Solutions of the first exam:
Stochastic Processes and Simulation II
May 30th 2023

Exercise 1: Poisson processes

(i) Give the definition of a compound Poisson process $X(t)$ and write an expression for its mean $\mathbb{E}[X(t)]$ and its variance $\mathrm{Var}(X(t))$.

Solution: A stochastic process $\{X(t), t \geqslant 0\}$ is said to be a compound Poisson process if it can be represented as

$$X(t) = \sum_{i=1}^{N(t)} Y_i, \qquad t \geqslant 0,$$

where $\{N(t), t \geqslant 0\}$ is a Poisson process, and $\{Y_i, i \geqslant 1\}$ is a family of i.i.d. random variables that is also independent of $\{N(t), t \geqslant 0\}$.
If $\{N(t), t \geqslant 0\}$ is a Poisson process with rate λ, then $\mathbb{E}[X(t)] = \lambda t\, \mathbb{E}[Y_1]$ and $\mathrm{Var}(X(t)) = \lambda t\, \mathbb{E}[Y_1^2]$.

Due to the current climate crisis, lots of families are migrating from climate affected countries seeking for better living conditions. Data show that most of the climate refugees migrates from Afghanistan, India and Pakistan. Assume that family are migrating from Afghanistan according to a Poisson process with rate 7 per day. Suppose also that the size of a family is $1, 2, 3, 4$, respectively with probability $1/8, 3/8, 3/8, 1/8$, independently for each family and from the migration process.

(ii) What is the expected value and variance of the number of refugees migrating from Afghanistan in a week?

Solution: We have that $\mathbb{E}[Y_1] = 2.5$ and $\mathbb{E}[Y_1^2] = 1\frac{1}{8} + 4\frac{3}{8} + 9\frac{3}{8} + 16\frac{1}{8} = 7$. Hence, $\mathbb{E}[X(7)] = 7 \cdot 7 \cdot 2.5 = 122.5$ and $\mathrm{Var}(X(7)) = 7 \cdot 7 \cdot 7 = 343$.

(iii) Using the central limit theorem, find the approximate probability that at least 1800 refugees migrate from Afghanistan in the next 100 days. The answer can be given in terms of $\phi(x) = \mathbb{P}(Z \leqslant x)$, where $Z \sim \mathcal{N}(0, 1)$ is a standard normal random variable.

Solution: We want to compute $\mathbb{P}(X(100) > 1800)$. Using the central limit theorem, we can approximate $\frac{X(100) - \mathbb{E}[X(100)]}{\sqrt{\mathrm{Var}(X(100))}}$ as a standard normal random variable $Z \sim \mathcal{N}(0, 1)$. We also know that $\mathbb{E}[X(100)] = 7 \cdot 100 \cdot 2.5 = 1750$ and that $\mathrm{Var}(X(100)) = 7 \cdot 100 \cdot 7 = 4900$. Hence

$$\begin{aligned}\mathbb{P}(X(100) > 1800) &= \mathbb{P}\left(\frac{X(100) - \mathbb{E}[X(100)]}{\sqrt{\mathrm{Var}(X(100))}} > \frac{1800 - \mathbb{E}[X(100)]}{\sqrt{\mathrm{Var}(X(100))}}\right) \\ &\approx \mathbb{P}\left(Z > \frac{1800 - 1750}{70}\right) = \mathbb{P}\left(Z > \frac{5}{7}\right) = 1 - \phi\left(\frac{5}{7}\right) \approx 0.24.\end{aligned}$$

Exercise 2: Renewal theory

(i) Let $\{N(t), t \geqslant 0\}$ be a renewal process with i.i.d. interarrival times X_n, $n \geqslant 1$, and let $\mu = \mathbb{E}[X_n]$. Prove that the rate of the renewal process $\frac{N(t)}{t}$ converges to $\frac{1}{\mu}$ almost surely as $t \to \infty$.

Solution: If $S_{N(t)}$ is the time of the last renewal prior to or at time t, while $S_{N(t)+1}$ is the time of the first renewal after time t, then

$$\frac{S_{N(t)}}{N(t)} \leqslant \frac{t}{N(t)} < \frac{S_{N(t)+1}}{N(t)}.$$

By the strong law of large numbers, since $N(t) \to \infty$, we have that the left-hand side term $\frac{S_{N(t)}}{N(t)}$ converges to μ a.s., and that the right-hand side term $\frac{S_{N(t)+1}}{N(t)}$, which can be written as $\frac{S_{N(t)+1}}{N(t)+1}\frac{N(t)+1}{N(t)} = \frac{S_{N(t)+1}}{N(t)+1}\left(1 + \frac{1}{N(t)}\right)$, converges to μ a.s.. Hence the term in the middle must also converge to μ a.s., and the result is proven by taking the reciprocal.

