MAS275 Probability Modelling

6 Poisson processes

6.1 Introduction

Poisson processes are a particularly important topic in probability theory. The one-dimensional Poisson process, which most of this section will be about, is a model for the random times of occurrences of instantaneous events; there are many examples of things whose random occurrences in time can be modelled by Poisson processes, for example customers arriving in a queue, incoming calls to a phone, eruptions of a volcano, and so on. Higher dimensional analogues of the basic Poisson process are also useful as models for random locations of objects in space; we will discuss this in section 6.9.

We will firstly recall some properties of the Poisson distribution before moving onto the definition of the process.

6.2 Properties of the Poisson distribution

Recall that the Poisson distribution with parameter λ, denoted $Po(\lambda)$, has probability function

$$p(x) = \frac{\lambda^x e^{-\lambda}}{x!}, x = 0, 1, 2, 3, \ldots$$

and a random variable with $Po(\lambda)$ distribution has $E(X) = \mathrm{Var}(X) = \lambda$.

Some further properties of the Poisson distribution follow.

Proposition 16. *The Poisson distribution is additive.*

More precisely, if X and Y are independent random variables with distributions $Po(\lambda)$ and $Po(\mu)$ respectively, then $X+Y$ has the distribution $Po(\lambda+\mu)$.

This property extends in an obvious way to more than two independent random variables.

Proof. See Exercise 17(b)(i). □

Proposition 17. *The Poisson distribution is a limit of binomial distributions.*

Specifically, if we consider the sequence of binomial distributions $Bi(n, \mu/n)$ for fixed μ, then the number of trials n is increasing but the probability of success on a trial is μ/n which is decreasing in proportion, so that the mean μ remains fixed. In the limit, the distribution is Poisson with parameter μ.

Proof. This was shown in MAS113, section 9.3. The proof is repeated below, but will not be covered in lectures.

To verify this, we look at the probability function of $Bi(n, \mu/n)$

$$\begin{aligned}\binom{n}{x}\left(\frac{\mu}{n}\right)^x\left(1-\frac{\mu}{n}\right)^{n-x} &= \frac{n(n-1)\dots(n-x+1)}{x!}\frac{\mu^x}{n^x}\left(1-\frac{\mu}{n}\right)^n\left(1-\frac{\mu}{n}\right)^{-x} \\ &= 1\left(1-\frac{1}{n}\right)\dots\left(1-\frac{(x-1)}{n}\right)\frac{\mu^x}{x!}\left(1-\frac{\mu}{n}\right)^n\left(1-\frac{\mu}{n}\right)^{-x} \\ &\to 1\cdot\frac{\mu^x}{x!}e^{-\mu}\cdot 1\end{aligned}$$

as $n \to \infty$. This is the probability function of $Po(\mu)$. □

6.3 The basic Poisson process

We now move to a continuous time scale which is usually regarded as starting at time zero, so that it consists of the positive real numbers, and denote our time variable typically by t. When we refer to a time interval, we adopt the convention that it excludes the left hand end point but includes the right hand end-point, say

$$(u, v] = \{t : u < t \le v\}.$$

The Poisson process is described in terms of the random variables $N_{u,v}$ for $0 \le u \le v$, where $N_{u,v}$ is the number of occurrences in the time interval $(u, v]$. The process has one parameter, which is a positive number λ known as the **rate** of the process and which is meant to measure the average or expected number of occurrences per unit time.

The basic Poisson process is then defined by the following two assumptions:

(a). For any $0 \le u \le v$, the distribution of $N_{u,v}$ is Poisson with parameter $\lambda(v-u)$.

(b). If $(u_1, v_1], (u_2, v_2], \ldots, (u_k, v_k]$ are disjoint time intervals then $N_{u_1,v_1}, N_{u_2,v_2}, \ldots, N_{u_k,v_k}$ are independent random variables.

Note that by assumption (a) and the mean of a Poisson random variable, $N_{u,v}$ has mean $\lambda(v-u)$, which gives the correct interpretation to the rate λ as described above.

