Week 7: Continuous-time Markov chains
Poisson process

Our goal in this exercise is to model and study the arrival of commuters to a subway station during the time interval $(0, T]$. To that end, we divide the interval in n subintervals of duration h, so that the i-th subinterval is $((i-1)h, ih]$, for $i = 1, \ldots, n$, and $T = nh$. For sufficiently small h, we can assume that there will be at most one arrival in each subinterval and that the probability of a single arrival occurring in that interval is λh. In other words, we assume that costumers do not arrive simultaneously and that the probability of arrival is proportional to the length of the interval. Formally, for $N_i(h)$ be the number of costumers that arrive in the i-th subinterval of duration h, we have

$$\mathbb{P}\left[N_i(h) = 0\right] = 1 - \lambda h, \qquad \mathbb{P}\left[N_i(h) = 1\right] = \lambda h, \quad \text{and} \quad \mathbb{P}\left[N_i(h) = k\right] = 0, \text{ for all } k > 1. \tag{1}$$

In principle, there is no reason to believe that the arrival of a customer in the i-th subinterval is independent of the arrival of a customer in the j-th time interval. Nevertheless, if the number of potential customers is very large, it is reasonable to assume that arrivals in different subintervals are independent. At the very least, it is a good approximation. We can therefore write

$$\mathbb{P}\left[N_i(h) = \ell, N_j(h) = k\right] = \mathbb{P}\left[N_i(h) = k\right] \mathbb{P}\left[N_j(h) = \ell\right], \quad \text{for } i \neq j \text{ and } k, \ell \in \{0, 1\}. \tag{2}$$

Expressions (1) and (2) define the stochastic process of arrivals of passengers to the subway station. Implicit in this definition is the assumption that T is sufficiently small so that the probability of a customer arriving does not depend on time.

We are interested in using this model to analyze two metrics: (i) the number of customers arriving by time $0 < t \leq T$, denoted $N(t) = \sum_{i=1}^{\lfloor t/h \rfloor} N_i(h)$, for $N(0) = 0$, and (ii) the time T_1 elapsed until the first customer arrives at the subway station. Notice from (1) that $N(t)$ must satisfy

$$\mathbb{P}\left[N(t+h) - N(t) = 1 \mid N(t)\right] = \lambda h.$$

Also, observe that T_1 is related to $N(t)$ since it describes the time at which $N(t)$ transitions from zero to one. Hence,

$$T_1 = \min\{t : N(t) = 1\}.$$

You will show in what follows that for sufficiently small h, T_1 is exponentially distributed with parameter λ and $N(t)$ is Poisson with parameter λt for all $t \in (0, T]$.

A Simulating $N(t)$. Write a MATLAB function that simulates this arrival process. Use $T =$ 10 minutes, $\lambda = 1$ customer per minute, and $n = 10^3$. Compare the histogram of $N(T)$ obtained from 10^4 experiments with the pmf of a Poisson with parameter λT. Also compare the histogram of $N(T)/2$ with the pmf of a Poisson with parameter $\lambda T/2$.

B The distribution of $N(t)$. The comparisons in part A should have yielded accurate fits. Discuss why this is the case in light of the Poisson approximation of the binomial distribution you saw in Homework 2. Argue that it implies the pmf of $N(t)$ is Poisson with parameter λt for all t, i.e.,

$$\mathbb{P}\left[N(t) = k\right] = e^{-\lambda t}\frac{(\lambda t)^k}{k!}. \tag{3}$$

C Simulating T_1. Using the function you wrote in part A, compute a histogram of T_1 from 10^4 experiments. Compare it with the pdf of an exponential with parameter λ.

D The distribution of T_1. You should have observed a good fit in part C. Using the fact that the probability of having no arrivals by time t can be obtained from (3) as $e^{-\lambda t}$, argue that T_1 is exponentially distributed with parameter λ.

Hint: It is easier to think about the cdf for this exercise. Indeed, notice that we have $T_1 > t$ if and only if there are no arrivals by time t, i.e., $N(t) = 0$.