Problem 11.1

Two thimbles (like tiny cups) are under a dripping roof. At the end of each second, thimble A receives 1 drop of water with probability 1, and thimble B receives 1 drop with probability 2/3 and 0 drops otherwise. By a complicated automatic mechanism, right before a 4th drop lands in thimble A, both thimbles are emptied. While the thimbles are being emptied, they miss catching the drops that would have otherwise landed inside the thimbles.

(a) Set up a Markov chain model, and draw the state transition diagram, and indicate the transition probabilities.

(b) If both thimbles were empty when you started watching, what is the probability that both thimbles contain exactly 1 drop after exactly 10,001 seconds?

Solutions:

1. The state of the system has to capture all the relevant information about the system. In this case, this is the number of drops that are in thimble A and in thimble B. Since the state must capture all the relevant information, we cannot say that it is just the number of drops in A or just the number of drops in B. Instead, we must think of the state of the system as the combined description of the number of drops in A and the number of drops in B.

Since at every point in time we get a drop in A, we know that the number of drops in B must be less than or equal to the number of drops in A since B sometimes gets a drop and sometimes does not. So, let us denote the state of the system as a pair \((x, y)\) where \(x\) represents the number of drops in A, and \(y\) represents the number of drops in B.

The automatic mechanism basically implies that the maximum number of drops in A is 3. And since B receives a drop at any time with probability \(\frac{2}{3}\), we get the following Markov chain.
2. Notice that this Markov chain is periodic with period 4. A periodic Markov chain has no steady-state probabilities. We are asked to find the probability that there is exactly one drop in both thimbles after exactly 10,001 seconds when we started observing when both of them were empty.

Relating the above word explanation to states in our Markov chain, we are asked to find the probability that we end up in state (1, 1) after 10,001 transitions if we started at state (0, 0). Notice that after 10,001 transitions, we will be in either state (1, 0) or in state (1, 1). So the probability we end up in state (1, 1) is $\frac{2}{3}$.

**Problem 11.2**

At the UC Berkeley book store, there is only one cashier. Due to the limited space, she allows only $M$ customers to line up before her at any time. If a customer finds there are $M$ customers there including the one being served by the cashier, he will leave the book store immediately. Every minute, exactly one of the following occurs: (i) one new customer arrives with probability $p$; or (ii) one existing customer leaves with probability $kq$, where $k$ is the number of customers in the book store; or (iii) no new customer arrives and no existing customer leaves with probability $1 - p - kq$ if there is at least one customer in the book store and with probability $1 - p$ otherwise.

(a) This problem can be modeled as a birth-death process (see end of §6.3). Define appropriate states and draw the transition probability graph.

(b) After the book store has been open for a long time, you walk into the book store. Calculate how many customers you expect to see in line.

**Solutions:**
1. We define a Markov chain with states 0, 1, · · · , M, corresponding to the number of customers in the house. Assume that $Mq < 1$, the transition probability graph is given as follows,

2. For the above Markov Chain, the local balance equations are

$$\pi_i p = \pi_{i+1} (i+1)q, \quad i = 0, 1, \cdots, M - 1.$$ 

We define $\rho = p/q$, and obtain $\pi_{i+1} = \frac{\rho^{i+1}}{i+1} \pi_i$, which leads to

$$\pi_i = \frac{\rho^i}{i!} \pi_0, \quad i = 0, 1, \cdots, M - 1.$$ 

By using the normalization equation, $1 = \pi_0 + \pi_1 + \cdots + \pi_M$, we obtain

$$1 = \pi_0 (1 + \frac{\rho^1}{1!} + \frac{\rho^2}{2!} + \cdots + \frac{\rho^M}{M!}),$$

and

$$\pi_0 = \frac{1}{\sum_{k=0}^{M} \frac{\rho^k}{k!}}.$$ 

Using the equation $\pi_i = \frac{\rho^i}{i!} \pi_0$, the steady-state probabilities are

$$\pi_i = \frac{\frac{\rho^i}{i!}}{\sum_{k=0}^{M} \frac{\rho^k}{k!}},$$

Therefore, the average number of customers in the house is given by

$$\bar{N} = \sum_{i=0}^{M} i \pi_i = \rho \frac{\sum_{i=0}^{M-1} \frac{\rho^i}{i!}}{\sum_{k=0}^{M} \frac{\rho^k}{k!}}.$$ 

