1. Exercise 8.9: You would like to write a simulation that uses exponentially distributed random variables. Your system has a random number generator that provides independent, uniformly distributed numbers from the real interval $(0,1)$. Give a procedure that transforms a uniform random number as given to an exponentially distributed random variable with parameter $\lambda$.

In general, if we want to simulate a distribution with distribution function $F$ by using a uniform random variable $Z \sim U(0,1)$, then we can return such $X$ that $F(X) = Z$. Then it follows that

$$\Pr(X < x) = \Pr(F(X) \leq F(x)) = \Pr(Z \leq F(x)) = F(x)$$

and thus the distribution function of $X$ is $F$. The first equation holds since $F$ is non-negative and the third equation follows from the distribution of $Z$.

For exponential distribution with parameter $\lambda$ we have $F(x) = 1 - e^{-\lambda x}$. Therefore, by setting $X = F^{-1}(Z) = \frac{\ln(1-Z)}{-\lambda}$ the result $X$ is distributed as desired. Since $Z \sim 1 - Z$ when $Z \sim U(0,1)$, we can instead use a simpler form $X = \ln Z - \lambda$.

2. Exercise 8.13: A digital camera needs two batteries. You buy a pack of $n$ batteries, labeled 1 to $n$. Initially, you install batteries 1 and 2. Whenever a battery is drained, you immediately replace the drained battery with the lowest numbered unused battery. Assume that each battery lasts for an amount of time that is exponentially distributed with mean $\mu$ before being drained, independent of all other batteries. Eventually, all the batteries but one will be drained.

a) Find the probability that the battery numbered $i$ is the one that is not eventually drained.

By the memorylessness property, every time a battery is replaced both the new and the old battery last an $\text{Exp}(1/\mu)$-distributed amount of time independently of each others and the past. By symmetry (or by Lemma 8.5), the probability that a fixed one of those two batteries lasts longer is $1/2$. For the battery numbered $i > 1$ to be the last one standing it must last longer than the other battery $n - i + 1$ times and $n - 1$ times for $i = 1$. The probability that this happens is

$$\Pr(\text{battery numbered } i \text{ is not drained}) = \begin{cases} 2^{i-n-1}, & \text{for } i > 1, \\ 2^{1-n}, & \text{for } i = 1. \end{cases}$$

b) Find the expected time your camera will be able to run with this pack of batteries.

Each moment the camera has two batteries, the lasting time of both being independently $\text{Exp}(1/\mu)$-distributed. By Lemma 8.5, the time before the first one of the batteries is drained is exponentially distributed with parameter $1/\mu + 1/\mu = 2/\mu$, the expected time being $\mu/2$. The camera runs out of batteries when this has happened $n - 1$ times. By linearity of expectation, the expected total running time is thus $(n - 1)\mu/2$.

3. Exercise 8.15:

a) Let $X_1, X_2, \ldots$ be a sequence of independent exponential random variables, each with mean 1. Given a positive real number $k$, let $N$ be defined by

$$N = \min \left\{ n : \sum_{i=1}^{n} X_i > k \right\}.$$

That is, $N$ is the smallest number for which the sum of the first $N$ of the $X_i$ is larger than $k$. Determine $E[N]$. 
By Theorem 8.11, the sequence $X_1, X_2, \ldots$ of exponential interarrival times defines a poisson process $\{M(t) \mid t \geq 0\}$ with rate $\lambda = 1$ and $M(t) \sim \text{Poisson}(t)$. But now we have that

$$N = M(k) + 1,$$

since by the definition of $M(k)$ exactly this many arrivals is the minimum that exceeds time $k$. Thus, we get

$$E[N] = E[M(k)] + 1 = k + 1.$$

b) Let $X_1, X_2, \ldots$ be a sequence of independent uniform random variables on the interval $(0, 1)$. Given a positive real number $k$ with $0 < k < 1$, let $N$ be defined by

$$N = \min \left\{ n : \prod_{i=1}^{n} X_i < k \right\}.$$

That is, $N$ is the smallest number for which the product of the first $N$ of the $X_i$ is smaller than $k$. Determine $E[N]$. (Hint: You may find Exercise 8.9 helpful.)