To show that a process actually exists which satisfies these assumptions still requires a bit of work. At this point, we can observe that the special properties of Poisson distributions are important: if say $u < v < w$ then

$$N_{u,w} = N_{u,v} + N_{v,w}$$

and because all three of these random variables are to have Poisson distributions, and the two on the right hand side are independent, the assumptions can only work because of the additivity property of Poisson random variables. We will see a bit more on justifying the existence of a process satisfying the assumptions at the end of section 6.8.

We can also ask about why the assumption of Poisson distributions might make sense in a modelling context. To see this, approximate the continuous time scale by dividing $(0, t]$ up into n small intervals, and assume that there is the same small probability of an occurrence in each, assuming that the probability of more than one occurrence in an interval is negligible. In order to fix the expected number of occurrences at λt, this probability must be set equal to $\lambda t/n$. Then the number of occurrences in the interval has the binomial distribution $Bi(n, \lambda t/n)$. As we let n tend to infinity, the small intervals become smaller and so the approximation to the continuous time scale becomes closer, and, by the relationship between the Poisson and Binomial distributions (see Proposition 17) the distribution of the number of occurrences approaches the Poisson distribution with parameter λt. This suggests the Poisson distribution as a sensible model in the genuinely continuous time setting.

You might like to think about which other distributions you have encountered might have assumptions similar to (a) and (b) which work.

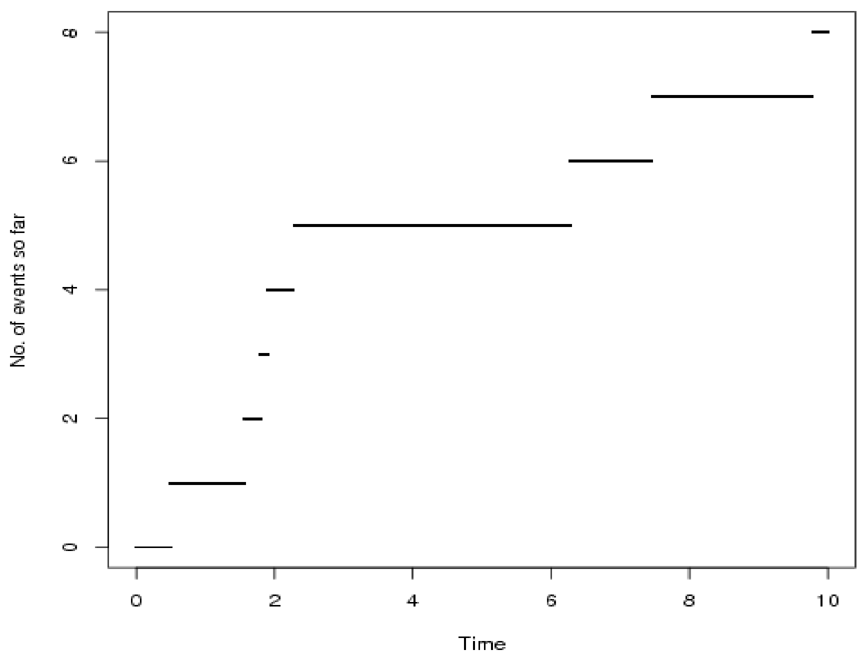

Figure 1: A simulation of a Poisson process with rate 1 up to time 10

Example 33. *Volcanic eruptions*

6.4 Inter-occurrence times

Let T_1 denote the length of time until the first occurrence, T_2 denote the length of time between the first and second occurrences, and so on, so that T_n represents the time between occurrences $n-1$ and n. These random variables are called **inter-occurrence times**. We first show that these cannot be zero, and we will then show which distribution they have.

