The Poisson process
Math 217 Probability and Statistics
Prof. D. Joyce, Fall 2014

Poisson processes. In a Poisson process, things happen occur uniformly randomly over time. We've used the word "event" to mean one thing so far, namely a subset of a sample space, but we'll also use it for each occurrence that occurs in a Poisson process.

An example of random events over time is radioactive decay. Suppose we have a mass of some radioactive substance. A Geiger counter can listen to decays of this substance, and it sounds like random clicks. Each event is the decay of a particular atom that can be detected by a Geiger counter.

We'll use one parameter for Poisson process, λ, the rate of events per unit time. As usual, we'll use the variable t for time.

The events in a Poisson process are not regularly spaced over time but random in some way. Unlike the Bernoulli process in which time is discrete, the Poisson process uses continuous time. There will be a few probability distributions related to the Poisson process analogous to those associated to the Bernoulli distribution.

- The Poisson distribution. It gives the number of events in unit time. Although the rate of events per unit time is λ, in any unit time interval there be some particular number 0, 1, 2, etc. The probability that there are n events in unit time will depend on the parameter λ. This is a discrete distribution. This is analogous to the binomial distribution for a Bernoulli process.

- The exponential distribution. This specifies the time until the next event. It's a continuous distribution analogous to the geometric distribution for a Bernoulli process.

- The gamma distribution. It specifies the time until the rth event. It's also a continuous distribution, and it's analogous to the negative binomial distribution for a Bernoulli process.

- The beta distribution. In a Poisson process, if $\alpha + \beta$ events occur in a time interval, then the fraction of that interval until the αth event occurs has a continuous distribution called the beta distribution. It's not quite analogous to the hypergeometric distribution for a Bernoulli process; the assumption is analogous, but the question is different.

We'll analyze the Poisson and exponential distributions today.

Axioms for the Poisson process. Poisson analyzed this process by means of axioms. We'll do the same thing. We'll make assumptions that we reflect some simple properties that we expect the process to have, then derive the above distributions from those assumptions. There are three axioms.

1. The numbers of events in two nonoverlapping regions are independent.

2. The probability of an event occurring in an interval $[a, a+h]$ of short length h is approximately proportional to h, with proportionality ratio λ, called the rate of events. More precisely,
$$\lim_{h \to 0} \frac{P(\text{an event in } [a, a+h])}{h} = \lambda.$$

3. The probability of two events occurring in a short interval $[t, t+h]$ is much smaller than the probability of one event. More precisely,
$$\lim_{h \to 0} \frac{P(\text{two or more events in } [a, a+h])}{P(\text{an event in } [a, a+h])} = 0.$$

Development from the axioms. We'll use these axioms to develop various related probability distributions. Let's use the notation $P_0(t)$ to stand for the probability that no events take place in an interval of length t. We'll figure out the derivative P_0' in terms of P_0.

First, note that
$$P_0(t + h) = P_0(t)\, P_0(h).$$

That's because if there are no events in the interval $[0, t+h]$, then there are none in the two intervals $[0, t]$ and $[t, t+h]$, and since those are nonoverlapping regions, by axiom 1 the probabilities are independent, and so multiply. Therefore,

$$
\begin{aligned}
P_0'(t) &= \lim_{h \to 0} \frac{P_0(t+h) - P_0(t)}{h} \\
&= \lim_{h \to 0} \frac{P_0(t)\, P_0(h) - P_0(t)}{h} \\
&= P_0(t) \lim_{h \to 0} \frac{P_0(h) - 1}{h}
\end{aligned}
$$

Let's pause for a moment. The probability $P_0(h)$ that no events occur in $[t, t+h]$ equals $1 - P(\text{at least one event in } [t, t+h])$. Therefore

$$P_0'(t) = P_0(t) \lim_{h \to 0} \frac{-P(\text{at least one event in } [t, t+h])}{h}$$

By axiom 2, that last limit is equal to λ. Now we have the equation

$$P_0'(t) = -\lambda P_0(t)$$

which is a differential equation. In fact, it's the exponential differential equation, perhaps the most important differential equation there is. You've probably seen it in the form $y' = ky$ and you know its general solution is $y = Ae^{kt}$. Translating that into the current notation, $P_0'(t) = -\lambda P_0(t)$ has the general solution $P_0(t) = Ae^{-\lambda t}$. Since we know that $P_0(0) = 1$, therefore the constant A is equal to 1. We've derived from the axioms this result:

$$P_0(t) = e^{-\lambda t}.$$

Next, let $P_1(t)$ be the probability that exactly one event takes place in an interval of length t. Then

$$
\begin{aligned}
P_1'(t) &= \lim_{h \to 0} \frac{P_1(t+h) - P_1(t)}{h} \\
&= \lim_{h \to 0} \frac{P_1(t)\,P_0(h) + P_0(t)\,P_1(h) - P_1(t)}{h} \\
&= \lim_{h \to 0} \frac{P_1(t)\,P_0(h) - P_1(t)}{h} + \lim_{h \to 0} \frac{P_0(t)\,P_1(h)}{h} \\
&= P_1(t) \lim_{h \to 0} \frac{P_0(h) - 1}{h} + P_0(t) \lim_{h \to 0} \frac{P_1(h)}{h} \\
&= -\lambda P_1(t) + \lambda P_0(t).
\end{aligned}
$$

Note that axiom 3 was used in the last equation. We now have the differential equation

$$
P_1'(t) = -\lambda P_1(t) + \lambda e^{-\lambda t}.
$$

That's not such an elementary equation as the first one, but it's what's called a linear differential equation, and it can be solved by elementary methods. Along with the initial value $P_1(0) = 0$, there's a unique solution which is

$$
P_1(t) = \lambda t e^{-t}.
$$

In general, if we let $P_n(t)$ be the probability that exactly n events take place in an interval of length t, we can derive a differential equation

$$
P_n'(t) = -\lambda P_n(t) + \lambda P_{n-1}(t)
$$

which has the solution

$$
P_n(t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}.
$$

This is the Poisson distribution. If X is the number of events in an interval of length t, then

$$
P(X{=}n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}.
$$

The exponential distribution.   We've already studied the exponential distribution, but now we can justify the claims we made for it. Let T be the time of the first event after $t = 0$. Then T is what we've called the exponential distribution. It's very closely related to the function $P_0(t)$ we found above. Let $F_T(t)$ denote the cumulative distribution function for the exponential distribution. Then

$$
\begin{aligned}
F_T(t) &= P(T \le t) \\
&= 1 - P(T \ge t) \\
&= 1 - P_0(t) = 1 - e^{-\lambda t}
\end{aligned}
$$

That gives us the justification for the formula we've been using for the exponential distribution. Differentiating it gives us the density function

$$
f_T(t) = \lambda e^{-\lambda t}.
$$

The exponential density function is

$$f(t) = \begin{cases} \lambda e^{-\lambda t} & \text{for } t > 0 \\ 0 & \text{otherwise.} \end{cases}$$

We can find the corresponding cumulative distribution function by integration.

$$(t) = \int_{-\infty}^{t} f(x)\, dx = \int_{0}^{t} \lambda e^{-\lambda x}\, dx = -e^{-\lambda x}\Big|_{0}^{t} = -e^{-\lambda t} + e^{-\lambda 0} = 1 - e^{-\lambda t}$$

Math 217 Home Page at `http://math.clarku.edu/~djoyce/ma217/`