**Problem 11.3**

Sam and Pat are playing foosball. When they begin, the score is 0-0. To make things interesting, if the score ever becomes tied, it is instantly reset to 0-0. Starting from any score, the probability that Sam gets the next point is $\frac{1}{3}$.

(a) Suppose the game stops when one player’s score reaches 2. Draw an appropriate Markov chain that describes the game, and identify all transient, recurrent, and periodic states. Find $P[\text{Pat wins}]$. 

3
(b) Now suppose instead that the game stops when a total of 3 points have been scored (note that this stopping condition does not explicitly depend on the score). The score still resets to 0-0 when the game is tied, and the player with the most points when the game ends wins. Draw an appropriate Markov chain that describes the game.

**Solutions:**

1. (a) Denote by \((x, y)\) the score of Sam and Pat respectively, a Markov chain that describes the game is

\[
\begin{array}{c}
1/3 & 2/3 & 2/3 & 2/3 \\
1/3 & 0/0 & 0/0 & 0/0 \\
1/3 & 1/3 & 0/1 & 0/2 \\
1/3 & 1/3 & 1/3 & 0/3 \\
\end{array}
\]

Note that the game ends when either state \((0, 2)\) or \((2, 0)\) is entered.

(b) Since we have a finite number of states, a state is recurrent if and only if it is accessible from all the states that are accessible from it, and therefore, states \((0, 0)\), \((0, 1)\) and \((1, 0)\) are transient and states \((0, 2)\) and \((2, 0)\) are recurrent.

(c) The probability that Pat wins is the probability that we get absorbed to state \((0, 2)\). Setting up the equations, we solve for \(a_{(1,0)}\), \(a_{(0,0)}\) and \(a_{(0,1)}\)

\[
\begin{align*}
a_{(0,1)} &= \frac{2}{3} + \frac{1}{3}a_{(0,0)} \\
a_{(0,0)} &= \frac{2}{3}a_{(1,0)} + \frac{1}{3}a_{(1,0)} \\
a_{(1,0)} &= \frac{2}{3}a_{(0,0)}
\end{align*}
\]

which yields that following

\[
\begin{align*}
a_{(0,1)} &= \frac{14}{15} \\
a_{(0,0)} &= \frac{12}{15} \\
a_{(1,0)} &= \frac{8}{15}
\end{align*}
\]

Therefore the probability of Pat winning is equal to \(12/15 = 0.8\).

2. The question can be interpreted in two ways: if the score is reset to 0-0 when the game is tied, then an appropriate Markov chain is
If the score is not reset, then the chain would look like

![Markov Chain Diagram]

Problem 11.4
Let $X$ be the height in meters of a randomly selected Canadian. Bo is interested in estimating $h = E[X]$. Being sure that no Canadian is taller than 3 meters, Bo decides to use 1.5 meters as a conservative (large) value for the standard deviation of $X$. To estimate $h$, Bo computes the average $H$ of the heights of $n$ Canadians that he selects at random.

(a) Compute $E[H]$ and $\text{var}(H)$ in terms of $h$ and Bo’s 1.5 meter bound for $\text{std}(X)$.

(b) Compute the minimum value of $n$ (with $n > 0$) such that the standard deviation of $H$ will be less than 0.01 meters.

(c) Say Bo would like to be 99% sure that his estimate is within 5 centimeters of the true average height of Canadians. Using the Chebyshev inequality, calculate the minimum value of $n$ required.