Theorem 18. *The probability that in $(0,t]$ two occurrences of a Poisson process with rate λ occur at exactly the same time is zero.*

Proof. Consider dividing $(0,t]$ into n small intervals, as in the justification of the Poisson model above. Each of these small intervals will be of the form $((i-1)t/n, it/n]$ for some i, and has length t/n, so the number of occurrences in any small interval has a Poisson distribution with parameter $\lambda t/n$. Hence the probability there are at least two occurrences in a given small interval is

$$1 - e^{-\frac{\lambda t}{n}} - \frac{\lambda t}{n} e^{-\frac{\lambda t}{n}} = 1 - e^{-\frac{\lambda t}{n}} \left(1 + \frac{\lambda t}{n}\right).$$

Let Y_n be the number of the small intervals which have at least two occurrences. By the independence assumptions, Y_n will have a Binomial distribution:

$$Y_n \sim Bin\left(n, 1 - e^{-\frac{\lambda t}{n}} \left(1 + \frac{\lambda t}{n}\right)\right),$$

and hence

$$P(Y_n = 0) = \left(e^{-\frac{\lambda t}{n}} \left(1 + \frac{\lambda t}{n}\right)\right)^n = e^{-\lambda t} \left(1 + \frac{\lambda t}{n}\right)^n.$$

As $\left(1 + \frac{\lambda t}{n}\right)^n \to e^{\lambda t}$ as $n \to \infty$, we have that $P(Y_n = 0) \to 1$ as $n \to \infty$. But if there were a probability $p > 0$ that there were two occurrences at exactly the same time, we would have $P(Y_n = 0) < 1 - p$ for all n, so this cannot be the case. Hence the probability of there being two occurrences at exactly the same time is 0. □

Theorem 19. *Inter-occurrence times are independent of each other, and are exponentially distributed with parameter λ.*

Proof. (Sketch) That T_1 has this distribution is straightforward to show. First, $P(T_1 \leq t)$ is the probability that the first occurrence has happened by time t, so is the probability that there is at least one occurrence in $(0, t]$. Hence

$$P(T_1 \leq t) = P(N_{0,t} \geq 1) = 1 - P(N_{0,t} = 0) = 1 - e^{-\lambda t},$$

which is the distribution function of an exponential distribution with parameter λ.

Now consider the probability that T_n is at most t, conditional on the values of $T_1, T_2, \ldots, T_{n-1}$,

$$P(T_n \leq t | T_1 = t_1, T_2 = t_2, \ldots, T_{n-1} = t_{n-1}).$$

(There are some technical details to deal with the fact that we are conditioning on an event of probability zero, which is why this proof is a "sketch".) Similarly to the above argument, this is the probability that there is at least one occurrence between times $t_1 + t_2 + \ldots + t_{n-1}$ and $t_1 + t_2 + \ldots + t_{n-1} + t$, which again is $1 - e^{-\lambda t}$. So, regardless of the values taken by $T_1, T_2, \ldots, T_{n-1}$, the conditional distribution of T_n is exponential with parameter λ, which implies the result. □

6.5 Variable rate Poisson process

We can make the basic Poisson process more flexible and realistic as a model by allowing the rate of the process to vary with time, $\lambda(t)$ say. This can take into account, for example, the fact that traffic is heavier at rush hours, the rate of emission of particles from a radioactive isotope declines with time, and so on.

The only change to the definition of the Poisson process is that the assumption (a) is replaced by the following:

For any $0 \le u \le v$, the distribution of $N_{u,v}$ is Poisson with parameter $\int_u^v \lambda(t)dt$.

This assumption generalises that of the constant rate case, and gives the correct interpretation of "rate" when this rate is varying. If $\lambda(t)$ is a constant we recover the basic Poisson process.

