CSE 312: Foundations of Computing II

Section 7: Markov Chains, CLT Solutions

1. Faulty Machines

You are trying to use a machine that only works on some days. If on a given day, the machine is working it will break down the next day with probability $0 < b < 1$, and works on the next day with probability $1 - b$. If it is not working on a given day, it will work on the next day with probability $0 < r < 1$ and not work the next day with probability $1 - r$.

(a) In this problem we will formulate this process as a Markov chain. First, let X_t be a random variable that denotes the state of the machine at time t. Then, define a state space $\mathcal{S}$ that includes all the possible states that the machine can be in. Lastly, for all $A, B \in \mathcal{S}$ find $\mathbb{P}(X_{t+1} = A \mid X_t = B)$ (A and B can be the same state).

Solution:

Formally, a Markov chain is defined by a state space $\mathcal{S}$ and a transition probability matrix. The two possible states of the machine are "working" and "broken". So, $\mathcal{S} = \{W, B\}$. Let X_t be the state of the process at time t. Then we can define the following transition probabilities:

$$\mathbb{P}(X_{t+1} = W \mid X_t = W) = 1 - b \qquad \mathbb{P}(X_{t+1} = B \mid X_t = W) = b$$

$$\mathbb{P}(X_{t+1} = W \mid X_t = B) = r \qquad \mathbb{P}(X_{t+1} = B \mid X_t = B) = 1 - r$$

We can also represent the TPM with the following matrix:

$$P = \begin{bmatrix} 1-b & b \\ r & 1-r \end{bmatrix}$$

where the ijth entry is probability that the machine is in the jth state at time $t+1$ given it was in state i at time t. (Here state 1 is working and state 2 is broken.)

(b) Suppose that on day 1, the machine is working. What is the probability that it is working on day 3?

Solution:

We are trying to find $\mathbb{P}(X_3 = W \mid X_1 = W)$. From the law of total probability, and then plugging in the values from our transition matrix:

$$\begin{aligned}
P(X_3 = W \mid X_1 = W) &= \sum_{i \in \mathcal{S}} \mathbb{P}(X_3 = W \mid X_1 = W, X_2 = i)\, \mathbb{P}(X_2 = i \mid X_1 = W) \\
&= \mathbb{P}(X_3 = W \mid X_2 = W)\, \mathbb{P}(X_2 = W \mid X_1 = W) + \mathbb{P}(X_3 = W \mid X_2 = B)\, \mathbb{P}(X_2 = B \mid X_1 = W) \\
&= \mathbb{P}(X_3 = W \mid X_2 = W)\, (1 - b) + \mathbb{P}(X_3 = W \mid X_2 = B)\, b \\
&= (1 - b)(1 - b) + rb \\
&= (1 - b)^2 + rb
\end{aligned}$$

(c) As $n \to \infty$, what does the probability that the machine is working on day n converge to? To get the answer, solve for the *stationary distribution*.

Solution:

The stationary distribution is the row vector $\pi = [\pi_W \ \pi_B]$ such that $\pi P = \pi$. The entries in the vector π_W and π_B can be interpreted as the probabilities that the machine works or is broken converge to. As such, $\pi_W + \pi_B = 1$. Additionally, multiplying the stationary distribution by the TPM gives us the following two equations:

$$\pi_W = \pi_W(1-b) + \pi_B r \qquad \pi_B = \pi_W b + \pi_B(1-r)$$

Solving each for π_W and π_B gives us the following solutions for the stationary distribution:

$$\pi_W = \frac{r}{b+r} \qquad \pi_B = \frac{b}{b+r}$$

So, as $n \to \infty$ the probability that the machine works on day n is $\pi_W = \frac{r}{b+r}$

2. Another Markov chain

Suppose that the following is the transition probability matrix for a 4 state Markov chain (states 1,2,3,4).

$$P = \begin{bmatrix} 0 & 1/2 & 1/2 & 0 \\ 1/3 & 0 & 0 & 2/3 \\ 1/3 & 1/3 & 0 & 1/3 \\ 1/5 & 2/5 & 2/5 & 0 \end{bmatrix}$$

(a) What is the probability that $X_2 = 4$ given that $X_0 = 4$?

