12

The Poisson process

In many random phenomena we encounter, it is not just one or two random variables that play a role but a whole collection. In that case one often speaks of a random *process*. The Poisson process is a simple kind of random process, which models the occurrence of random points in time or space. There are numerous ways in which processes of random points arise: some examples are presented in the first section. The Poisson process describes in a certain sense the *most random way* to distribute points in time or space. This is made more precise with the notions of homogeneity and independence.

12.1 Random points

Typical examples of the occurrence of random time points are: arrival times of email messages at a server, the times at which asteroids hit the earth, arrival times of radioactive particles at a Geiger counter, times at which your computer crashes, the times at which electronic components fail, and arrival times of people at a pump in an oasis.

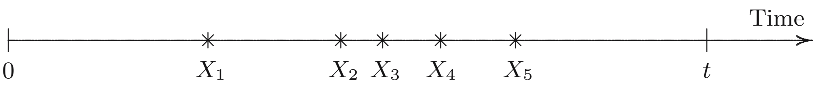
Examples of the occurrence of random points in space are: the locations of asteroid impacts with earth (2-dimensional), the locations of imperfections in a material (3-dimensional), and the locations of trees in a forest (2-dimensional).

Some of these phenomena are better modeled by the Poisson process than others. Loosely speaking, one might say that the Poisson process model often applies in situations where there is a very large population, and each member of the population has a very small probability to produce a point of the process. This is, for instance, well fulfilled in the Geiger counter example where, in a huge collection of atoms, just a few will emit a radioactive particle (see [28]). A property of the Poisson process—as we will see shortly—is that points may lie arbitrarily close together. Therefore the tree locations are not so well modeled by the Poisson process.

12.2 Taking a closer look at random arrivals

A well-known example that is usually modeled by the Poisson process is that of calls arriving at a telephone exchange—the exchange is connected to a large number of people who make phone calls now and then. This will be our leading example in this section.

Telephone calls arrive at random times $X_1, X_2, \ldots$ at the telephone exchange during a time interval $[0, t]$.



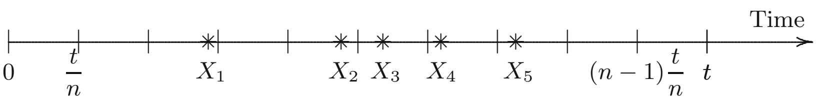
The two basic assumptions we make on these random arrivals are

1. (*Homogeneity*) The rate λ at which arrivals occur is constant over time: in a subinterval of length u the expectation of the number of telephone calls is λu.
2. (*Independence*) The numbers of arrivals in disjoint time intervals are independent random variables.

Homogeneity is also called weak stationarity. We denote the total number of calls in an interval I by $N(I)$, abbreviating $N([0, t])$ to N_t. Homogeneity then implies that we require

$$\mathrm{E}[N_t] = \lambda t.$$

To get hold of the *distribution* of N_t we divide the interval $[0, t]$ into n intervals of length t/n. When n is large enough, every interval $I_{j,n} = ((j-1)\,t/n, j\,t/n]$ will contain either 0 or 1 arrival: For such a large n (which also satisfies



$n > \lambda t$), let R_j be the number of arrivals in the time interval $I_{j,n}$. Since R_j is 0 or 1, R_j has a $Ber(p_j)$ distribution for some p_j. Recall that for a Bernoulli random variable $\mathrm{E}[R_j] = 0 \cdot (1 - p_j) + 1 \cdot p_j = p_j$. By the homogeneity assumption, for each j

$$p_j = \lambda \cdot \text{length of } I_{j,n} = \frac{\lambda t}{n}.$$

Summing the number of calls in the intervals gives the total number of calls, hence

$$N_t = R_1 + R_2 + \cdots + R_n.$$

By the independence assumption, the R_j are independent random variables, therefore N_t has a $Bin(n,p)$ distribution, with $p = \lambda t/n$.

Remark 12.1 (About this approximation). The argument just given seems pretty convincing, but actually R_j does *not* have a Bernoulli distribution, whatever the value of n. A way to see this is the following. Every interval $I_{j,n}$ is a union of the two intervals $I_{2j-1,2n}$ and $I_{2j,2n}$. Hence the probability that $I_{j,n}$ contains two calls is at least $(\lambda t/2n)^2 = \lambda^2 t^2/4n^2$, which is larger than zero.
Note however, that the probability of having two arrivals is of smaller order than the probability that R_j takes the value 1. If we add a third assumption, namely that the probability of two or more calls arriving in an interval $I_{j,n}$ tends to zero faster than $1/n$, then the conclusion below on the distribution of N_t is valid.