(d) If we agree that no Canadians are taller than three meters, why is it correct to use 1.5 meters as an upper bound on the standard deviation for $X$, the height of any Canadian selected at random?

Solutions:
Note that $n$ is deterministic and $H$ is a random variable.

1. Use $X_1, X_2, \ldots$ to denote the (random) measured heights.

$$H = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

$$E[H] = \frac{E[X_1 + X_2 + \cdots + X_n]}{n} = \frac{nE[X]}{n} = h$$

$$\sigma_H = \sqrt{\text{var}(H)} = \sqrt{\frac{n \text{var}(X)}{n^2}} = \frac{1.5}{\sqrt{n}}$$

(var of sum of independent r.v.s is sum of vars)
2. We solve $\frac{1.5}{\sqrt{n}} < 0.01$ for $n$ to obtain $n > 22500$.

3. Apply the Chebyshev inequality to $H$ with $\mathbb{E}[H]$ and $\text{var}(H)$ from part (a):

$$
\mathbb{P}(|H - h| \geq t) \leq \left( \frac{\sigma_H}{t} \right)^2
$$

$$
\mathbb{P}(|H - h| < t) \geq 1 - \left( \frac{\sigma_H}{t} \right)^2
$$

To be “99% sure” we require the latter probability to be at least 0.99. Thus we solve

$$
1 - \left( \frac{\sigma_H}{t} \right)^2 \geq 0.99
$$

with $t = 0.05$ and $\sigma_H = \frac{1.5}{\sqrt{n}}$ to obtain

$$
n \geq \left( \frac{1.5}{0.05} \right)^2 \frac{1}{0.01} = 90000.
$$

4. The variance of a random variable increases as its distribution becomes more spread out. In particular, if a random variable is known to be limited to a particular closed interval, the variance is maximized by having 0.5 probability of taking on each endpoint value. In this problem, random variable $X$ has an unknown distribution over $[0, 3]$. The variance of $X$ cannot be more than the variance of a random variable that equals 0 with probability 0.5 and 3 with probability 0.5. This translates to the standard deviation not exceeding 1.5.

In fact, this argument can be made more rigorous as follows.

First, we have

$$
\text{var}(X) \leq \mathbb{E}[(X - \frac{3}{2})^2] = \mathbb{E}[X^2] - 3\mathbb{E}[X] + \frac{9}{4}
$$

since $\mathbb{E}[(X - a)^2]$ is minimized when $a$ is the mean (i.e., the mean is the least-squared estimator).

Second, we also have

$$
0 \leq \mathbb{E}[X(3 - X)] = \mathbb{E}[X] - \mathbb{E}[X^2]
$$

since the variable has support in $[0, 3]$. Adding the above two inequalities, we have

$$
\text{var}(X) \leq \frac{9}{4}
$$

or equivalently, $\sigma_X \leq \frac{3}{2}$.

**Problem 11.5**

Consider a normal variable $Z \sim N(\mu, \sigma^2)$.

(a) Use the Chernoff approach to show that $\ln \mathbb{P}[Z \geq \mu + \epsilon] \leq \min_{t>0} \left[ \frac{\sigma^2 t^2}{2} - \epsilon t \right]$ for any $\epsilon > 0$. Conclude that $\mathbb{P}[Z \geq \mu + \epsilon] \leq \exp(-\epsilon^2 / (2\sigma^2)$.
(b) Let \( Z_1, Z_2, \ldots \) be independent and identically distributed as \( Z(0, \sigma^2) \). Show that the probability \( \mathbb{P}[\max\{Z_1, \ldots, Z_n\} \geq \sqrt{3\sigma^2 \log n}] \to 0 \), as \( n \to +\infty \). (Hint: Union bound combined with part (a) could be useful.)

