitions (including transitions that do not alter the state) needed to reach state 3, starting from state 1.

Exercise 12 Consider a Markov chain with state space $\{1, 2, 3, 4, 5, 6\}$ and transition probability matrix

$$\begin{pmatrix} \frac{1}{3} & 0 & 0 & 0 & \frac{2}{3} & 0 \\ 0 & \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{4}{5} & \frac{1}{5} & 0 & 0 \end{pmatrix}.$$

(a) Draw the corresponding transition diagram, and determine the classes of communicating states.
(b) Verify – for all possible initial states – whether there exists a limiting distribution and, if so, determine this distribution.
(c) What is the probability that state 1 is ever reached, starting in state 2?

Exercises Stochastic Performance Modelling
Instruction 3

Exercise 13 Two types of consultations occur at a database according to two independent Poisson processes: 'read' consultations occur at rate λ_R and 'write' consultations at rate λ_W.

(a) What is the probability that the time between two consecutive 'read' consultations is longer than t?
(b) What is the probability that the first consultation occurring after time $t = 0$ is a 'read' consultation?
(c) What is the probability that during the time interval $[0, t]$ at most three 'write' consultations occur?
(d) What is the probability that during the time interval $[0, t]$ at least two consultations occur in total?
(e) Determine the distribution of the number of 'read' consultations occurring during the time interval $[0, t]$, given that in total n consultations occurred during this time interval.

Exercise 14 The home page of the extremely popular T.E. Acher is frequently consulted by many students worldwide. New 'visits' to the home page by students occur according to a Poisson process with an average of 10 visits per hour. Mr. Acher is also highly respected by colleagues. The average number of colleagues who visit the home page per hour is 2 (according to a Poisson process as well).

(a) What is the probability that the home page is visited at least twice during one hour?
(b) What is the probability that the home page is not visited at all over the course of 15 minutes?

A student who visits the home page 'clicks' on the link to an overview of Mr. Acher's research activities with probability 2/5.

(c) Determine the probability that during one work day (8 hours) exactly 1 student consults the research overview.

Exercise 15 A production facility consists of two machines, M1 and M2. M1 makes electronic parts of type A. The time it takes M1 to make a part is exponentially distributed with a mean production time of 3 minutes. M2 makes electronic parts as well, the production time being exponentially distributed with a mean of 6 minutes. The material used by M2 varies over time and, as a consequence, the parts produced by M2 are of type A or B, each with probability $\frac{1}{2}$.

(a) What is the probability that M2 makes at least 3 parts during a half-hour period? (The types of the parts produced are not relevant.)
(b) Suppose a customer requests a type-A part when this part type is not available from stock. What is the probability that the customer must wait longer than 5 minutes until the next type-A part is available?

Both machines occasionally produce defective parts (this can not be detected until the part is put into use). A part produced by M1 is defective with probability $\frac{1}{5}$; for M2 the percentage of defective parts is even as high as 40%, regardless of the part type.

(c) What is the probability that a customer for a type-A part receives a defective part?
(d) Suppose customers request parts according to a Poisson process of rate λ. What is the probability that no part is produced during the period between two consecutive customer requests?

Exercise 16 Consider the Markov chain with state space $\{1, 2, 3, 4, 5, 6\}$ and transition probability matrix

$$P = \begin{pmatrix} 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \frac{1}{4} & 0 & \frac{3}{4} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{4} & 0 & 0 & \frac{3}{4} & 0 \end{pmatrix}.$$

(a) Draw the corresponding transition diagram, and determine the classes of communicating states.
(b) Determine the steady-state distribution of this Markov chain.
(c) What is the probability that, starting from state 1, the process *never* makes a direct transition from state 1 to state 3?
(d) Calculate the expected number of transitions needed to reach state 2, starting from state 1.

Exercise 17 Consider a Markov process with state space $\{1, 2, 3\}$ and infinitesimal generator

$$Q = \begin{pmatrix} -1 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -1 & \frac{1}{2} \\ \frac{3}{2} & \frac{3}{2} & -3 \end{pmatrix}.$$

(a) Determine the steady-state distribution of this Markov process.
(b) Also determine the steady-state distribution of the underlying Markov chain.