We have found that (at least in first approximation)

$$\mathrm{P}(N_t = k) = \binom{n}{k}\left(\frac{\lambda t}{n}\right)^k \left(1 - \frac{\lambda t}{n}\right)^{n-k} \qquad \text{for } k = 0, \ldots, n.$$

In this analysis n is a rather artificial parameter, of which we only know that it should not be "too small." It therefore seems a good idea to get rid of n by letting n go to infinity, hoping that the probability distribution of N_t will settle down. Note that

$$\lim_{n\to\infty} \binom{n}{k}\frac{1}{n^k} = \lim_{n\to\infty} \frac{n}{n} \cdot \frac{n-1}{n} \cdots \frac{(n-k+1)}{n} \cdot \frac{1}{k!} = \frac{1}{k!},$$

and from calculus we know that

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

Since certainly

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

we obtain, combining these three limits, that

$$\lim_{n\to\infty} \mathrm{P}(N_t = k) = \lim_{n\to\infty} \binom{n}{k}\frac{1}{n^k} \cdot (\lambda t)^k \cdot \left(1 - \frac{\lambda t}{n}\right)^n \cdot \left(1 - \frac{\lambda t}{n}\right)^{-k} = \frac{(\lambda t)^k}{k!}\mathrm{e}^{-\lambda t}.$$

Since

$$\mathrm{e}^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} = \mathrm{e}^{-\lambda t}\mathrm{e}^{\lambda t} = 1,$$

we have indeed run into a probability distribution on the numbers $0, 1, 2, \ldots$. Note that all these probabilities are determined by the single value λt. This motivates the following definition.

DEFINITION. A discrete random variable X has a *Poisson distribution* with parameter μ, where $\mu > 0$ if its probability mass function p is given by

$$p(k) = \mathrm{P}(X = k) = \frac{\mu^k}{k!}\,\mathrm{e}^{-\mu} \quad \text{for } k = 0, 1, 2, \ldots.$$

We denote this distribution by $Pois(\mu)$.

Figure 12.1 displays the graphs of the probability mass functions of the Poisson distribution with $\mu = 0.9$ (left) and the Poisson distribution with $\mu = 5$ (right).

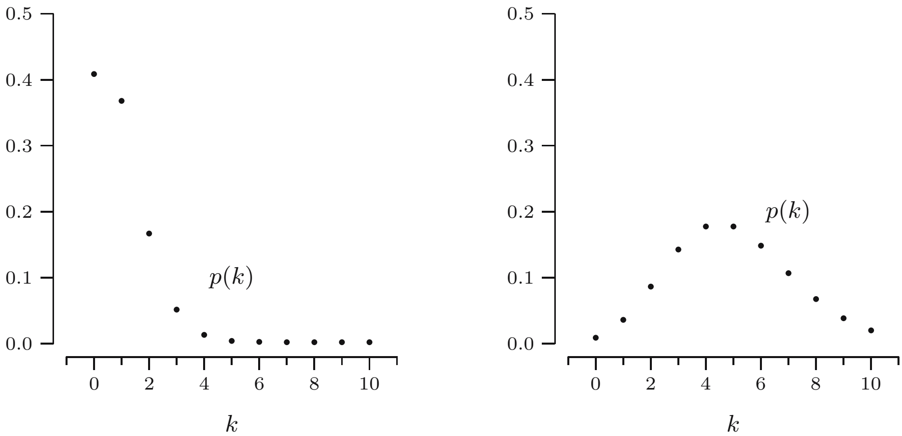


Fig. 12.1. The probability mass functions of the $Pois(0.9)$ and the $Pois(5)$ distributions.

QUICK EXERCISE 12.1 Consider the event "exactly one call arrives in the interval $[0, 2s]$." The probability of this event is $\mathrm{P}(N_{2s} = 1) = \lambda \cdot 2s \cdot \mathrm{e}^{-\lambda \cdot 2s}$. But note that this event is the same as "there is exactly one call in the interval $[0, s)$ and no calls in the interval $[s, 2s]$, *or* no calls in $[0, s)$ and exactly one call in $[s, 2s]$." Verify (using assumptions 1 and 2) that you get the same answer if you compute the probability of the event in this way.

