In this lecture, we use many of the new concepts we learned in the last few lectures, and illustrate them in the context of a very important mathematical model called the *Poisson process*. It is commonly used to model the number of phone calls to a hotline, the number of arrivals in an emergency room, the number of earthquakes per year, etc. A key is that for all these phenomena, the events are rare enough to be counted, and have measurable delays between them.

The word process used above refers to the fact that the random phenomenon one is studying involves a whole collection of random variables. We call such a random phenomenon a *random process*.

General setup

Here is how one may intuitively introduce the Poisson process. Imagine you are at the front desk of a hospital emergency service, you set your stop watch to zero, and measure the time until the first patient comes. This can be viewed as a realization of a continuous random variable X_1. You let the stop watch run, and write down the time at which the second patient comes in. This can be viewed as a realization of a second continuous random variable X_2. In this setup, you may therefore consider k random variables X_1, X_2, ..., X_k, corresponding to the arrival times of the first k patients.

In this lecture, we will construct the probability distributions of the X_i. To do so, we will first have to make two mathematical assumptions about this real-life situation, and then derive the probability distributions of a few other mathematical quantities before being able to study the X_i themselves.

1 Number of arrivals

1.1 Applicability of the Poisson process

The Poisson process is a good model when the following two hypotheses are satisfied:

1. **The distinct numbers of arrivals in disjoint time intervals are independent random variables**.

   In other words, the fact that 4 patients arrived in the last 5 minutes does not influence the number of patients who will come in the next five minutes.

2. **The rate λ at which arrivals occur is constant over time (in an averaged sense)**: whatever the time interval of length r, the mean number of arrivals in that time interval is λr. This property is called homogeneity (of the time interval), or weak stationarity.

It is empirically observed that these hypotheses are indeed satisfied to a good approximation in emergency services in many developed countries (except in times of epidemic crises).

1.2 Number of arrivals

We are interested in the total number of arrivals N_t in the time interval $[0, t]$ for some predetermined time t.

N_t is a *discrete* random variable. To characterize its distribution, we fine tune the experiment by subdividing the interval $[0, t]$ in n equal subintervals of length $\frac{t}{n}$, and consider the random variable R_j corresponding to the number of arrivals in the interval of time $[(j-1)\frac{t}{n}, j\frac{t}{n}]$.

By construction, we have

$$N_t = R_1 + R_2 + \ldots + R_n$$

Now, since the arrivals are rare enough to be counted, we can make n large enough that for any j, R_j can take only two possible values, 0 or 1, because the interval $[(j-1)\frac{t}{n}, j\frac{t}{n}]$ is then short enough.

For any j, R_j is therefore a *Bernoulli random variable* with parameter p_j. Let us express p_j in terms of the relevant parameters in the experiment.

For a Bernoulli random variable, we know that we have

$$E[R_j] = p_j$$

and by the homogeneity assumption presented in the previous section, we have

$$E[R_j] = \lambda\frac{t}{n}$$

Therefore, for any j we have

$$p_j = \lambda \frac{t}{n}$$

Note: If you read this derivation very carefully, you may notice what appears to be an inconsistency: if the two hypotheses of independence and of homogeneity apply, then there should be a finite probability that there are at least two arrivals in any interval of time $[(j-1)\frac{t}{n}, j\frac{t}{n}]$ (to see this, try to understand what would happen if we split this interval in two equal halves). It therefore seems like it is not strictly true that the R_j are Bernoulli random variables.

However, the latter statement does indeed become mathematically true in the limit $\frac{t}{n} \to 0$ we will shortly take, which makes our entire derivation mathematically valid.

If you are interested in this subtle issue, you may have a look at the top of p.169 in the textbook for more details.

In summary, N_t is the sum of n independent Bernoulli random variables $R_1, R_2, \ldots, R_n$ with parameter $p = \lambda\frac{t}{n}$. As we saw in lecture 3, this means that N_t has a $\text{Bin}(n, \lambda\frac{t}{n})$ distribution:

$$p_{N_t}(k) = \binom{n}{k} \left(\frac{\lambda t}{n}\right)^k \left(1 - \frac{\lambda t}{n}\right)^{n-k} \quad , \quad k = 0, 1, \ldots, n$$

The accuracy of the characterization of the distribution of N_t improves if we make the length $\frac{t}{n}$ of the subintervals smaller and smaller. We are therefore interested in the limit when $n \to +\infty$:

$$\lim_{n \to +\infty} p_{N_t}(k) = \lim_{n \to +\infty} \left[\binom{n}{k} \left(\frac{\lambda t}{n}\right)^k\right] \lim_{n \to +\infty} \left[\left(1 - \frac{\lambda t}{n}\right)^n\right] \lim_{n \to +\infty} \left[\left(1 - \frac{\lambda t}{n}\right)^{-k}\right]$$

Let us look at each limit separately.