Suppose that at one of the border control stations between Afghanistan and Iran, all the refugees that arrive from Afghanistan receive help to reach the nearest city in Iran thanks to a bus service provided by some volunteers authorized by the governments. Assume that refugees arrive according to a Poisson process with rate λ. Assume that, as soon as there are N refugees at the station, a bus picks them all up and departs. The bus service association incurs a cost at a rate of nc per unit time whenever there are n refugees waiting at the station.

(ii) Describe the problem in terms of a renewal reward process. State and use the renewal reward theorem to compute the long-run average cost.

Solution: We can model the problem as a renewal reward process where the renewals/cycles are described by the arrivals of buses and the reward is a cost which is paid gradually through a cycle at rate nc per unit time.
The renewal reward theorem says that, if the expected reward in a cycle $\mathbb{E}[R]$ and the expected cycle length $\mathbb{E}[X]$ are finite, then

(i) $\frac{R(t)}{t} \to \frac{\mathbb{E}[R]}{\mathbb{E}[X]}$ almost surely as $t \to \infty$;

(ii) $\frac{\mathbb{E}[R(t)]}{t} \to \frac{\mathbb{E}[R]}{\mathbb{E}[X]}$ as $t \to \infty$.

In our model, the long-run average cost is then given by $\frac{\mathbb{E}[R]}{\mathbb{E}[X]} = \frac{\mathbb{E}[\text{cost in a cycle}]}{\mathbb{E}[\text{length of a cycle}]}$. The expected cost up to the arrival of the first refugee is 0. The expected cost between the first and the second arrival is $\frac{c}{\lambda}$, and in general the expected cost between the k-th and the $k+1$-st arrival is $\frac{kc}{\lambda}$. Hence we have that $\mathbb{E}[R] = \sum_{k=0}^{N-1} \frac{kc}{\lambda} = \frac{cN(N-1)}{2\lambda}$. Moreover, we have that $\mathbb{E}[X] = \frac{N}{\lambda}$. The long-run avarage cost is then

$$\frac{\mathbb{E}[R]}{\mathbb{E}[X]} = \frac{\frac{cN(N-1)}{2\lambda}}{\frac{N}{\lambda}} = \frac{c(N-1)}{2}.$$

Consider now a slightly different model and assume that, as soon as there are N refugees at the station, a bus is called and takes T units of time to arrive. Again, when it arrives all the refugees are picked up.

(iii) What is the expected cost in a cycle?

Solution: The expected cost in a cycle can now be splitted in the sum of three type of costs R_1, R_2 and R_3. The first cost is cost up to the arrival of the N-th refugees (which we calculated above), hence $\mathbb{E}[R_1] = \frac{cN(N-1)}{2\lambda}$. The second cost is the cost coming from the N refugees at the station while waiting for the bus to arrive once it is called. This cost is simply $\mathbb{E}[R_2] = cNT$. The third cost is the cost of the refugees arriving during time T that the bus takes to arrive. Let M be the number of new arrivals during this time. By the order statistic property, they are uniformly distributed in the interval of length T, hence their expected cost is $\mathbb{E}[R_3] = c\,\mathbb{E}[M]\frac{T}{2} = \frac{c\lambda T^2}{2}$. Summing up these three types of cost, we have that

$$\mathbb{E}[R] = \frac{cN(N-1)}{2\lambda} + cNT + \frac{c\lambda T^2}{2}.$$

(iv) Bonus (2 points): What is the long-run average cost?

Solution: The expected length of a cycle is now $\mathbb{E}[X] = \frac{N}{\lambda} + K$. Hence the long-run average cost is

$$\frac{\mathbb{E}[R]}{\mathbb{E}[X]} = \frac{\frac{cN(N-1)}{2\lambda} + cNT + \frac{c\lambda T^2}{2}}{\frac{N}{\lambda} + K} = \frac{cN(N-1) + 2c\lambda TN + c\lambda^2 T^2}{2(N + T\lambda)}.$$

Exercise 3: Queueing theory

When the refugees arrive at the border control station they are sent to a registration desk to fill in some documents. Suppose that there is only one active desk, that the refugees arrive at a Poisson rate λ independently of each other, and that the time it takes for each registration is exponentially distributed with mean $1/\mu$, independently of everything else.