We do have a hint[1] about what the expectation and variance of a Poisson random variable might be: since $\mathrm{E}[N_t] = \lambda t$ for all n, we anticipate that the limiting Poisson distribution will have expectation λt. Similarly, since N_t has a $Bin(n, \frac{\lambda t}{n})$ distribution, we anticipate that the variance will be

[1] This is really not more than a hint: there are simple examples where the distributions of random variables converge to a distribution whose expectation is different from the limit of the expectations of the distributions! (cf. Exercise 12.14).

$$\lim_{n\to\infty} \mathrm{Var}(N_t) = \lim_{n\to\infty} n \cdot \frac{\lambda t}{n} \cdot \left(1 - \frac{\lambda t}{n}\right) = \lambda t.$$

Actually, the expectation of a Poisson random variable X with parameter μ is easy to compute:

$$\begin{aligned}\mathrm{E}[X] &= \sum_{k=0}^{\infty} k \frac{\mu^k}{k!} \mathrm{e}^{-\mu} = \mathrm{e}^{-\mu} \sum_{k=1}^{\infty} \frac{\mu^k}{(k-1)!} \\ &= \mu \mathrm{e}^{-\mu} \sum_{k=1}^{\infty} \frac{\mu^{k-1}}{(k-1)!} = \mu \mathrm{e}^{-\mu} \sum_{j=0}^{\infty} \frac{\mu^j}{j!} = \mu.\end{aligned}$$

In a similar way the variance can be determined (see Exercise 12.8), and we arrive at the following rule.

THE EXPECTATION AND VARIANCE OF A POISSON DISTRIBUTION. Let X have a Poisson distribution with parameter μ; then

$$\mathrm{E}[X] = \mu \quad \text{and} \quad \mathrm{Var}(X) = \mu.$$

12.3 The one-dimensional Poisson process

We will derive some properties of the sequence of random points $X_1, X_2, \ldots$ that we considered in the previous section. What we derived so far is that for any interval $(s, s+t]$ the number $N((s, s+t])$ of points X_i in that interval is a random variable with a $Pois(\lambda t)$ distribution.

Interarrival times

The differences

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

are called interarrival times. Here we define $T_1 = X_1$, the time of the *first* arrival. To determine the probability distribution of T_1, we observe that the event $\{T_1 > t\}$ that the first call arrives after time t is the same as the event $\{N_t = 0\}$ that no calls have been made in $[0, t]$. But this implies that

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

Therefore T_1 has an exponential distribution with parameter λ.

To compute the joint distribution of T_1 and T_2, we consider the conditional probability that $T_2 > t$, given that $T_1 = s$, and use the property that arrivals in different intervals are independent:

$$\begin{aligned}\mathrm{P}(T_2 > t \,|\, T_1 = s) &= \mathrm{P}(\text{no arrivals in } (s, s+t] \,|\, T_1 = s)\\ &= \mathrm{P}(\text{no arrivals in } (s, s+t])\\ &= \mathrm{P}(N((s, s+t]) = 0) = \mathrm{e}^{-\lambda t}.\end{aligned}$$

Since this answer does not depend on s, we conclude that T_1 and T_2 are independent, and

$$\mathrm{P}(T_2 > t) = \mathrm{e}^{-\lambda t},$$

i.e., T_2 also has an exponential distribution with parameter λ. Actually, although the conclusion is correct, the method to derive it is not, because we conditioned on the event $\{T_1 = s\}$, which has zero probability. This problem could be circumvented by conditioning on the event that T_1 lies in some small interval, but that will not be done here. Analogously, one can show that the T_i are independent and have an $Exp(\lambda)$ distribution. This nice property allows us to give a simple definition of the one-dimensional Poisson process.

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

Note that the connection with N_t is as follows: N_t is equal to the number of X_i that are smaller than (or equal to) t.

QUICK EXERCISE 12.2 We model the arrivals of email messages at a server as a Poisson process. Suppose that on average 330 messages arrive per minute. What would you choose for the intensity λ in messages per second? What is the expectation of the interarrival time?

An obvious question is: what is the distribution of X_i? This has already been answered in Chapter 11: since X_i is a sum of i independent exponentially distributed random variables, we have the following.

> THE POINTS OF THE POISSON PROCESS. For $i = 1, 2, \ldots$ the random variable X_i has a $Gam(i, \lambda)$ distribution.