- $$\lim_{n \to +\infty} \binom{n}{k} \left(\frac{\lambda t}{n}\right)^k = \lim_{n \to +\infty} \frac{\overbrace{n(n-1) \cdot (n-k+1)}^{k \text{ terms}}}{\underbrace{n \cdot n \cdot \ldots n}_{k \text{ terms}}} \frac{1}{k!} \lambda^k t^k = \frac{(\lambda t)^k}{k!}$$

- $$\lim_{n \to +\infty} \left(1 - \frac{\lambda t}{n}\right)^n = \lim_{n \to +\infty} e^{n \ln(1 - \frac{\lambda t}{n})} = \lim_{n \to +\infty} e^{-\frac{n\lambda t}{n}} = e^{-\lambda t}$$

- $$\lim_{n \to +\infty} \left(1 - \frac{\lambda t}{n}\right)^{-k} = 1$$

Hence, in the limit when $n \to +\infty$,

$$\boxed{p_{N_t}(k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}} \tag{1}$$

Comparison with page 6 of Lecture 3 then tells us that N_t is a *Poisson* random variable with parameter λt. We may write $N_t \sim \text{Poisson}(\lambda t)$.

It is important to observe that we could reproduce all the steps of our derivation for the case when the time interval does not start at 0, but instead at some $s > 0$, so that the interval would be $[s, s+t]$. In that case too the number of arrivals N_t in the span $[s, s+t]$ would be such that $N_t \sim \text{Poisson}(\lambda t)$. This is a natural consequence of the hypotheses of independence and of homogeneity. We will use this result shortly.

2 The Poisson process

Now that we characterized N_t, let us see what we can say about the random variables $X_1, X_2, \ldots, X_k$ corresponding to the time of arrival of the different patients.

2.1 Interarrival times

By definition, X_1 is the time T_1 of the first arrival. For $i \geq 2$, let

$$T_i = X_i - X_{i-1}$$

which is the time between successive arrivals.

Let us see if we can characterize the distributions of the T_i, as we did for N_t.

For any $t \in \mathbb{R}$

$$F_{T_1}(t) = P(T_1 \leq t) = 1 - P(T_1 > t) = 1 - P(N_t = 0) = 1 - e^{-\lambda t}$$

From Probem Set 5, we know that this means that T_1 has an *exponential* distribution with parameter λ: $T_1 \sim \mathrm{Exp}(\lambda)$.

Let us continue. For any $s > 0$, any $t > 0$,

$$P(T_2 \leq t) = 1 - P(T_2 > t) = 1 - P(N_{t+s} = 1 | N_s = 1) = 1 - \frac{P(N_{t+s} = 1, N_s = 1)}{P(N_s = 1)}$$

To make further progress, we may write the event $\{N_{t+s} = 1, N_s = 1\}$ as the event $\{N_{t+s} - N_s = 0, N_s = 1\}$, i.e. $\{N_{[s,t+s]} = 0, N_s = 1\}$, where $N_{[s,t+s]}$ is the number of arrivals in the time interval $[s, t+s]$.

Now, since the number of arrivals in disjoint time intervals are independent random variables,

$$P(N_{[s,t+s]} = 0, N_s = 1) = P(N_{[s,t+s]} = 0) P(N_s = 1) = e^{-\lambda t} P(N_s = 1)$$

We conclude that

$$P(T_2 \leq t) = 1 - e^{-\lambda t} \frac{P(N_s = 1)}{P(N_s = 1)} = 1 - e^{-\lambda t}$$

In other words, T_2 also has an *exponential* distribution, with parameter λ: $T_2 \sim \mathrm{Exp}(\lambda)$.

Similar reasoning for all the T_i would always lead to the same conclusion: $T_i \sim \mathrm{Exp}(\lambda)$.

This property is our *definition* of what we call the Poisson process.

2.2 Definition of the one-dimensional Poisson process

The one-dimensional **Poisson process** with intensity λ is a **sequence** X_1, X_2, $X_3, \ldots$ of random variables with the property that **the interarrival times** X_1, $X_2 - X_1$, $X_3 - X_2, \ldots$ **are independent random variables, each with an Exp(λ) distribution**.

Note 1: The definition just told us that for a Poisson process, $T_1 = X_1$, $T_2 = X_2 - X_1$, $T_3 = X_3 - X_2$, ... all have an exponential distribution. You proved in Problem Set 5 that they are therefore all *memoryless random variables*.

For example, that means that for any $s > 0$, and any $t > 0$,

$$P(T_2 > s + t | T_2 > t) = P(T_2 > s)$$

Note 2: One may wonder how the arrival times $X_1, X_2, X_3, \ldots$ themselves are distributed. To do so, we observe that

$$\begin{aligned} X_1 &= T_1 \\ X_2 &= T_1 + T_2 \\ X_3 &= T_1 + T_2 + T_3 \\ \vdots \quad &\quad \vdots \end{aligned}$$

The X_i can thus be written as the sum of independent random variables with exponential distribution. You will show in recitation that we then know that the random variable X_i has a *gamma distribution* with parameters i and λ, written $\mathrm{Gam}(i, \lambda)$, which corresponds to the following probability density function:

$$f_{X_i}(x) = \lambda \frac{(\lambda x)^{i-1} e^{-\lambda x}}{(i-1)!} \qquad , \quad x \geq 0$$

2.3 Location of points

Let us now flip the viewpoint somewhat: say we are interested in a particular interval of time $[0, t]$, and know that there have been p arrivals in that time interval. The question is: where are these p arrivals located in the time interval?