Exercise 18 Consider the Markov chain of Exercise 11. From this Markov *chain* we construct a Markov *process*: each visit to state i consists of an exponentially distributed sojourn time with parameters $\mu_1 = 1/2$, $\mu_2 = 1/3$ and $\mu_3 = 1$ in states 1, 2 and 3, respectively. Observe that it is possible for both states 1 and 3 to return to this state after such an exponentially distributed sojourn time.

(a) Determine the fraction of time spent in each of the states $i = 1, 2, 3$.
(b) Determine the generator matrix of this Markov process (use the answer to question (a)).
(c) Calculate the expected *amount of time* to reach state 3, starting from state 1. (Note that this is different from the expected *number of transitions*.)

Exercises Stochastic Performance Modelling
Instruction 4

Exercise 19 A repair man fixes broken TV sets. Broken TV sets arrive at his repair shop according to a Poisson process, with an average of 10 broken TV sets per work day (8 hours). The repair times are exponentially distributed with a mean of 30 minutes.

(a) What is the fraction of time that the repair man has no work to do?
(b) How many TV sets are, on average, at his repair shop?
(c) What is the mean sojourn time (waiting time plus repair time) of a TV set?

Exercise 20 A total of N lorries drive back and forth between a loading platform and an unloading platform of a transshipment terminal. Lorries queue up at the loading platform to be loaded by a (single) crane, with exponentially distributed loading times (per lorry) with parameter λ, all loading times being independent of each other. Similarly, items are removed from the lorries by a single crane at the unloading platform, with (independent) exponentially distributed unloading times with parameter μ, $\lambda \neq \mu$. Lorry driving times between the two platforms may be neglected (with respect to loading and unloading times), and it is further assumed that there is an abundance of items at the loading platform and that the system is in equilibrium.

(a) What is the probability that k of the N lorries are at the loading platform?
(b) The dynamics of this model are identical to that of a well-known queueing model. Describe that queueing model.

Exercise 21 Consider an M/M/1 queue with arrival rate λ and service rate μ, with $\mu > \lambda$. Let C_n be the expected time for the system to empty, starting with n customers, $n = 0, 1, \ldots$ (so $C_0 = 0$).

(a) Show that the C_n's satisfy the following recursive relationship:

$$C_n = \frac{1}{\lambda+\mu} + \frac{\mu}{\lambda+\mu}C_{n-1} + \frac{\lambda}{\lambda+\mu}C_{n+1}, \quad n = 1, 2, \ldots.$$

(b) Show that $C_1 = \frac{1}{\mu-\lambda}$.
Hint: Argue that C_1 is the expected time that the server is working without interruption and that $1/\lambda$ is the expected time that the system is empty without interruption; then $\frac{C_1}{C_1+1/\lambda}$ must be equal to λ/μ.
(c) Argue that the expected time to decrease the queue length from 2 customers to 1 customer is equal to C_1, so that $C_2 = 2C_1$, and in general $C_n = nC_1$. Verify this by substitution into the recursive relationship.

Exercise 22 A repair facility for automatic copiers has three repair men. Repair requests occur according to a Poisson process with a rate of $\lambda = 5$ per day. The repair times are exponentially distributed with a mean of $1/\mu = 0.5$ day. The system is assumed to be in equilibrium.

(a) Determine the transition diagram with the corresponding transition rates.
(b) What is the mean number of copiers at the repair facility?
(c) What is the mean out-of-order time (waiting time plus repair time) of a copier at the repair facility?
(d) What is the mean number of active repair men? What is the utilization factor of a repair man, i.e., the fraction of time that the repair man is busy?

Exercises Stochastic Performance Modelling
Instruction 5

Exercise 23 Two types of customers arrive at a post office with a single counter. Customers of type 1 are patient and always join the waiting line (if any) and wait for service. Customers of type 2 are less patient and only queue up if (upon arrival) there are less than K customers at the post office. Type-1 customers arrive according to a Poisson process with rate λ, and type-2 customers arrive according to an independent Poisson process with rate γ. Customer service times are type-independent and exponentially distributed with mean $1/\mu$.