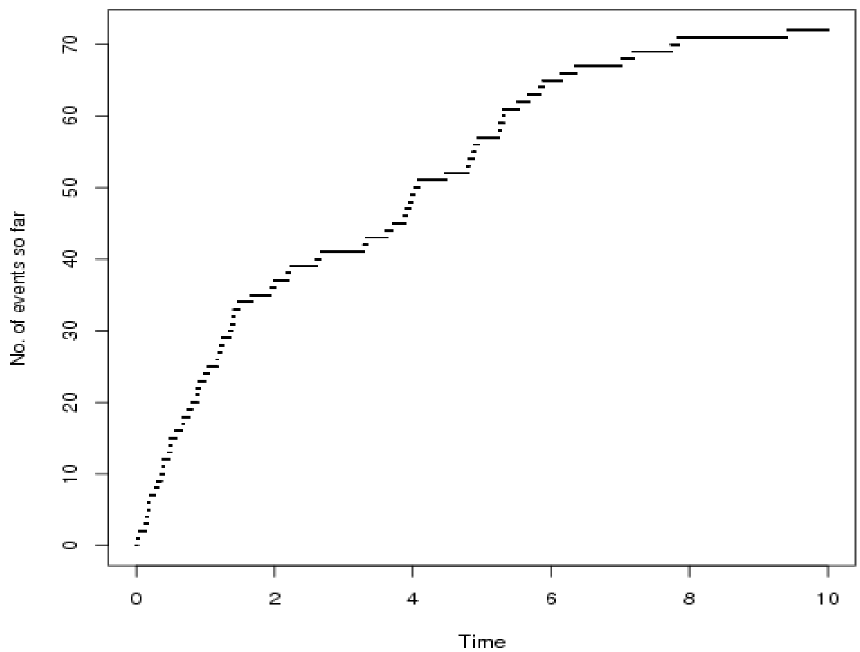

Figure 2: A simulation of a variable rate Poisson process with rate $30/(t+1)$ up to time 10

The independence assumption (b) still holds when the rate is variable; this still works because of the additivity property of Poisson random variables, and also because of the additivity of integrals, namely that if $u < v < w$ then

$$\int_u^w \lambda(t)dt = \int_u^v \lambda(t)dt + \int_v^w \lambda(t)dt.$$

Example 34. *Email arrivals*

6.6 Superposition

In some situations we have more than one Poisson process running. If these are independent, then the process formed by combining them is also a Poisson process.

Theorem 20. *Let $(N_{u,v})$ and $(M_{u,v})$ be independent Poisson processes with (possibly variable) rates $\lambda(t)$ and $\mu(t)$ respectively. Then $(N_{u,v} + M_{u,v})$ also forms a Poisson process, with rate $\lambda(t) + \mu(t)$.*

Proof. Because $(N_{u,v})$ and $(M_{u,v})$ are Poisson processes,

$$N_{u,v} \sim Po\left(\int_u^v \lambda(t)\, dt\right)$$

and

$$M_{u,v} \sim Po\left(\int_u^v \mu(t)\, dt\right).$$

Because they are independent, the additivity of the Poisson distribution (Proposition 16) tells us that

$$\begin{aligned} N_{u,v} + M_{u,v} &\sim Po\left(\int_u^v \lambda(t)\, dt + \int_u^v \mu(t)\, dt\right) \\ &= Po\left(\int_u^v (\lambda(t) + \mu(t))\, dt\right). \end{aligned}$$

The independence of the number of occurrences in disjoint intervals in the combined process follows from the same property of the two original processes.

□

6.7 Marking and thinning

Sometimes the occurrences in a Poisson process may be categorised as each belonging to one of a number of types. This is sometimes referred to as

marking: think of each occurrence as being given a random "mark". Specifically we will assume that each occurrence in a Poisson process with (possibly variable) rate $\lambda(t)$ is given, independently of everything else, one of k different marks with probabilities $p_1, p_2, \ldots, p_k$ respectively. Write the total number of occurrences in $(u, v]$ as $N_{u,v}$ (as before) and write the number of occurrences of type i in $(u, v]$ as $N_{u,v}^{(i)}$.

The following result about Poisson random variables will be useful.

Lemma 21. *Let X be a Poisson random variable with parameter μ, and imagine that, conditional on $X = x$, we have x objects each of which is of one of k types. Assume further that each of these objects is of type i with probability p_i, independently of the other objects. Let the number of objects of type i be Y_i.*

The (unconditional) joint distribution of $Y_1, Y_2, \ldots, Y_k$ is such that they are independent Poisson, with parameters $p_1\mu, p_2\mu, \ldots$ and $p_k\mu$ respectively.