Solution:

Let's denote the state space $\mathcal{S} = \{1, 2, 3, 4\}$. Using the law of total probability we can determine that

$$\begin{aligned} \mathbb{P}(X_2 = 4 \mid X_0 = 4) &= \sum_{i \in \mathcal{S}} \mathbb{P}(X_2 = 4 \mid X_0 = 4, X_1 = i)\, \mathbb{P}(X_1 = i \mid X_0 = 4) \\ &= \sum_{i \in \mathcal{S}} \mathbb{P}(X_2 = 4 \mid X_1 = i)\, \mathbb{P}(X_1 = i \mid X_0 = 4) \\ &= 0 + \frac{2}{5} \cdot \frac{2}{3} + \frac{2}{5} \cdot \frac{1}{3} + 0 \\ &= \frac{2}{5} \end{aligned}$$

(b) Write down the system of equations that the stationary distribution must satisfy and solve them.

Solution:

The stationary distribution is the row vector $\pi = [\pi_1 \ \pi_2 \ \pi_3 \ \pi_4]$ such that $\pi P = \pi$. We know that $\pi_1 + \pi_2 + \pi_3 + \pi_4 = 1$. Additionally, multiplying the stationary distribution by the TPM gives us the following equations:

$$\begin{aligned} \pi_1 &= \frac{1}{3}\pi_2 + \frac{1}{3}\pi_3 + \frac{1}{5}\pi_4 \\ \pi_2 &= \frac{1}{2}\pi_1 + \frac{1}{3}\pi_3 + \frac{2}{5}\pi_4 \\ \pi_3 &= \frac{1}{2}\pi_1 + \frac{2}{5}\pi_4 \\ \pi_4 &= \frac{2}{3}\pi_2 + \frac{1}{3}\pi_3 \end{aligned}$$

Solving for each π_i gives us the following solutions for the stationary distribution:

$$\pi_1 = \frac{46}{206} \qquad \pi_2 = \frac{60}{206} \qquad \pi_3 = \frac{45}{206} \qquad \pi_4 = \frac{55}{206}$$

3. Three tails

You flip a fair coin until you see three tails in a row. Model this as a Markov chain with the following states:

- S: start state, which we are only in before flipping any coins.
- H: We see a heads, which means no streak of tails currently exists.
- T: We've seen exactly one tail in a row so far.
- TT: We've seen exactly two tails in a row so far.
- TTT: We've accomplished our goal of seeing three tails in a row and stop flipping.

(a) Write down the transition probability matrix.

Solution:

$$P = \begin{bmatrix} 0 & 1/2 & 1/2 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

(b) Write down the system of equations whose variables are $D(s)$ for each state $s \in \{S, H, T, TT, TTT\}$, where $D(s)$ is the expected number of steps until state TTT is reached starting from state s. Solve this system of equations to find $D(S)$.

Solution:

Using the law of total expectation and the TPM above we can set up and solve the following system of equations:

$$\begin{aligned} D(TTT) &= 0 \\ D(TT) &= 1 + \frac{1}{2}D(H) + \frac{1}{2}D(TTT) = \frac{1}{2}D(H) + 1 \\ D(T) &= 1 + \frac{1}{2}D(H) + \frac{1}{2}D(TT) = \frac{3}{4}D(H) + \frac{3}{2} \\ D(H) &= 1 + \frac{1}{2}D(H) + \frac{1}{2}D(T) = \frac{7}{8}D(H) + \frac{7}{4} \\ D(S) &= 1 + \frac{1}{2}D(H) + \frac{1}{2}D(T) = \frac{7}{8}D(H) + \frac{7}{4} \end{aligned}$$

Solving for $D(H)$ gives us that $D(H) = 14$, which allows as to solve for the rest of the expected number of steps, $D(TT) = 8$, $D(T) = 12$, $D(S) = 14$. So, we expect to flip 14 coins before we flip three tails in a row.