(a) Determine the set of equilibrium equations, and calculate the equilibrium distribution.
(b) Determine an expression for the fraction of customers that impatiently leave the post office without having received service.

Exercise 24 At a post office with a single counter customers arrive according to a Poisson process with a rate of 60 customers per hour. Half of the customers have a service time that is the sum of a fixed time of 15 seconds and an exponentially distributed time with a mean of 15 seconds. The other half have an exponentially distributed service time with a mean of 1 minute.

Determine the mean waiting time and the mean queue length.

Exercise 25 Consider the M/G/1 queue with arrival rate $\lambda = \frac{1}{3}$ and with the following hyper-exponential service time distribution:

$$B(x) = \frac{1+a}{2}\left(1 - e^{-\frac{1+a}{2}x}\right) + \frac{1-a}{2}\left(1 - e^{-\frac{1-a}{2}x}\right), \quad x \geq 0,$$

with $0 \leq a < 1$.

(a) Determine the mean and the variance of the service times.
(b) For which values of a is the expected waiting time of a customer larger than the expected duration of a busy period?
(c) Verify that the behavior of this queue when $a \to 1$ is very different from the case $a = 1$.

Exercise 26 Consider an M/G/1 queue with arrival rate λ, mean service time $\mathbb{E}[B]$ and second moment of the service times $\mathbb{E}[B^2]$.

(a) Suppose $\lambda = \frac{3}{2}$, $\mathbb{E}[B] = \frac{1}{2}$ and $\mathbb{E}[B^2] = \frac{1}{2}$. Determine the expected values of the waiting time, the sojourn time, the queue length, the number of customers in the system and the busy period.
(b) Suppose that from measurements it is observed that the expected waiting time equals 5, that the traffic load ρ equals $\frac{2}{3}$, and that the mean number of customers in system is 8. Determine the arrival rate and the first two moments of the service time distribution. What are the expected values of the queue length and the sojourn time?

Exercises Stochastic Performance Modelling
Instruction 6

Exercise 27 Verify Example 6.2.1 in the lecture notes.

Exercise 28 A processor is used for two types of jobs: primary and secondary jobs. Primary jobs are generated according to a Poisson process with rate λ, and require exponentially distributed service times with mean $1/\mu$. We assume that $\rho := \lambda/\mu < 1$. When no primary jobs are available, the processor can render service to secondary jobs, of which there is always at least one waiting in line. As soon as a primary job arrives, the secondary job in service is interrupted. However, it takes a random amount of time T before the service of the primary job can start (for instance, because data of the secondary job needs to be saved on a hard disk). The set-up time T is exponentially distributed with mean $1/\theta$. When service of a primary job is completed, the first arrived primary job waiting in line is taken into service immediately (in that case there is no additional set-up time).

(a) Determine the arrival relation for the mean waiting time of a primary job. (Recall that the arrival relation expresses the mean waiting/sojourn time in terms of the mean number of customers found in the system upon arrival and their remaining service times.)
(b) Determine the mean waiting time and sojourn time of primary jobs, as well as the mean *total* number of primary jobs in the system and the mean number of waiting primary jobs.
(c) Why is the increase in waiting time (and sojourn time) compared to the M/M/1 queue without set-up times not smaller than $\mathbb{E}[T] = 1/\theta$?

Now assume that service times and set-up times have general distributions with means $\mathbb{E}[B]$ and $\mathbb{E}[T]$ and second moments $\mathbb{E}[B^2]$ and $\mathbb{E}[T^2]$. As before, we assume that $\rho := \lambda\mathbb{E}[B] < 1$.

(d) What is the probability that the processor is working on secondary jobs when a primary job arrives? And what is the probability that the processor is in the process of conducting a set-up when a primary job arrives?
(e) What is the arrival relation for the mean waiting time in this case?
(f) Determine the mean waiting time and the mean sojourn time of primary jobs.

Exercise 29 A machine mounts electronic components on three different types of printed circuit boards, type A, B and C boards say. On average 54 type-A boards arrive per hour, 48 type-B boards, and 18 type-C boards. The arrival processes are Poisson. The mounting times are exactly 20 seconds for type A, 30 seconds for type B, and 40 seconds for type C. The boards are processed in order of arrival.