Proof. For $y_1, y_2, \ldots y_k = 0, 1, 2, \ldots$, and letting

$$x = y_1 + y_2 + \ldots + y_k,$$

$$\begin{aligned} P(Y_1 = y_1, Y_2 = y_2, \ldots, Y_k = y_k) &= P(X = x)P(Y_1 = y_1, Y_2 = y_2, \ldots, Y_k = y_k | X = x) \\ &= e^{-\mu}\frac{\mu^x}{x!}\frac{x!}{y_1!y_2!\ldots y_k!}p_1^{y_1}p_2^{y_2}\ldots p_k^{y_k} \\ &= e^{-p_1\mu}\frac{(p_1\mu)^{y_1}}{y_1!}e^{-p_2\mu}\frac{(p_2\mu)^{y_2}}{y_2!}\ldots e^{-p_k\mu}\frac{(p_k\mu)^{y_k}}{y_k!} \end{aligned}$$

as required. □

Theorem 22. *For each i, the process given by $(N_{u,v}^{(i)})$ (counting the occurrences which are type i) is a Poisson process with rate $\lambda(t)p_i$, and the k processes for the different types are independent of each other.*

Proof. This essentially follows from Lemma 21 together with the independence properties of Poisson processes. □

It is even possible to allow the probabilities of the marks to be dependent on time, say $p_1(t), p_2(t), \ldots, p_k(t)$. Then the generalised result is that the marked processes are independent Poisson processes with variable rates $p_1(t)\lambda(t), p_2(t)\lambda(t), \ldots, p_k(t)\lambda(t)$ respectively.

One special case of marking is where $k = 2$ and the process of marking consists of either retaining the occurrence, with probability p, or deleting it, with probability $q = 1 - p$. Then the process of retained points is Poisson with rate $p\lambda$, and in this context the property is often known as the **thinning** property.

Example 35. *University applications*

6.8 Conditioning on the number of occurrences in an interval

Sometimes we know how many occurrences there are in a given interval, and are interested in how they are distributed within the interval.

Theorem 23. *Assume that we have a Poisson process with constant rate λ. Given that there are n occurrences in the time interval $(0, t]$ say, the positions of these occurrences are distributed as a random sample of size n from the uniform distribution on that interval.*

Note that this implies that, conditional on there being n occurrences in $(0, t]$, the number of occurrences in any interval $(u, v] \subseteq (0, t]$ (so $0 \leq u < v \leq t$) has a $Bi(n, (v - u)/t)$ distribution, as each of the n occurrences would have probability $(v - u)/t$ of being in $(u, v]$, independently of the others. We will prove this latter version of the statement.

Proof. Note that $N_{0,t} \sim Po(\lambda t)$, $N_{u,v} \sim Po(\lambda(v - u))$ and $N_{0,t} - N_{u,v} = N_{0,u} + N_{v,t} \sim Po(\lambda(t - (v - u))$, and the latter two random variables are independent. Thus, for $0 \le a \le n$,

$$\begin{aligned}
P(N_{u,v}=a|N_{0,t}=n) &= \frac{P(N_{u,v}=a, N_{0,t}=n)}{P(N_{0,t}=n)} \\
&= \frac{P(N_{u,v}=a, N_{0,t}-N_{u,v}=n-a)}{P(N_{0,t}=n)} \\
&= \frac{P(N_{u,v}=a)P(N_{0,t}-N_{u,v}=n-a)}{P(N_{0,t}=n)} \\
&= \frac{(\lambda(v-u))^a e^{-\lambda(v-u)}}{a!} \frac{(\lambda(t-(v-u)))^{n-a} e^{-\lambda(t-(v-u))}}{(n-a)!} \frac{n!}{(\lambda t)^n e^{-\lambda t}} \\
&= \frac{e^{-\lambda(v-u)} e^{-\lambda(t-(v-u))}}{e^{-\lambda t}} \frac{n!}{a!(n-a)!} \frac{\lambda^a (v-u)^a \lambda^{n-a} (t-(v-u))^{n-a}}{(\lambda t)^n} \\
&= \binom{n}{a} \left(\frac{v-u}{t}\right)^a \left(1-\frac{v-u}{t}\right)^{n-a},
\end{aligned}$$

which is the probability that a $Bi(n,(v-u)/t)$ random variable takes the value a, as required. □

This result generalises to the variable rate case, but the uniform distribution is replaced by the distribution which has p.d.f.

$$f(s) = \frac{\lambda(s)}{\int_0^t \lambda(x)dx},$$

which is the distribution on $(0,t]$ whose density is proportional to the rate of the original process. So the number of occurrences in any interval $(u,v] \subseteq (0,t]$ has a

$$Bi\left(n, \frac{\int_u^v \lambda(s)ds}{\int_0^t \lambda(s)ds}\right)$$

distribution. The proof is essentially the same.