(c) Write down the system of equations whose variables are $\gamma(s)$ for each state $s \in \{S, H, T, TT, TTT\}$, where $\gamma(s)$ is the expected number of heads seen before state TTT is reached. Solve this system to find $\gamma(S)$, the expected number of heads seen overall until getting three tails in a row.

Solution:

Like in the previous part we can use LTE and the TPM to set up and solve the follwoing system of equations:

$$\begin{aligned}
\gamma(TTT) &= 0 \\
\gamma(TT) &= 0.5\gamma(H) + 0.5\gamma(TTT) = 0.5\gamma(H) \\
\gamma(T) &= 0.5\gamma(H) + 0.5\gamma(TT) = 0.75\gamma(H) \\
\gamma(H) &= 1 + 0.5\gamma(H) + 0.5\gamma(T) = 0.875\gamma(H) + 1 \\
\gamma(S) &= 0.5\gamma(H) + 0.5\gamma(T) = 0.875\gamma(H)
\end{aligned}$$

Solving for $\gamma(H)$ gives us $\gamma(H) = 8$. This allows us to solve for the other expected values which are $\gamma(TT) = 4$, $\gamma(T) = 6$, $\gamma(S) = 7$. So, we expect to see 7 heads before we flip three tails in a row.

4. CLT example

Let X be the sum of 100 real numbers, and let Y be the same sum, but with each number rounded to the nearest integer before summing. If the roundoff errors are independent and uniformly distributed between -0.5 and 0.5, what is the approximate probability that $|X - Y| > 3$?

Solution:

Let $X = \sum_{i=1}^{100} X_i$, and $Y = \sum_{i=1}^{100} r(X_i)$, where $r(X_i)$ is X_i rounded to the nearest integer. Then, we have

$$X - Y = \sum_{i=1}^{100} X_i - r(X_i)$$

Note that each $X_i - r(X_i)$ is simply the round off error, which is distributed as $Unif(-0.5, 0.5)$. Since $X - Y$ is the sum of 100 i.i.d. random variables with mean $\mu = 0$ and variance $\sigma^2 = \frac{1}{12}$, $X - Y \approx W \sim \mathbb{N}(0, \frac{100}{12})$ by the Central Limit Theorem. For notational convenience let $Z \sim \mathbb{N}(0, 1)$

$$\begin{aligned}
\mathbb{P}(|X - Y| > 3) &\approx \mathbb{P}(|W| > 3) && \text{[CLT]} \\
&= \mathbb{P}(W > 3) + \mathbb{P}(W < -3) && \text{[No overlap between } W > 3 \text{ and } W < -3\text{]} \\
&= 2\,\mathbb{P}(W > 3) && \text{[Symmetry of normal]} \\
&= 2\,\mathbb{P}\left(\frac{W}{\sqrt{100/12}} > \frac{3}{\sqrt{100/12}}\right) \\
&\approx 2\,\mathbb{P}(Z > 1.039) && \text{[Standardize } W\text{]} \\
&= 2\,(1 - \Phi(1.039)) \approx 0.29834
\end{aligned}$$

5. Tweets

A prolific twitter user tweets approximately 350 tweets per week. Let's assume for simplicity that the tweets are independent, and each consists of a uniformly random number of characters between 10 and 140. (Note that this is a discrete uniform distribution.) Thus, the central limit theorem (CLT) implies that the number of characters tweeted by this user is approximately normal with an appropriate mean and variance. Assuming this normal approximation is correct, estimate the probability that this user tweets between 26,000 and 27,000 characters in a particular week. (This is a case where continuity correction will make virtually no difference in the answer, but you should still use it to get into the practice!).