The distribution of points

Another interesting question is: if we know that n points are generated in an interval, where do these points lie? Since the distribution of the number of points only depends on the length of the interval, and not on its location, it suffices to determine this for an interval starting at 0. Let this interval be $[0, a]$. We start with the simplest case, where there is one point in $[0, a]$: suppose that $N([0, a]) = 1$. Then, for $0 < s < a$:

$$\begin{aligned}
\mathrm{P}(X_1 \le s \,|\, N([0,a]) = 1) &= \frac{\mathrm{P}(X_1 \le s, N([0,a]) = 1)}{\mathrm{P}(N([0,a]) = 1)} \\
&= \frac{\mathrm{P}(N([0,s]) = 1, N((s,a]) = 0)}{\mathrm{P}(N([0,a]) = 1)} \\
&= \frac{\lambda s \mathrm{e}^{-\lambda s} \mathrm{e}^{-\lambda(a-s)}}{\lambda a \mathrm{e}^{-\lambda a}} \\
&= \frac{s}{a}.
\end{aligned}$$

We find that conditional on the event $\{N([0,a]) = 1\}$, the random variable X_1 is uniformly distributed over the interval $[0,a]$.

Now suppose that it is given that there are two points in $[0,a]$: $N([0,a]) = 2$. In a way similar to what we did for *one* point, we can show that (see Exercise 12.12)

$$\mathrm{P}(X_1 \le s, X_2 \le t \mid N([0,a]) = 2) = \frac{t^2 - (t-s)^2}{a^2}.$$

Now recall the result of Exercise 9.17: if U_1 and U_2 are two independent random variables, both uniformly distributed over $[0,a]$, then the joint distribution function of $V = \min(U_1, U_2)$ and $Z = \max(U_1, U_2)$ is given by

$$\mathrm{P}(V \le s, Z \le t) = \frac{t^2 - (t-s)^2}{a^2} \quad \text{for } 0 \le s \le t \le a.$$

Thus we have found that, if we forget about their order, the two points in $[0,a]$ are independent and uniformly distributed over $[0,a]$. With somewhat more work, this generalizes to an arbitrary number of points, and we arrive at the following property.

LOCATION OF THE POINTS, GIVEN THEIR NUMBER. Given that the Poisson process has n points in the interval $[a,b]$, the locations of these points are independently distributed, each with a uniform distribution on $[a,b]$.

12.4 Higher-dimensional Poisson processes

Our definition of the one-dimensional Poisson process, starting with the interarrival times, does not generalize easily, because it is based on the ordering of the real numbers. However, we can easily extend the assumptions of independence, homogeneity, and the Poisson distribution property. To do this we need a higher-dimensional version of the concept of length. We denote the k-dimensional volume of a set A in k-dimensional space by $m(A)$. For instance, in the plane $m(A)$ is the area of A, and in space $m(A)$ is the volume of A.

Definition. The k-dimensional *Poisson process* with intensity λ is a collection $X_1, X_2, X_3, \ldots$ of random points having the property that if $N(A)$ denotes the number of points in the set A, then

1. (*Homogeneity*) The random variable $N(A)$ has a Poisson distribution with parameter $\lambda m(A)$.
2. (*Independence*) For disjoint sets $A_1, A_2, \ldots, A_n$ the random variables $N(A_1)$, $N(A_2), \ldots, N(A_n)$ are independent.

Quick exercise 12.3 Suppose that the locations of defects in a certain type of material follow the two-dimensional Poisson process model. For this material it is known that it contains on average five defects per square meter. What is the probability that a strip of length 2 meters and width 5 cm will be without defects?

In Figure 7.4 the locations of the buildings the architect wanted to distribute over a 100-by-300-m terrain have been generated by a two-dimensional Poisson process. This has been done in the following way. One can again show that given the total number of points in a set, these points are uniformly distributed over the set. This leads to the following procedure: first one generates a value n from a Poisson distribution with the appropriate parameter (λ times the area), then one generates n times a point uniformly distributed over the 100-by-300 rectangle.