Example 36. *Conditioning on number of events*

Note that we can actually reverse this idea to construct a Poisson process, for example for simulation purposes, or to convince ourselves that Poisson processes really exist. Assuming that we want to construct a variable rate Poisson process with rate $\lambda(t)$, we can do the following:

- Divide the positive real line up into intervals $(n-1, n]$ for each positive integer n.
- To each of these intervals $(n-1, n]$ assign a Poisson random variable X_n with parameter $\int_{n-1}^{n} \lambda(t)\, dt$. These Poisson random variables should be independent of each other. (This assumes this integral is finite; if for one of the intervals it is not we will need to be more careful.)
- If $X_n = 0$, then there will be no occurrences in the interval $(n-1, n]$; if $X_n = x > 0$, then we create a random sample of x random variables on $(n-1, n]$ with probability density function $f(s) = \frac{\lambda(s)}{\int_{n-1}^{n} \lambda(x) dx}$, in a similar manner to above. The values of these random variables will give the times of occurrences in the interval.

It is not too hard to show, using the marking and additivity properties of the Poisson distribution, that a process of occurrences constructed in this way will satisfy the assumptions with which we defined the Poisson process.

6.9 The spatial Poisson process

An important generalisation of the basic Poisson process is to replace the time scale with a space, and the aim is to model a random scattering of points in this space. The space may be one-dimensional – for example if we wish to consider defects on a length of cable – and in that case it looks like the time scale, or it may be in a higher dimension – two dimensions for positions of spots of rain on a pavement, three dimensions for positions of stars in space, for example.

To generalise the assumptions we made for the basic Poisson process, we need an analogue of the length of a time interval. The natural way to do this is to consider length in one dimension, area in two dimensions, volume in three dimensions, and so on. We will refer to length, area or volume, as appropriate, as **measure**, and denote the measure of the set A by $|A|$. [For those who are familiar with the mathematical concept of a measure, we can use other measures on our space here in place of length, area or volume.]

The behaviour of the process can be described by random variables $N(A)$ for subsets A with finite measure: $N(A)$ represents the number of points of the process which fall inside the set A. The parameter λ is in this context called the **density** of the process.

A **spatial Poisson process** is now defined to be a process which satisfies the following assumptions, which are generalisations of those we used in the time setting.

(a). For any set A of finite measure, $N(A)$ has the Poisson distribution with parameter $\lambda|A|$.

(b). If $A_1, A_2, \ldots, A_k$ are disjoint sets of finite measure, then $N(A_1), N(A_2), \ldots, N(A_k)$ are independent random variables.

The assumptions work for essentially the same reasons as before, notably the additivity of independent Poisson random variables. Most of the properties of the basic Poisson process have analogues in more than one dimension:

Superposition Two independent spatial Poisson processes with rates λ and μ can be combined to form a spatial Poisson process with rate $\lambda + \mu$.

Marking If the points of a spatial Poisson process with rate λ are given independent marks (from $1, 2, \ldots, k$) with probabilities $p_1, p_2, \ldots, p_k$ then the points with mark i form a spatial Poisson process with rate λp_i, and the processes corresponding to the different marks are independent.

Conditioning If we know that there are n points of the process in A, then the conditional distribution of the location of the points is that of a random sample of size n from the uniform distribution on A. In particular, if $B \subseteq A$, then the number of points in B is Binomial with parameters n and $|B|/|A|$.

The proofs of all of these properties are natural generalisations of the proofs of the one-dimensional versions.

It is also possible to simulate from a spatial Poisson process in a similar way to the one described for the variable rate time Poisson process.