Solution:

Let X be the total number of characters tweeted by a twitter user in a week. Let $X_i \sim Unif(10, 140)$ be the number of characters in the ith tweet (since the start of the week). Since X is the sum of 350 i.i.d. rvs with mean $\mu = 75$ and variance $\sigma^2 = 1430$, $X \approx N = \mathbb{N}(350 \cdot 75, 350 \cdot 1430)$. Thus,

$$\mathbb{P}(26,000 \leq X \leq 27,000) \approx \mathbb{P}(25,999.5 \leq N \leq 27,000.5)$$

Standardizing this gives the following formula

$$\begin{aligned}\mathbb{P}(25,999.5 \leq N \leq 27,000.5) &\approx \mathbb{P}\left(-0.3541 \leq \frac{N - 350 \cdot 75}{\sqrt{350 \cdot 1430}} \leq 1.0608\right) \\ &= \mathbb{P}\left(-0.3541 \leq Z \leq 1.0608\right) \\ &= \Phi(1.0608) - \Phi(-0.3541) \\ &\approx 0.4923\end{aligned}$$

So the probability that this user tweets between 26,000 and 27,000 characters in a particular week is approximately 0.4923.

6. Poisson CLT practice

Suppose $X_1, \ldots, X_n$ are iid Poisson(λ) random variables, and let $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$, the sample mean. How large should we choose n to be such that $\mathbb{P}(\frac{\lambda}{2} \leq \overline{X}_n \leq \frac{3\lambda}{2}) \geq 0.99$? Use the CLT and give an answer involving $\Phi^{-1}(\cdot)$. Then evaluate it exactly when $\lambda = 1/10$ using the Φ table on the last page.

Solution:

We know $\mathbb{E}[X_i] = \text{Var}(X_i) = \lambda$. By the CLT, $\overline{X}_n \approx \mathcal{N}\left(\lambda, \frac{\lambda}{n}\right)$, so we can standardize this normal approximation.

$$\mathbb{P}(\frac{\lambda}{2} \leq \overline{X}_n \leq \frac{3\lambda}{2}) \approx \mathbb{P}(\frac{-\lambda/2}{\sqrt{\lambda/n}} \leq Z \leq \frac{\lambda/2}{\sqrt{\lambda/n}}) = \Phi\left(\frac{\lambda/2}{\sqrt{\lambda/n}}\right) - \Phi\left(\frac{-\lambda/2}{\sqrt{\lambda/n}}\right)$$

$$= \Phi\left(\frac{\lambda/2}{\sqrt{\lambda/n}}\right) - \left(1 - \Phi\left(\frac{\lambda/2}{\sqrt{\lambda/n}}\right)\right) = 2\Phi\left(\frac{\lambda/2}{\sqrt{\lambda/n}}\right) - 1 \geq 0.99 \to \Phi\left(\frac{\lambda/2}{\sqrt{\lambda/n}}\right) \geq 0.995$$

$$\to \frac{\sqrt{\lambda}}{2}\sqrt{n} \geq \Phi^{-1}(0.995) \to n \geq \frac{4}{\lambda}\left[\Phi^{-1}(0.995)\right]^2$$

We have $\lambda = \frac{1}{10}$ and from the table, $\Phi^{-1}(0.995) \approx 2.575$ so that $n \geq \frac{4}{1/10} \cdot 2.575^2 = 265.225$. So $n = 266$ is the smallest value that will satisfy the condition.

7. Bad Computer (More Faulty Machines)

Each day, the probability your computer crashes is 10%, independent of every other day. Suppose we want to evaluate the computer's performance over the next 100 days.

(a) Let X be the number of crash-free days in the next 100 days. What distribution does X have? Identify $\mathbb{E}[X]$ and $Var(X)$ as well. Write an exact (possibly unsimplified) expression for $\mathbb{P}(X \geq 87)$.

Solution:

Since X counts the number of crash-free days (successes) in 100 days (trials), where each trial is a success with probability 0.9, we can see that X is binomial with $n = 100$ and $p = 0.9$, or $X \sim$ Binomial$(100, 0.9)$. Hence, $\mathbb{E}[X] = np = 90$ and $Var(X) = np(1-p) = 9$. Finally,

$$\mathbb{P}(X \geq 87) = \sum_{k=87}^{100} \binom{100}{k} (0.9)^k (1 - 0.9)^{100-k}$$

(b) Approximate the probability of at least 87 crash-free days out of the next 100 days using the Central Limit Theorem. Use continuity correction.