Actually one can generate a higher-dimensional Poisson process in a way that is very similar to the natural way this can be done for the one-dimensional process. Directly from the definition of the one-dimensional process we see that it can be obtained by consecutively generating points with exponentially distributed gaps. We will explain a similar procedure for dimension two. For $s > 0$, let

$$M_s = N(C_s),$$

where C_s is the circular region of radius s, centered at the origin. Since C_s has area πs^2, M_s has a Poisson distribution with parameter $\lambda \pi s^2$. Let R_i denote the distance of the ith closest point to the origin. This is illustrated in Figure 12.2.

Note that R_i is the analogue of the ith arrival time for the one-dimensional Poisson process: we have in fact that

$$R_i \le s \qquad \text{if and only if} \qquad M_s \ge i.$$

In particular, with $i = 1$ and $s = \sqrt{t}$,

$$\mathrm{P}\big(R_1^2 \le t\big) = \mathrm{P}\Big(R_1 \le \sqrt{t}\Big) = \mathrm{P}\big(M_{\sqrt{t}} > 0\big) = 1 - \mathrm{e}^{-\lambda \pi t}.$$

In other words: R_1^2 is $Exp(\lambda\pi)$ distributed. For general i, we can similarly write

$$\mathrm{P}\big(R_i^2 \le t\big) = \mathrm{P}\Big(R_i \le \sqrt{t}\Big) = \mathrm{P}\big(M_{\sqrt{t}} \ge i\big).$$

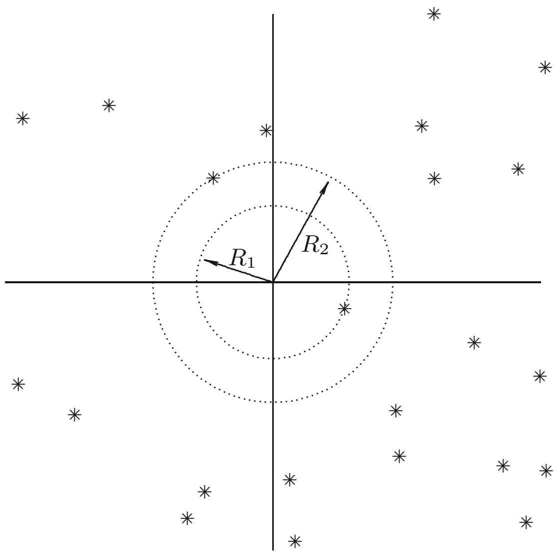


Fig. 12.2. The Poisson process in the plane, with the two circles of the two points closest to the origin.

So

$$\mathrm{P}\big(R_i^2 \leq t\big) = 1 - \mathrm{e}^{-\lambda\pi t} \sum_{j=0}^{i-1} \frac{(\lambda\pi t)^j}{j!},$$

which means that R_i^2 has a $Gam(i, \lambda\pi)$ distribution—as we saw on page 157. Since gamma distributions arise as sums of independent exponential distributions, we can also write

$$R_i^2 = R_{i-1}^2 + T_i,$$

where the T_i are independent $Exp(\lambda\pi)$ random variables (and where $R_0 = 0$). Note that this is quite similar to the one-dimensional case. To simulate the two-dimensional Poisson process from a sequence $U_1, U_2, \ldots$ of independent $U(0,1)$ random variables, one can therefore proceed as follows (recall from Section 6.2 that $-(1/\lambda)\ln(U_i)$ has an $Exp(\lambda)$ distribution): for $i = 1, 2, \ldots$ put

$$R_i = \sqrt{R_{i-1}^2 - \frac{1}{\lambda\pi}\ln(U_{2i})};$$

this gives the distance of the ith point to the origin, and then put the point on this circle according to an angle value generated by $2\pi U_{2i-1}$. This is the correct way to do it, because one can show that in polar coordinates the radius and the angle of a Poisson process point are independent of each other, and the angle is uniformly distributed over $[0, 2\pi]$. The latter is called the *isotropy* property of the Poisson process.

12.5 Solutions to the quick exercises

12.1 The probability of exactly one call in $[0, s)$ and no calls in $[s, 2s]$ equals

$$\begin{aligned}\mathrm{P}(N([0,s))=1, N([s,2s])=0) &= \mathrm{P}(N([0,s))=1)\,\mathrm{P}(N([s,2s])=0)\\ &= \mathrm{P}(N([0,s))=1)\,\mathrm{P}(N([0,s])=0)\\ &= \lambda s\mathrm{e}^{-\lambda s}\cdot \mathrm{e}^{-\lambda s},\end{aligned}$$

because of independence and homogeneity. In the same way, the probability of exactly one call in $[s, 2s]$ and no calls in $[0, s)$ is equal to $\mathrm{e}^{-\lambda s}\cdot\lambda s\mathrm{e}^{-\lambda s}$. And indeed: $\lambda s\mathrm{e}^{-\lambda s}\cdot\mathrm{e}^{-\lambda s}+\mathrm{e}^{-\lambda s}\cdot\lambda s\mathrm{e}^{-\lambda s} = 2\lambda s\mathrm{e}^{-\lambda\cdot 2s}$.