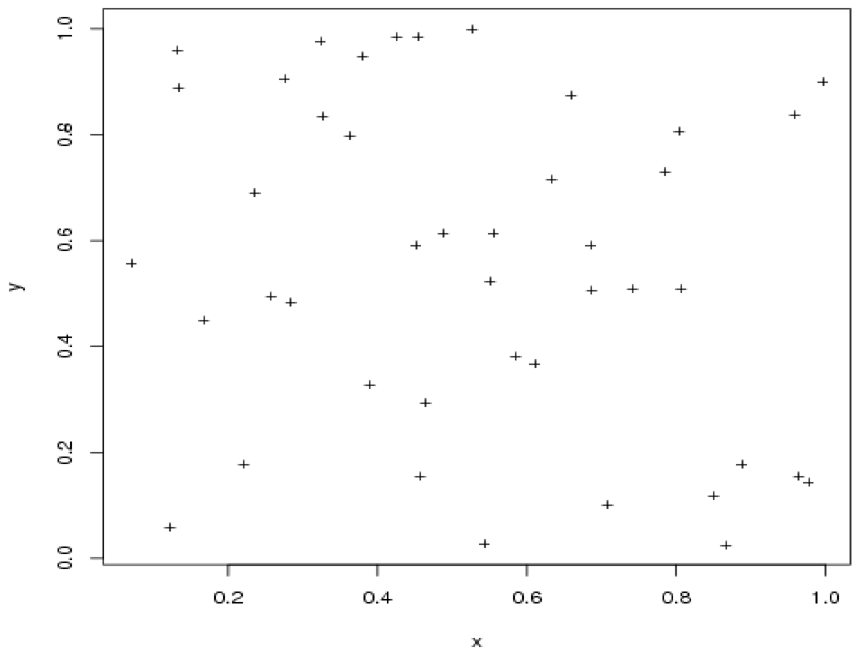

Figure 3: A simulation of a spatial Poisson process with rate 40 on a unit square

One property of the basic Poisson process which does not naturally carry over is the joint distribution of inter-occurrence times, as there is no natural ordering of points in two or more dimensions, and so the analogue of inter-occurrence times does not exist. However, it is possible to use a similar idea to calculate the distribution of the distance to the nearest point in the process from a given point.

Example 37. *Trees in a forest*

It is also possible to define variable rate spatial Poisson processes; the parameter of the Poisson distribution giving the number of points in a set A will be the integral over A of the rate function, just as for the one-dimensional case, but the integral is now a multidimensional one.

An actual scatter of points may be **clustered** relative to the true randomness of the Poisson process – for example, positions of plants each of which self-propagates within a local area – or **regular** relative to the Poisson process – for example, positions of birds' nests when there is a territorial effect inhibiting nests from being too close together.

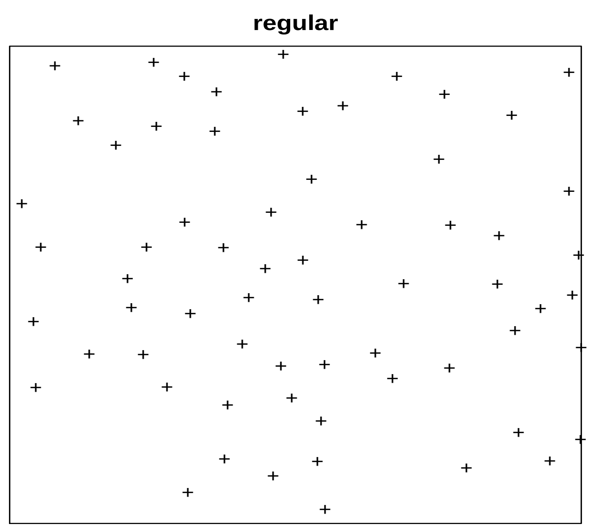

Figure 4: A simulation of a spatial process more regular than a Poisson process

6.10 Compound Poisson processes*

Suppose that events occur at random times but also each event carries with it some numerical value, and the chief interest is in the sum of these numerical values over a period of time. Examples might be claims on an insurance company, which occur at random times but they differ in size; or fatalities in