Let us first answer the question for the case when we know that there is only one arrival: $N_t = 1$. We are interested in the following conditional probability:

$$P(X_1 \leq x | N_t = 1) = \frac{P(X_1 \leq x, N_t = 1)}{P(N_t = 1)} = \frac{P(N_x = 1, N_{[x,t]} = 0)}{P(N_t = 1)}$$

where, in agreement with previous notation, $N_{[x,t]}$ is the number of arrivals in the interval $[x, t]$.

Now, by hypothesis, N_x and $N_{[x,t]}$ are independent random variables, so

$$P(N_x = 1, N_{[x,t]} = 0) = P(N_x = 1)P(N_{[x,t]} = 0) = \lambda x e^{-\lambda x} e^{-\lambda(t-x)} = \lambda x e^{-\lambda t}$$

Hence,

$$P(X_1 \leq x | N_t = 1) = \frac{\lambda x e^{-\lambda t}}{\lambda t e^{-\lambda t}} = \frac{x}{t}$$

We find that this is the expression for the cumulative distribution of a uniform random variable on the interval $[0, t]$. In other words, given the event $\{N_t = 1\}$, the random variable X_1 is **uniformly distributed** on $[0, t]$.

In recitation, you will prove an analog result for the case $\{N_t = 2\}$: X_1 and X_2 are (to within a reordering) **uniformly distributed** over the interval $[0, t]$, and independent (conditional to the event $\{N_t = 2\}$).

The result can be fully generalized to the following theorem:

Given that the Poisson process has p points in the interval $[a, b]$, the locations of these points in $[a, b]$ are **independently distributed, each with a uniform distribution** on $[a, b]$.

3 Worked example

Queuing at a post office can be accurately modeled by a Poisson process with arrival rate $\lambda = 30$ customers per hour. Find

1. The expected number of arrivals in the first 10 minutes of an hour.
2. The probability of 4 or fewer arrivals in the first 10 minutes of an hour.
3. The probability that there are 3 customers in the first 10 minutes of an hour, and 5 customers in the next 20 minutes.
4. The expected time between the arrival of the 6^{th} customer and the 7^{th} customer.

<u>Solutions</u>

Foreword: For this problem, and many others, it will be useful to know the expected value of a Poisson random variable. Let us derive that formula here.

Let X be a random variable with a Poisson distribution:

$$p_X(k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

Then,

$$
\begin{aligned}
E[X] &= \sum_{k=0}^{+\infty} k \cdot \frac{\lambda^k}{k!} e^{-\lambda} \\
&= \sum_{k=1}^{+\infty} \frac{\lambda^k}{(k-1)!} e^{-\lambda} \\
&= \sum_{k=1}^{+\infty} \frac{\lambda^k}{(k-1)!} e^{-\lambda} \\
&= \lambda \sum_{k=1}^{+\infty} \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda} \\
&\overset{p=k-1}{=} \lambda \underbrace{\sum_{p=0}^{+\infty} \frac{\lambda^p}{p!} e^{-\lambda}}_{=1\text{ , see Lecture 3, p.7}} \\
&= \lambda
\end{aligned}
$$

We are now ready to solve the problem.

1. We have $E[N_{10}] = \lambda \cdot 10 = \frac{1}{2} \cdot 10 = 5$ arrivals.
2. The probability is

$$
\begin{aligned}
P(N_{10}=0) + P(N_{10}=1) + P(N_{10}=2) + P(N_{10}=3) + P(N_{10}=4) &= \frac{5^0}{0!}e^{-5} + \frac{5}{1!}e^{-5} + \frac{5^2}{2!}e^{-5} + \frac{5^3}{3!}e^{-5} + \frac{5^4}{4!}e^{-5} \\
&= e^{-5}\left(1 + 5 + \frac{25}{2} + \frac{125}{6} + \frac{625}{24}\right) \\
&\approx 0.4405
\end{aligned}
$$

3. We are looking for

$$
\begin{aligned}
P(N_{10}=3, N_{[10,30]}=5) &\overset{\text{disjoint intervals}}{=} P(N_{10}=3)P(N_{[10,30]}=5) \\
&\overset{\text{homogeneity property}}{=} P(N_{10}=3)P(N_{20}=5) \\
&= \left(e^{-5}\frac{125}{6}\right)\left(e^{-10}\frac{10^5}{5!}\right) \\
&= e^{-15}\frac{12500000}{6!} \approx 0.0053
\end{aligned}
$$

4. We are looking for $E[T_7]$. We know that T_7 has an exponential distribution with parameter $\lambda = \frac{1}{2}$. From Problem Set 5, we then know that

$$
E[T_7] = \frac{1}{\lambda} = 2 \text{ minutes}
$$