Let $Y_i = -\ln X_i$ for $i = 1, 2, \ldots$. Then $Y_i \sim \text{Exp}(1)$ as we saw in Exercise 8.9. But now taking the logarithm both sides we observe that the inequation $\prod_{i=1}^{n} X_i < k$ is equivalent to $\sum_{i=1}^{n} Y_i > -\ln k$ and thus

$$N = \min \left\{ n : \sum_{i=1}^{n} Y_i > -\ln k \right\}.$$

By using the result from the first part the expectation is then

$$E[N] = -\ln k + 1.$$

4. **Exercise 8.16:** There are $n$ tasks that are given to $n$ processors. Each task has two phases, and the time for each phase is given by an exponentially distributed random variable with parameter 1. The times for all phases and for all tasks are independent. We say that a task is half-done if it has finished one of its two phases.

a) Derive an expression for the probability that there are $k$ tasks that are half-done at the instant when exactly one task becomes completely done.

From independency and memoryless distribution it follows that the probability, that the phase finished next belongs to a given task, is always the same for all (remaining) tasks. We can model the situation using balls and bins: $i$th ball goes to bin $j$ if the phase that is finished $i$th belongs to task $j$. We are asked to find the probability of the event that the first $k+1$ balls go to different bins and then the $k+2$:th ball goes to one of the $k+1$ nonempty bins. The probability of this is

$$p_k = \frac{1}{n} \cdot \left( 1 - \frac{1}{n} \right) \left( 1 - \frac{2}{n} \right) \cdots \left( 1 - \frac{k}{n} \right) \frac{k+1}{n} = \frac{n!(k+1)}{(n-k-1)!n^{k+2}}.$$

Using the same approach as in Section 5.1 of the book we can approximate the above:

$$p_k \approx \frac{k+1}{n} \exp \left( -\frac{k^2}{2n} \right).$$

This is maximized when $k \approx \sqrt{n}$, which is also approximately the expectation of the exact distribution (proof omitted).

b) Derive an expression for the expected time until exactly one task becomes completely done.

Above we saw that the probability that there are $k$ half-done tasks when the first task is finished is $p_k$. If this happens, then exactly $k + 2$ phases must have completed. The expected time before $k + 2$ phases are
completed is \((k + 2)/n\) since we select \(k + 2\) times the minimum of \(n\) exponentially distributed variables. The total expected time is therefore

\[
\sum_{k=0}^{n-1} \frac{k+2}{n} p_k = \sum_{k=0}^{n-1} \frac{n!(k+1)(k+2)}{(n-k-1)!n^{k+3}}.
\]

Asymptotically this decreases like \(1/\sqrt{n}\).

Another way to approach this is as follows. Let \(X\) be the time that is needed before the first task becomes completely done and let \(N_i(t)\) be the number phases completed for \(i\)th task at time \(t\). Now \(X > t\) if and only if \(N_i(t) \leq 1\) for all \(i\). Since the times for all phases are independent \(\text{Exp}(1)\) random variables, \(N_i(t)\)s are independent Poisson processes with rate \(\lambda = 1\). Therefore,

\[
\Pr(X > t) = \Pr(N_1(t) \leq 1, N_2(t) \leq 1, \ldots, N_n(t) \leq 1) = \Pr(N_i(t) \leq 1)^n.
\]

\[
= \left( \Pr(N_i(t) = 0) + \Pr(N_i(t) = 1) \right)^n
= \left( e^{-t^1} + e^{-t^1} \right)^n = e^{-tn(t+1)^n}
\]

Since \(X\) is continuous, by using Lemma 8.1, the expectation of \(X\) is

\[
E[X] = \int_0^\infty \Pr(X \geq t) \, dt = \int_0^\infty \Pr(X > t) \, dt = \int_0^\infty e^{-tn(t+1)^n} \, dt
\]

By substituting \(t = s/n - 1\) we have \(s = n(t+1)\) and \(dt = ds/n\) and thus

\[
E[X] = \int_n^\infty e^{-n(\frac{s}{n} - 1)} \frac{s^n}{n} \, ds
= \frac{e^n}{n^{n+1}} \int_n^\infty e^{-s} s^n \, ds
= \frac{e^n}{n^{n+1}} \Gamma(n + 1, n),
\]

where \(\Gamma(a, x)\) is the upper incomplete gamma function.

c) Explain how this problem is related to the birthday paradox.