**Solutions:**

(a) Let \( X = e^{sZ} \) for \( s > 0 \). Because \( X \) is a strictly increasing function of \( Y \), the following probability must be true:

\[
\mathbb{P}(Z \geq \mu + \epsilon) = \mathbb{P}(X \geq e^{s(\mu + \epsilon)})
\]

Apply Markov inequality on the right-hand-side.

\[
\mathbb{P}(X \geq e^{s(\mu + \epsilon)}) \leq \frac{\mathbb{E}(X)}{e^{s(\mu + \epsilon)}} = \frac{M_Y(s)}{e^{s(\mu + \epsilon)}} = e^{\sigma^2 s^2/2 - s\epsilon}
\]

In the last expression, we can choose its minimum and the inequality still holds, which is

\[
\ln \mathbb{P}(Z \geq \mu + \epsilon) \leq \min_{t>0} \left[ \frac{\sigma^2 t^2}{2} - t\epsilon \right]
\]

Moreover, the minimizer of the quadratic function is actually \( t = \frac{\epsilon}{\sigma^2} \) and the corresponding minimal value is \(-\frac{\epsilon^2}{2\sigma^2}\). It leads to

\[
\mathbb{P}(Z \geq \mu + \epsilon) \leq \exp\left( -\frac{\epsilon^2}{2\sigma^2} \right)
\]

(b)

\[
\mathbb{P}(\max(Z_1, \ldots, Z_n) \geq \sqrt{3\sigma^2 \log n}) = 1 - (1 - \mathbb{P}(Z_1 \geq \sqrt{3\sigma^2 \log n}))^n \\
\leq 1 - (1 - \exp\left( -\frac{3\sigma^2 \log n}{2\sigma^2} \right))^n \\
= 1 - (1 - n^{-1.5})^n \\
= 1 - (1 - n^{-1.5})^{n^{1.5} \times n^{-0.5}}
\]

where the limit of \((1 - n^{-1.5})^{n^{1.5}}\) is \(1/e\) when \( n \to +\infty \). And

\[
\lim_{n \to +\infty} (1/e)^{n^{-0.5}} = (1/e)^0 = 1
\]

Thus,

\[
\lim_{n \to +\infty} \mathbb{P}(\max(Z_1, \ldots, Z_n) \geq \sqrt{3\sigma^2 \log n}) = 0
\]
Problem 11.6
We are laying out 25 plastic planks lengthwise, trying to make a path of about 1000 meters. The plastic planks are made in molds, and any variation in the lengths of the planks is due entirely to variation between different molds. The length in meters, $X$, of any particular mold used for making planks is independent of the length of all other molds. $X$ is uniformly distributed between $40 - \sqrt{3}$ and $40 + \sqrt{3}$ meters. $X$ has an expected value of 40 meters and a standard deviation of 1 meter. Assume every plank has exactly the same length as its mold. What is the probability that the resulting path will be within $1000 \pm 7.5$ meters if we use 25 planks:
(a) all made from the same mold?
(b) each made from a different mold?

Explain the difference between the answers to parts (a) and (b).

Solutions:

1. When using just one mold, the length of the path is $25X$ and the desired probability is

$$P(|25X - 1000| < 7.5) = P(|X - 40| < 0.3) = \frac{\sqrt{3}}{10} \approx 0.1732.$$

2. When using separate molds with lengths $X_1, X_2, \ldots, X_{25}$ the desired probability is

$$P\left(\left|\sum_{i=1}^{25} X_i - 1000\right| < 7.5\right) = P\left(\left|\frac{\sum_{i=1}^{25} X_i - 1000}{\sqrt{25}}\right| < \frac{7.5}{\sqrt{25}}\right)$$
$$\approx P\left|Z\right| < 1.5 \quad \text{where } Z \text{ is a standard normal r.v. (CLT)}$$
$$= \Phi(1.5) - \Phi(-1.5) = 2\Phi(1.5) - 1 \approx 0.8664.$$

Intuitively, adding independent instances of mold lengths “averages out” the variations and gives higher probability of a total path length close to the mean.