(a) Calculate for each type of printed circuit board the mean sojourn time and also calculate the mean overall sojourn time.

Now suppose that there are priority rules in effect. Type-A boards have the highest priority, type-B boards have intermediate priority, and type-C boards have the lowest priority. The priorities are non-preemptive.

(b) Calculate for each type of printed circuit board the mean sojourn time and also calculate the mean overall sojourn time.

Exercise 30 Consider a single processor which handles three types of tasks. Type-i tasks arrive according to a Poisson process with rate λ_i, and the service requirements of type-i tasks are exponentially distributed with mean $1/\mu_i$, $i = 1, 2, 3$, with $\lambda_1/\mu_1 + \lambda_2/\mu_2 + \lambda_3/\mu_3 < 1$. Tasks are served in order of arrival.

(a) Argue why the system may be viewed as an $M/H_3/1$ queue, where H_3 represents a hyper-exponential service requirement distribution.
(b) Determine the mean waiting time of type-1 tasks.
(c) Suppose that at some point in time there are at least three tasks present in the system. What is the probability that the first two waiting tasks (thus excluding the task in service) belong to different types?
(d) Now suppose that the tasks are served according to a preemptive-resume priority strategy, which assigns the highest priority to type-1 tasks, the next highest priority to type-2 tasks, and the lowest priority to type-3 tasks. Determine the mean waiting time of an arbitrary task.

Exercises Stochastic Performance Modelling
Instruction 7

Exercise 31 A total of K jobs circulate in a closed network of four queues (Q_0, Q_1, Q_2, Q_3), each with a single server. At each of the four queues, the jobs are served in order of arrival, and the service times are exponentially distributed with means $\frac{1}{\mu_0} = 2$, $\frac{1}{\mu_1} = \frac{1}{3}$ and $\frac{1}{\mu_2} = \frac{1}{\mu_3} = \frac{1}{2}$, respectively. Each job must sequentially undergo service in Q_1, Q_2 and Q_3, but sometimes the service in Q_1 is not successful, in which case an additional service must be performed at Q_0. A service at Q_1 is successful with probability $\frac{2}{3}$, in which case the job is forwarded to Q_2, then (always) to Q_3 and then (always) back to Q_1. If the service at Q_1 is not successful, then the job is sent to Q_0 and from there, after receiving service, returned to Q_1 in order to be processed there again.

(a) Determine the relative number of visits to each of the four queues.
(b) Determine the joint equilibrium distribution of the numbers of jobs at the four queues (including jobs possibly in service)? Also indicate how the normalizing constant can be computed.
(c) Argue without any calculations that only Q_0 saturates, if the number of jobs in the network grows large ($K \to \infty$), i.e., argue that in that case only in Q_0 the (mean) number of jobs tends to infinity.
Hint: compare the (relative) loads of the queues.
(d) Choose $K = 2$. Use the *Mean-Value Analysis* algorithm to determine the mean number of jobs at each of the four queues (including jobs possibly in service).

Exercise 32 A total of K jobs circulate in a closed network of four queues (Q_1, Q_2, Q_3, Q_4), each with a single server. A job completed at Q_1 always moves to Q_2. Upon service completion at Q_2, a job moves with probability $\frac{1}{2}$ to Q_1, with probability $\frac{1}{4}$ to Q_3, and with probability $\frac{1}{4}$ to Q_4. Upon service completion at Q_3 or Q_4, jobs always return (with probability 1) to Q_2. At each of the four queues, the jobs are served in order of arrival, and the service times are exponentially distributed with means 12 ms (milliseconds), 2 ms, 8 ms and 8 ms, respectively.

(a) Use the so-called traffic equations to determine the relative number of visits to each of the four queues.
(b) Determine the joint equilibrium distribution of the numbers of jobs at the four queues (including jobs possibly in service).

For parts (c) and (d) we assume that $K = 2$ (a total of two jobs in the network).

(c) Determine the probability distribution of the number of jobs present at Q_1 (including the job possibly in service), and use this to determine the throughputs at Q_1 and Q_2.
(d) Use the *Mean-Value Analysis* algorithm to determine the mean number of jobs at each of the four queues (including jobs possibly in service).