The balls and bins analogy relates this directly to the birthday paradox: bins (tasks) correspond to different days and balls (completing phases) correspond to persons. When the first bin gets second ball (the first task is fully completed) there are approximately \(\sqrt{n}\) bins with ball (half-done tasks).

5. **Exercise 8.19:** You are waiting at the bus stop to catch a bus across town. There are actually \(n\) different bus lines you can take, each following a different route. Which bus you decide to take will depend on which bus gets to the bus stop first. As long as you are waiting, the time you have to wait for a bus on the \(i\)th line is exponentially distributed with mean \(\mu_i\) minutes. Once you get on a bus on the \(i\)th line, it will take you \(t_i\) minutes to get across town.

Design an algorithm for deciding – when a bus arrives – whether or not you should get on the bus, assuming your goal is to minimize the expected time to cross town. (Hint: You want to determine the set of buses that you want to take as soon as they arrive. There are \(2^n\) possible sets, which is far too large for an efficient algorithm. Argue that you need only consider a small number of these sets.)

Let \(S = \{1, \ldots, n\}\) be the set of all buses and let \(B \subseteq S\) be the set of buses that you take if they arrive. We want to find the set \(B\) that minimizes the expected total traveling time \(E[T_B]\)
The total traveling time consists of waiting time and the time spent in the bus. The waiting time \( W_i \) for bus line \( i \) is \( \text{Exp}(\lambda_i) \) distributed, where \( \lambda_i = 1/\mu_i \). Assume that we have fixed \( B \). Since the waiting times for different lines are independent, by Lemma 8.5 the time before the first of these arrives is \( W_B = \min_{i \in B} W_i \sim \text{Exp}(\sum_{i \in B} \lambda_i) \). The expected time before the first of the buses in \( B \) arrives is thus \( 1/\sum_{i \in B} \lambda_i \). Furthermore, by Lemma 8.5, the probability that the first arriving bus is \( j \) is \( \lambda_j/\sum_{i \in B} \lambda_i \). Denoting by \( R_B \) the time spent riding the bus, the expected total time is

\[
E[T_B] = E[W_B] + E[R_B] = \frac{1}{\sum_{i \in B} \lambda_i} + \sum_{j \in B} \frac{\lambda_j}{\sum_{i \in B} \lambda_i} t_j = \frac{1 + \sum_{i \in B} t_i \lambda_i}{\sum_{i \in B} \lambda_i}.
\]

But how to find the optimal \( B \) without explicitly testing all possible choices? Assume that the buses are sorted by traveling time, that is, \( t_1 \leq t_2 \leq \cdots \leq t_n \). Now clearly if \( i \in B \), then also \( j \in B \) for all \( j < i \). (If you take bus \( i \) if it arrives first, then of course you should also take a faster bus \( j < i \) if that happens to arrive first.) Therefore, \( B = \{1, \ldots, k\} \) for some \( k \in \{1, \ldots, n\} \) and only \( n \) different sets need to be considered.

Finally, it is good to note that the following greedy algorithm does not work: when bus \( i \) arrives, you take it if \( t_i < \mu_j + t_j \) for all \( j \neq i \). Why does not this work? While the expected total time for any single bus is longer than the traveling time for \( i \), it is likely that some of the other buses arrives quicker than expected. That is, the expected waiting time for any other bus is smaller than for the fastest single bus. The latter is \( \min_{j \neq i} t_j \) while the former is \( 1/\sum_{j \neq i} (1/\mu_j) \).

6. Exercise 8.20: Given a discrete space, continuous time Markov process \( X(t) \), we can derive a discrete time Markov chain \( Z(t) \) by considering the states the process visits. That is, let \( Z(0) = X(0) \), let \( Z(1) \) be the state that process \( X(t) \) first moves to after time \( t = 0 \), let \( Z(2) \) be the next state process \( X(t) \) moves to, and so on. (If Markov process \( X(t) \) makes a transition from state \( i \) to state \( k \), which can occur when \( p_{ij} \neq 0 \) in the associated transition matrix, then the Markov chain \( Z(t) \) should also make a transition from state \( i \) to state \( k \).)

a) Suppose that, in the process \( X(t) \), the time spent in state \( i \) is exponentially distributed with parameter \( \theta_i = \theta \) (which is the same for all \( i \)). Further suppose that the process \( X(t) \) has a stationary distribution. Show that the Markov chain \( Z(t) \) has the same stationary distribution.