12.2 Because there are 60 seconds in a minute, we have $60\lambda = 330$. It follows that $\lambda = 5\frac{1}{2}$. Since the interarrival times have an $Exp(\lambda)$ distribution, the expected time between messages is $1/\lambda = 0.18$ second.

12.3 The intensity of this process is $\lambda = 5$ per m^2. The area of the strip is $2\cdot(1/20) = 1/10\,\mathrm{m}^2$. Hence the probability that no defects occur in the strip is $\mathrm{e}^{-\lambda\cdot(\text{area of strip})} = \mathrm{e}^{-5\cdot(1/10)} = \mathrm{e}^{-1/2} = 0.60$.

12.6 Exercises

12.1 ⊞ In each of the following examples, try to indicate whether the Poisson process would be a good model.

- **a.** The times of bankruptcy of enterprises in the United States.
- **b.** The times a chicken lays its eggs.
- **c.** The times of airplane crashes in a worldwide registration.
- **d.** The locations of worngly spelled words in a book.
- **e.** The times of traffic accidents at a crossroad.

12.2 The number of customers that visit a bank on a day is modeled by a Poisson distribution. It is known that the probability of no customers at all is 0.00001. What is the expected number of customers?

12.3 Let N have a $Pois(4)$ distribution. What is $\mathrm{P}(N = 4)$?

12.4 Let X have a $Pois(2)$ distribution. What is $\mathrm{P}(X \le 1)$?

12.5 ⊡ The number of errors on a hard disk is modeled as a Poisson random variable with expectation one error in every Mb, that is, in every 2^{20} bytes.

- **a.** What is the probability of at least one error in a sector of 512 bytes?
- **b.** The hard disk is an 18.62-Gb disk drive with 39 054 015 sectors. What is the probability of at least one error on the hard disk?

12.6 ⊞ A certain brand of copper wire has flaws about every 40 centimeters. Model the locations of the flaws as a Poisson process. What is the probability of two flaws in 1 meter of wire?

12.7 ⊞ The Poisson model is sometimes used to study the flow of traffic ([15]). If the traffic can flow freely, it behaves like a Poisson process. A 20-minute time interval is divided into 10-second time slots. At a certain point along the highway the number of passing cars is registered for each 10-second time slot. Let n_j be the number of slots in which j cars have passed for $j = 0, \ldots, 9$. Suppose that one finds

j	0	1	2	3	4	5	6	7	8	9
n_j	19	38	28	20	7	3	4	0	0	1

Note that the total number of cars passing in these 20 minutes is 230.

a. What would you choose for the intensity parameter λ?

b. Suppose one estimates the probability of 0 cars passing in a 10-second time slot by n_0 divided by the total number of time slots. Does that (reasonably) agree with the value that follows from your answer in **a**?

c. What would you take for the probability that 10 cars pass in a 10-second time slot?

12.8 ⊡ Let X be a Poisson random variable with parameter μ.

a. Compute $\mathrm{E}[X(X-1)]$.

b. Compute $\mathrm{Var}(X)$, using that $\mathrm{Var}(X) = \mathrm{E}[X(X-1)] + \mathrm{E}[X] - (\mathrm{E}[X])^2$.

12.9 Let Y_1 and Y_2 be independent Poisson random variables with parameter μ_1, respectively μ_2. Show that $Y = Y_1 + Y_2$ also has a Poisson distribution. Instead of using the addition rule in Section 11.1 as in Exercise 11.2, you can prove this without doing any computations by considering the number of points of a Poisson process (with intensity 1) in two disjoint intervals of length μ_1 and μ_2.

12.10 Let X be a random variable with a $Pois(\mu)$ distribution. Show the following. If $\mu < 1$, then the probabilities $\mathrm{P}(X = k)$ are strictly decreasing in k. If $\mu > 1$, then the probabilities $\mathrm{P}(X = k)$ are first increasing, then decreasing (cf. Figure 12.1). What happens if $\mu = 1$?

