Continuous Stochastic Processes - Exercise 1

Mark Sverdlov

May 2025

1 Exercise 1

Computing the transition semigroup of the described process directly is difficult. Thus, we first prove three general lemmas, and then use them to estimate the transition probability asymptotically when $t \to 0$. Finally, we use these estimates to calculate the local characteristics of the process.

Lemma 1.1. Let $X \sim \operatorname{Exp}(\lambda)$, then $P(X \leq t) = \lambda t + o(t)$ as $t \to 0$.

Proof. By direct calculation:

$$P(X \leq t) = 1 - \exp(-\lambda t) = 1 - 1 + \lambda t + o(t) = \lambda t + o(t) \tag{1}$$

where the estimate is done using Taylor's theorem. □

Lemma 1.2. Let $X \sim \operatorname{Exp}(\lambda)$ and $Y \sim \operatorname{Exp}(\mu)$ be independent variables. Then $P(X + Y \leq t) = o(t)$ as $t \to 0$.

Proof. First, by Fubini's theorem:

$$P(X + Y \leq t) = \int_0^t \mathrm{d}x \cdot \lambda x \exp(-\lambda x) \int_0^{t-x} \mathrm{d}y \cdot \mu y \exp(-\mu y) \tag{2}$$

Then, using routine calculation we arrive at the equation:

$$\int_0^t \mathrm{d}x \cdot \lambda x \exp(-\lambda x) \int_0^{t-x} \mathrm{d}y \cdot \mu y \exp(-\mu y) = 1 - \exp(-\lambda t) - \lambda \exp(-\mu t) \int_0^t \exp((\mu - \lambda) x) \, \mathrm{d}x \tag{3}$$

Here we break into two cases. If $\mu \neq \lambda$ then we calculate:

$$1 - \exp(-\lambda t) - \lambda \exp(-\mu t) \int_0^t \exp((\mu - \lambda) x) \, \mathrm{d}x = 1 - \exp(-\lambda t) - \frac{\lambda}{\mu - \lambda} \exp(-\lambda t) + \frac{\lambda}{\mu - \lambda} \exp(-\mu t) \tag{4}$$

and by utilizing Taylor's theorem three times we estimate that

$$1 - \exp(-\lambda t) - \frac{\lambda}{\mu - \lambda} \exp(-\lambda t) + \frac{\lambda}{\mu - \lambda} \exp(-\mu t) = o(t) \tag{5}$$

On the other hand, if $\lambda = \mu$, then:

$$1 - \exp(-\lambda t) - \lambda \exp(-\mu t) \int_0^t \exp((\mu - \lambda) x) \, \mathrm{d}x = 1 - \exp(-\lambda t) - \lambda \exp(-\lambda t) \, t \tag{6}$$

and again, utilizing Taylor's theorem twice, we arrive at the conclusion that:

$$1 - \exp(-\lambda t) - \lambda \exp(-\lambda t) \, t = o(t) \tag{7}$$

as required. □

Lemma 1.3. Let $X \sim \text{Exp}(\lambda)$, then $P(X > t) = 1 - \lambda t + o(t)$

Proof. First, note that $P(X > t) = \exp(-\lambda t)$. The lemma is a direct application of Taylor's theorem for $\exp(-\lambda t)$. $\square$

We now observe that for any $s \geq 0$, $t > 0$, the probability of two consecutive clock ringings in the interval $(s, s+t]$ is $o(t)$ (which is a direct application of lemma 1.2), and by routine calculation of probabilities, conditioning on the number of ringings in the interval $(s, s+t]$, we deduce:

Lemma 1.4. For any states i, j, we have:

$$P(X_{s+t} = j | X_s = i) = (1 - \lambda t) 1_{j=i} + P(j \text{ is the next state after } i) \lambda t + o(t) \tag{8}$$

Finally, utilizing the fact that the $o(t)$ is negligible in differentiation, we calculate the local characteristics exactly, and conclude:

$$A = \begin{pmatrix} -1 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 3 & -3 \end{pmatrix} \tag{9}$$

2 Exercise 2

By a theorem from class, since the chain is finite, a distribution π is stationary if and only if $\pi A = 0$. By routine calculation we deduce that $\pi = \left(\frac{3}{7}, \frac{3}{7}, \frac{1}{7}\right)$ is the unique stationary distribution of the chain.