(i) Specify what type of queueing model best describes the registration process. Write down the balance equations and show how they can be solved to compute the limiting probability P_0 that there are no refugees at the border control station. What condition must λ and μ satisfy in order for the limiting probabilities to exist?

Solution: The registration process can be described as an $M/M/1$ queueing model, where the arrival times are i.i.d. $\text{Exp}(\lambda)$ and the service times are i.i.d. $\text{Exp}(\mu)$. The balance equations are

$$\begin{aligned} \lambda P_0 &= \mu P_1, \\ (\lambda + \mu) P_n &= \lambda P_{n-1} + \mu P_{n+1}, \qquad n \geqslant 1. \end{aligned}$$

We have that $P_1 = \frac{\lambda}{\mu} P_0$ and $P_{n+1} = \frac{\lambda}{\mu} P_n + \left(P_n - \frac{\lambda}{\mu} P_{n-1}\right)$ for $n \geqslant 1$. Solving in terms of P_0, for $n \geqslant 2$, $P_n = \frac{\lambda}{\mu} P_{n-1} + \left(P_{n-1} - \frac{\lambda}{\mu} P_{n-2}\right) = \frac{\lambda}{\mu} P_{n-1} = \left(\frac{\lambda}{\mu}\right)^n P_0$. Since $1 = \sum_{n=0}^{\infty} P_n = \sum_{n=0}^{\infty} \left(\frac{\lambda}{\mu}\right)^n P_0 = \frac{P_0}{1-\frac{\lambda}{\mu}}$, we get that $P_0 = 1 - \frac{\lambda}{\mu}$. The condition to be satisfied is $\lambda < \mu$.

(ii) What is the average number of refugees at the station? What is the average number of refugees in the queue at the station?

Solution: The average number of refugees at the station is given by $L = \frac{\lambda}{\mu-\lambda}$. Since $W = \frac{L}{\lambda} = \frac{1}{\mu-\lambda}$ and $W_Q = W - \mathbb{E}[S] = W - \frac{1}{\mu} = \frac{\lambda}{\mu(\mu-\lambda)}$, the average number of refugees in queue at the station is given by $L_Q = \lambda W_Q = \frac{\lambda^2}{\mu(\mu-\lambda)}$.

(iii) Assume now that the refugees arrive on average every 10 minutes according to Poisson process, and that the time S it takes for each registration is not anymore exponentially distributed, but $\mathbb{E}[S] = \mathrm{Var}(S) = 5$ minutes. What is the average number of refugees at the station? What is the average number of refugees in the queue at the station?

Solution: The registration process can now be described as an $M/G/1$ queueing model, where the arrival times are i.i.d. $\mathrm{Exp}\left(\frac{1}{10}\right)$ and the service times are i.i.d. with mean $\mathbb{E}[S] = 5$. By the Pollaczek-Khintchine formula, we have that $W_Q = \frac{\lambda \mathbb{E}[S^2]}{2(1-\lambda\mathbb{E}[S])}$, where $\mathbb{E}[S^2] = \mathrm{Var}(S) + \mathbb{E}[S]^2 = 5 + 25 = 30$. Hence, $W_Q = \frac{\frac{1}{10}30}{2(1-\frac{1}{10}5)} = 3$ minutes. The average number of refugees in queue at the station is then

$$L_Q = \lambda W_Q = \frac{1}{10} 3 = \frac{3}{10}.$$

Since $W = W_Q + \mathbb{E}[S] = 3 + 5 = 8$ minutes, we also have that the average number of refugees at the station is

$$L = \lambda W = \frac{1}{10} 8 = \frac{4}{5}.$$

Exercise 4: Simulation

(i) Describe and prove the inverse transformation method to simulate a random variable with distribution function F.

Solution: The inverse transformation method says that, when F^{-1} is computable, we can simulate a random variable X from a continuous distribution F by simulating $U \sim U(0,1)$ and then setting $X = F^{-1}(U)$.
To prove it, just notice that, since F is monotone, we have

$$\mathbb{P}(X \leqslant a) = \mathbb{P}(F^{-1}(U) \leqslant a) = \mathbb{P}(U \leqslant F(a)) = F(a).$$

(ii) How can we simulate a Poisson random variable with mean λ starting from independent uniform random variables $U_1, U_2, \ldots$? Describe the method and argue that it gives the desired distribution.