Let \( \mathbf{P} = [p_{ij}] \) be the transition matrix of \( Z(t) \) and let \( \pi \) be the stationary distribution of \( X(t) \). By definition \( p_{ij} = p_{i,j} \). Using this, and the fact that \( \theta_i = \theta \) to the rate equations (equation (8.4) on page 212 of the book) we get that

\[
\pi_i \theta_i = \sum_k \pi_k \theta_k p_{k,i} \iff \pi_i = \sum_k \theta_k p_{k,i}
\]

for all \( i \). That is, \( \pi = \pi \mathbf{P} \), which means that \( \pi \) is the stationary distribution of \( Z(t) \).

b) Give an example showing that, if the \( \theta_i \) are not all equal, then the stationary distributions for \( X(t) \) and \( Z(t) \) may differ.

Let the transition matrix of \( X(t) \) (and \( Z(t) \)) be

\[
\mathbf{P} = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}
\]

and let \( \bar{\theta} = (1, 2) \). Now the stationary distribution of the embedded Markov chain \( Z \) is \( \pi = (1/2, 1/2) \), but the solution to the rate equation \( \pi' \theta_i = \sum_k \pi_k \theta_k p_{k,i} \) of \( X \) is \( \pi' = (2/3, 1/3) \).

7. Exercise 8.22: We can obtain a discrete time Markov chain from the \( M/M/1 \) queueing process in the manner described in Exercise 8.20. The discrete time chain tracks the number of customers in the queue. It is useful to allow departure events to occur with rate \( \lambda \) at the queue even when it is empty; this does not affect the queue behavior, but it gives transitions from state 0 to state 0 in the corresponding Markov chain.

a) Describe the possible transitions of this discrete-time chain an give their probabilities.
The states of the chain correspond to the number of customers in the queue. From state $i$ there are possible transitions to states $i-1$ (departure) and $i+1$ (arrival). The exception is state 0 from which there are transitions to 1 (arrival) and 0 (fake departure). The probabilities of arrival and departure are $\lambda/((\lambda+\mu)$ and $\mu/((\lambda+\mu)$ respectively. The transition probability matrix is therefore

$$
P = \frac{1}{\lambda+\mu} \begin{bmatrix} \mu & \lambda & 0 & 0 & 0 \\
\mu & 0 & \lambda & 0 & 0 \\
0 & \mu & 0 & \lambda & 0 \\
0 & 0 & \mu & 0 & \lambda \\
\vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.
$$

b) Show that the stationary distribution of this chain when $\lambda < \mu$ is the same as for the $M/M/1$ process.

The stationary distribution of the chain must satisfy $\pi_i = \sum_k \pi_k p_{i,k}$, that is

$$
\pi_0 = \frac{1}{\lambda+\mu} (\pi_0 \mu + \pi_1 \mu)
$$

and

$$
\pi_i = \frac{1}{\lambda+\mu} (\pi_{i-1} \lambda + \pi_{i+1} \mu) \quad \text{for } i > 0.
$$

From these it follows that

$$
\pi_i = \pi_0 \left( \frac{\lambda}{\mu} \right)^i
$$

(see the book, section 8.6.1, pages 213–214). In addition we know that $\sum_i \pi_i = 1$. Given $\lambda < \mu$, the solution to these equations is

$$
\pi_i = \left( 1 - \frac{\lambda}{\mu} \right) \left( \frac{\lambda}{\mu} \right)^i.
$$

But this is exactly the same as for the $M/M/1$ process.

c) Show that, in the case $\lambda = \mu$, there is no valid stationary distribution for the Markov chain.

From (1) we get that $\pi_i = \pi_0$ for all $i$, which implies that the stationary distribution should be uniform. But then $\sum_{i=0}^{\infty} \pi_i = \infty$ for $\pi_i > 0$ and $\sum_{i=0}^{\infty} \pi_i = 0$ for $\pi_i = 0$, which means that no such distribution exists.