Important: continuity correction says that if we are using the normal distribution to approximate

$$\mathbb{P}(a \leq \sum_{i=1}^{n} X_i \leq b)$$

where $a \leq b$ are integers and the X_i's are i.i.d. **discrete** random variables, then, as our approximation, we should use

$$\mathbb{P}(a - 0.5 \leq Y \leq b + 0.5)$$

where Y is the appropriate normal distribution that $\sum_{i=1}^{n} X_i$ converges to by the Central Limit Theorem.[1] For more details see pages 209-210 in the book.

Solution:

From the previous part, we know that $\mathbb{E}[X] = 90$ and $Var(X) = 9$.

$$\begin{aligned}\mathbb{P}(X \geq 87) &= \mathbb{P}(86.5 < X < 100.5) = \mathbb{P}(\frac{86.5-90}{3} < \frac{X-90}{3} < \frac{100.5-90}{3}) \\ &\approx \mathbb{P}(-1.17 < \frac{X-90}{3} < 3.5) \approx \Phi(3.5) + \Phi(1.17) - 1 \approx 0.9998 + 0.8790 - 1 = 0.8788\end{aligned}$$

Notice that, if you had used $86.5 < X$ in place of $86.5 < X < 100.5$, your answer would have been nearly the same, because $\Phi(3.5)$ is so close to 1.

(Note: This solution requires finding $\Phi(3.5)$, which you can find on a z-table such as this one. For Psets and quizzes, please ONLY use the Z-table linked on the website.)

8. Waffles

A new diner specializing in waffles opens on our street. It will be open 24 hours a day, seven days a week. It is assumed that the inter-arrival times between customers will be i.i.d. Exponential random variables with mean 10 minutes. Approximate the probability that the 120th customer will arrive after the first 21 hours of operation.

Solution:

We will let X_k denote the waiting time between the $(k-1)^{\text{th}}$ customer and the k^{th} customer. Note that each $X_k \sim Exp(6)$ since we have a rate of 6 customers per hour to get a mean of ten minutes.
This means that $\mathbb{E}[X_k] = \frac{1}{6}$ and $Var(X_k) = \frac{1}{36}$.
Then, we will use S_{120} to be the waiting time until the 120^{th} customer. Note that:

$$S_{120} = \sum_{k=1}^{120} X_k$$

By the central limit theorem we have:

$$S_{120} \approx \mathcal{N}(120 \cdot \frac{1}{6}, 120 \cdot \frac{1}{36}) = \mathcal{N}(20, (\frac{\sqrt{120}}{6})^2)$$

[1] The intuition here is that, to avoid a mismatch between discrete distributions (whose range is a set of integers) and continuous distributions, we get a better approximation by imagining that a discrete random variable, say W, is a continuous distribution with density function

$$f_W(x) := p_W(i) \quad \text{when } i - 0.5 \leq x < i + 0.5 \text{ and } i \text{ integer}$$

We do not need to use the continuity correction, because the random variable we are approximating is continuous. So we have:

$$\begin{aligned}
\mathbb{P}(S_{120} > 21) &= \mathbb{P}(\frac{S_{120} - 20}{\frac{\sqrt{120}}{6}} > \frac{21 - 20}{\frac{\sqrt{120}}{6}}) \\
&= \mathbb{P}(Z > 0.55) \\
&= 1 - \Phi(0.55) \\
&\approx 1 - 0.71 \\
&= 0.29
\end{aligned}$$

What if we asked for an exact answer instead though? Notice that for $C \sim Gamma(120, 6)$, which describes the actual probability:

$$\mathbb{P}(C > 21) = \int_{21}^{\infty} \frac{6^{120} \cdot x^{119}}{119!} e^{-6x} dx \approx 0.285$$