Exercise session 10 – Prob. & Stat. 2 — Dec 5, 2023

The Bernoulli process

1. (Sum of a geometric number of independent geometric random variables.)
Let $Y = X_1 + \cdots + X_N$, where the random variables X_i are geometric with parameter p and N is geometric with parameter q. Assume that the random variables $N, X_1, X_2, \ldots$ are independent. Show that Y is geometric with parameter pq. Hint: Interpret the various random variables in terms of a split Bernoulli process.

2. * (The bits in a uniform random variable form a Bernoulli process.)
Let $X_1, X_2, \ldots$ be a sequence of binary random variables taking values in the set $\{0, 1\}$. Let Y be a continuous random variable that takes values in the set $[0, 1]$. We relate X and Y by assuming that Y is the real number whose binary representation is $0.X_1X_2X_3\ldots$

More concretely

$$Y = \sum_{i \geq 1} 2^{-i} X_i.$$

(a) Suppose that the X_i form a Bernoulli process with parameter $p = 1/2$. Show that Y is uniformly distributed. [Hint: Consider the probability of the event $(i-1)/2^k < Y < i/2^k$, where i and k are positive integers.]

(b) Suppose that Y is uniformly distributed. Show that the $X_1, X_2, \ldots$ form a Bernoulli process with parameter $p = 1/2$.

The Poisson process

Recall the definition of the Poisson process by means of exponential waiting times. And also the theorem speaking about distribution of N_t, more precisely $N_{t_{k+1}} - N_{t_k} \sim Pois(\lambda(t_{k+1} - t_k))$. You will also use the following two results about merging and splitting of Poisson processes.

- Given a Poisson process with intensity $\lambda > 0$ and $p \in (0, 1)$. We create a new Poisson process by keeping each arrival with probability p. More precisely: Having defined times T_k as we have in the lecture, at time T_k we toss a coin (with probability p) to decide if something actually happened at time T_k. Then we define new sequences L'_k, T'_k N'_t based on the new arrival times.

  Then the new process is a Poisson process with intensity λp.

- Similarly: having two Poisson processes, with intensity λ_1 and λ_2, their merge is a Poisson process with intensity $\lambda_1 + \lambda_2$.

3. Customers depart from a bookstore according to a Poisson process with rate λ per hour. Each customer buys a book with probability p, independent of everything else.
(a) Find the distribution of the time until the first sale of a book.
(b) Find the probability that no books are sold during a particular hour.
(c) Find the expected number of customers who buy a book during a particular hour.

4. An athletic facility has 5 tennis courts. Pairs of players arrive at the courts and use a court for an exponentially distributed time with mean 40 minutes. Suppose a pair of players arrives and finds all courts busy and k other pairs waiting in queue.
(a) What is the expected time till they get a court?
(b) What is the probability that they get to play within 2 hours – assume $k = 0$.

5. By considering Poisson process, derive that for independent random variables $X_i \sim Pois(\lambda_i)$ we have

$$X_1 + \cdots + X_n \sim Pois(\lambda_1 + \cdots + \lambda_n).$$

6. * Consider a Poisson process. Given that a single arrival occurred in given interval $[0, t]$, show that the distribution of the arrival time is uniform over $[0, t]$.

Experiments

7. Choose one of the following ways to generate a Bernoulli process: generate a sequence of independent Bernoulli trials $X_1, X_2, \ldots$ and compute the waiting times L_t and number of arrivals N_t. OR generate a sequence of independent geometric RVs $L_1, L_2, \ldots$ and from this deduce T_t, L_t and N_t. Verify that the computed variables have the distribution it should have (by computing its variance and mean, or by plotting the distribution of the sampled one and of a separately generated samples from the correct distribution).

8. Generate a Poisson process: generate a sequence of independent exponential RVs $L_1, L_2, \ldots$ and from this deduce $N(t)$. Verify that it has Poisson distribution – as above, by sampling the variable many times, estimating the mean and variance and possibly by plotting the distribution. Or you may use Kolmogorov-Smirnov test (scipy.stats.kstest).