Solution: Generate the sequence $U_1, U_2, \ldots$ stopping at $N + 1 = \min\{n : \prod_{i=1}^{n} U_i < e^{-\lambda}\}$. The random variable N has the desired distribution.
Indeed, it can be shown that $N = \max\{n : \sum_{i=1}^{n} - \log(U_i) < \lambda\}$. Then, since $\log(U_i) \sim \text{Exp}(1)$, we can interpret N as the number of events by time λ of a Poisson process with rate 1, which is a Poisson random variable with mean λ.

Suppose that the refugees arrive at the border control station according to a nonhomogeneous Poisson process $\{N(t), t \geqslant 0\}$ with rate $\lambda(t) = 3t^2$.

(iii) Explain in detail how we can simulate the first T time units of the refugees arrival process by first simulating the random variable $N(T)$ and then simulating $N(T)$ random variables representing the arrival times.

Solution: For a nonhomogeneous Poisson process on $[0, T]$, given $N(T)$, the event times are i.i.d. with conditional distribution $F(t) = \frac{m(t)}{m(T)}$, where $m(t) = \int_0^t \lambda(s)\, ds$. In our case, we have that $m(t) = \int_0^t 3s^2\, ds = t^3$, hence the conditional distribution is $F(t) = \frac{t^3}{T^3}$. Since $N(T) \sim \text{Po}(m(T))$, we can simulate the nonhomogeneous Poisson process by first simulating $N(T)$ as described in (ii) and then simulating $N(T)$ random variables from their common distribution function $F(t) = \frac{t^3}{T^3}$. To do so, we can use the inverse transformation method to simulate each of them by simulating a uniform variable $U \sim U(0, 1)$ and then compute $F^{-1}(U) = TU^{1/3}$.

Exercise 5: Brownian motion

(i) Give the definition of a Brownian motion $\{X(t), t \geqslant 0\}$ with drift coefficient μ and variance parameter σ^2.

Solution: The process $\{X(t), t \geqslant 0\}$ is a Brownian motion with drift coefficient μ and variance parameter σ^2 if $X(0) = 0$, $\{X(t), t \geqslant 0\}$ has stationary and independent increments, and $X(t) \sim \mathcal{N}(\mu t, \sigma^2 t)$.

Data from this year show that, due to emissions of greenhouse gases from human activities, the global temperature has increased by 1.1 degrees Celsius compared with pre-industrial levels in the period 1850-1900. Recent studies predicted that it will be 2 degrees Celsius warmer than pre-industrial times by the year 2050.

(ii) Assume that the global temperature evolves according to a Brownian motion with drift $\{X(t), t \geqslant 0\}$ and that it is expected to reach the 2 degrees Celsius threshold exactly in 2050. What is the value of the drift coefficient μ in the unit Celsius/years?

Solution: The process is expected to increase by 0.9 degrees Celcius in 27 years from now, so the drift coefficient must be $\mu = \frac{0.9}{27} = \frac{1}{30}$.

Assume that we will be able to find solution to the climate crisis and completely cancel the effect of the drift in the year 2035, so that the global temperature will evolve according to a Brownian motion with drift μ until 2035 and then without drift until 2050.

(iii) If in 2035 the global temperature will be exactly at its mean value and if $\sigma = 1$, what is probability that it will still reach the 2 degrees Celsius threshold by 2050?

Solution: In 2035, so in 12 years from now, the global temperature is expected to increase by $12\frac{1}{30} = 0.4$ degrees Celsius, so it is expected to be 1.5 degrees Celcius warmer than pre-industrial times. It will then evolve according to a standard Brownian motion without drift. Note that the probability of hitting the threshold of 2 degrees Celsius before 2050 is equivalent to the probability that a standard Brownian motion $\{\bar{X}(t), t \geqslant 0\}$ starting at 0 increases by 0.5 degrees Celsius within 15 years. Recall that, if we let $T_a = \inf\{t \geqslant 0 : \bar{X}(t) \geqslant a\}$ be the hitting time of barrier a, then $\mathbb{P}(T_a \leqslant t) = \frac{2}{\sqrt{2\pi}} \int_{a/\sqrt{t}}^{\infty} e^{-y^2/2}\, dy$. Hence, we have that

$$\mathbb{P}(T_{0.5} \leqslant 15) = \frac{2}{\sqrt{2\pi}} \int_{0.5/\sqrt{15}}^{\infty} e^{-y^2/2}\, dy.$$