12.11 ⊞ Consider the one-dimensional Poisson process with intensity λ. Show that the number of points in $[0, t]$, *given* that the number of points in $[0, 2t]$ is equal to n, has a $Bin(n, \frac{1}{2})$ distribution.
Hint: write the event $\{N([0, s]) = k, N([0, 2s]) = n\}$ as the intersection of the (independent!) events $\{N([0, s]) = k\}$ and $\{N((s, 2s]) = n - k\}$.

12.12 We consider the one-dimensional Poisson process. Suppose for some $a > 0$ it is given that there are exactly two points in $[0, a]$, or in other words: $N_a = 2$. The goal of this exercise is to determine the joint distribution of X_1 and X_2, the locations of the two points, conditional on $N_a = 2$.

a. Prove that for $0 < s < t < a$

$$\begin{aligned} &\mathrm{P}(X_1 \leq s, X_2 \leq t, N_a = 2) \\ &\qquad = \mathrm{P}(X_2 \leq t, N_a = 2) - \mathrm{P}(X_1 > s, X_2 \leq t, N_a = 2)\,. \end{aligned}$$

b. Deduce from **a** that

$$\mathrm{P}(X_1 \leq s, X_2 \leq t, N_a = 2) = \mathrm{e}^{-\lambda a}\left(\frac{\lambda^2 t^2}{2!} - \frac{\lambda^2 (t-s)^2}{2!}\right).$$

c. Deduce from **b** that for $0 < s < t < a$

$$\mathrm{P}(X_1 \leq s, X_2 \leq t \mid N_a = 2) = \frac{t^2 - (t-s)^2}{a^2}.$$

12.13 Walking through a meadow we encounter two kinds of flowers, daisies and dandelions. As we walk in a straight line, we model the positions of the flowers we encounter with a one-dimensional Poisson process with intensity λ. It appears that about one in every four flowers is a daisy. Forgetting about the dandelions, what does the process of the *daisies* look like? This question will be answered with the following steps.

a. Let N_t be the total number of flowers, X_t the number of daisies, and Y_t be the number of dandelions we encounter during the first t minutes of our walk. Note that $X_t + Y_t = N_t$. Suppose that each flower is a daisy with probability $1/4$, independent of the other flowers. Argue that

$$\mathrm{P}(X_t = n, Y_t = m \mid N_t = n + m) = \binom{n+m}{n}\left(\frac{1}{4}\right)^n\left(\frac{3}{4}\right)^m.$$

b. Show that

$$\mathrm{P}(X_t = n, Y_t = m) = \frac{1}{n!}\frac{1}{m!}\left(\frac{1}{4}\right)^n\left(\frac{3}{4}\right)^m \mathrm{e}^{-\lambda t}(\lambda t)^{n+m},$$

by conditioning on N_t and using **a**.

c. By writing $\mathrm{e}^{-\lambda t} = \mathrm{e}^{-(\lambda/4)t}\mathrm{e}^{-(3\lambda/4)t}$ and summing over m, show that

$$\mathrm{P}(X_t = n) = \frac{1}{n!}\mathrm{e}^{-(\lambda/4)t}\left(\frac{\lambda t}{4}\right)^n.$$

Since it is clear that the numbers of daisies that we encounter in disjoint time intervals are independent, we may conclude from **c** that the process (X_t) is again a Poisson process, with intensity $\lambda/4$. One often says that the process (X_t) is obtained by *thinning* the process (N_t). In our example this corresponds to picking all the dandelions.

12.14 ⊡ In this exercise we look at a simple example of random variables X_n that have the property that their distributions converge to the distribution of a random variable X as $n \to \infty$, while it is *not* true that their expectations converge to the expectation of X. Let for $n = 1, 2, \ldots$ the random variables X_n be defined by

$$\mathrm{P}(X_n = 0) = 1 - \frac{1}{n} \quad \text{and} \quad \mathrm{P}(X_n = 7n) = \frac{1}{n}.$$

a. Let X be the random variable that is equal to 0 with probability 1. Show that for all a the probability mass functions $p_{X_n}(a)$ of the X_n converge to the probability mass function $p_X(a)$ of X as $n \to \infty$. Note that $\mathrm{E}[X]=0$.

b. Show that nonetheless $\mathrm{E}[X_n] = 7$ for all n.