3 Exercise 3

We present our simulation code on our Github[1]. There, we used the Gillespie algorithm to simulate 500 chains, each simulating 5 time units. Then, we sample each simulation at 200 equidistant time points in the time interval $[0, 5]$ and calculate the empirical frequency of the states at each time point along the different simulations. Since the simulations are independent and identically distributed, this results in an unbiased estimate of the marginal distribution, which is equal to $e_1 \exp(At)$ at time t. Furthermore, the exact marginal distribution of the states at each time point may be estimated when $t \to \infty$ as $\pi + O(\exp(\lambda_2 t))$, where λ_2 is the second largest eigenvalue of A which is equal to $-2 + \sqrt{3}$.

In the upper part of figure 1, we depict a trace plot of one simulation out of the 500 we simulated. In the bottom part, we show the L^2 difference between the estimated marginal distribution at each time point to the stationary distribution π that was calculated before, as well as the curve $\exp(\lambda_2 t)$, which characterizes the L^2 difference between the exact marginal distribution at time t to the stationary distribution. Keep in mind that the curve depicts an unbiased estimate of the marginal distribution, obtained by running 500 independent simulations, rather than the exact marginal distribution. Regardless, we see that the theoretical rate of convergence seems to hold even for the estimated marginal distribution as $t \to \infty$, which demonstrates the correctness of the theorems regarding the convergence rate. Of course, for practical scenarios where we use the estimated marginal distribution for some large t as an estimate of the stationary distribution, it's important to keep in mind that this estimate is separated from the stationary distribution with two degrees of separations: first - it's an estimate of the marginal distribution rather than the exact marginal distribution. Second, the marginal distribution is not exactly the stationary distribution, although it converges to it quite quickly for large t.

[1] https://github.com/MarkSverdlov/stochastic-processes-exercise-1

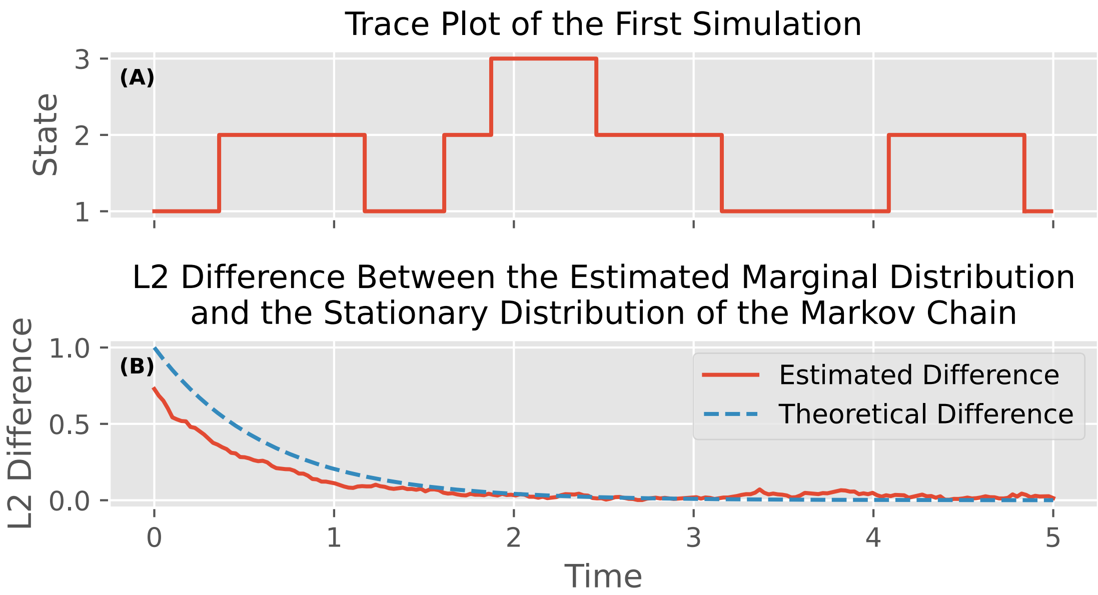

Figure 1: **(A)** depicts a trace plot of one Gillespie algorithm simulation over the Markov chain. **(B)** depicts the L^2 difference between the estimate of the marginal distribution to the stationary distribution. The dashed line shows the theoretical rate of convergence of the (exact) marginal distribution to the stationary distribution.