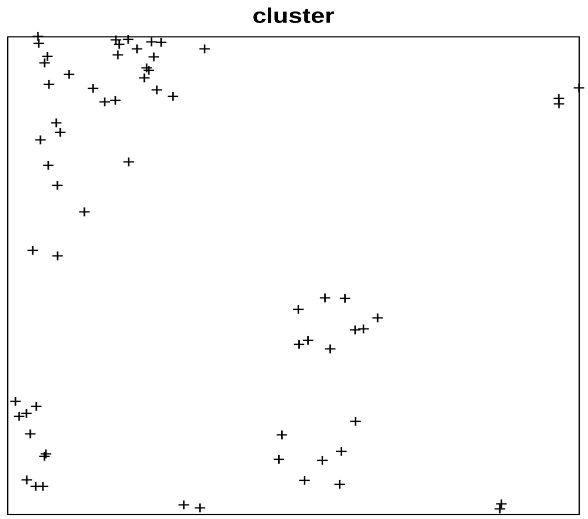

Figure 5: A simulation of a spatial process more clustered than a Poisson process

road accidents, where the accidents occur at random times but each accident may incur a number of deaths.

The simplest model for such situations is to take the times of occurrences as a basic Poisson process and then to assume that the sizes of the occurrences are random variables each with some known distribution, which may be discrete or continuous, these random variables being independent of each other and of the times of the occurrences. This gives what is known as a **compound Poisson process**.

We will use the following notation. $N(t)$ denotes the number of occurrences in the time interval $(0, t]$, previously written as $N_{0,t}$; the sizes of the occurrences in chronological order are denoted by $Y_1, Y_2, \ldots$. Then the sum of these over the time interval $(0, t]$ may be written

$$X(t) = \sum_{i=1}^{N(t)} Y_i.$$

If we think graphically of plotting $N(t)$ and $X(t)$ against t, then the graph of $N(t)$ jumps upwards by 1 at each occurrence and stays constant in between, whereas the graph of $X(t)$ jumps (upwards or downwards, since Y_i could be negative) by the random quantities $Y_1, Y_2, \ldots$ at the times of the occurrences, and stays constant in between.

The process is completely specified by the rate λ of the Poisson process and the common distribution of the random variables $Y_1, Y_2, \ldots$.

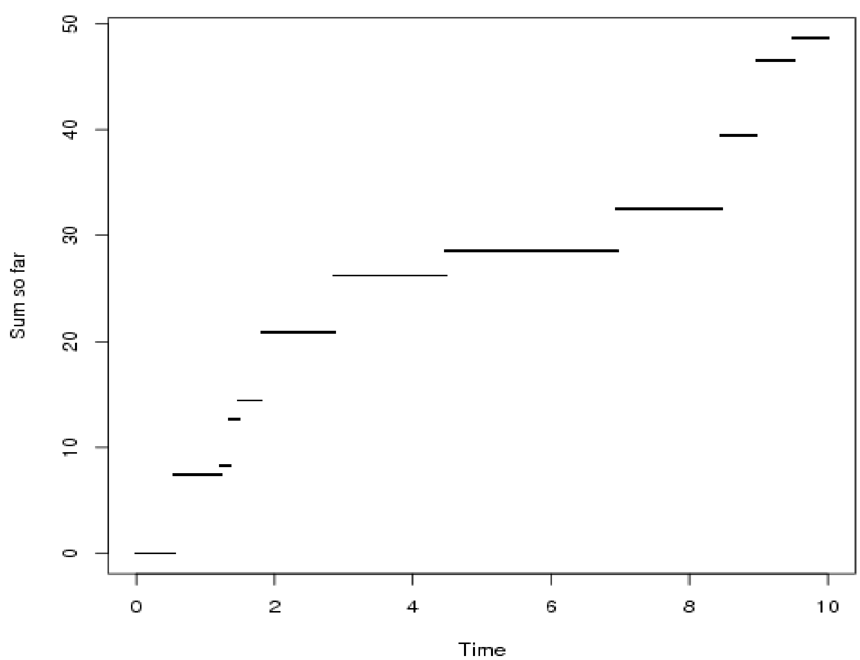

Figure 6: A simulation of a compound Poisson process with rate 1 and jumps having χ^2_4 distribution up to time 10

We can combine the ideas of compound Poisson processes with variable rate and spatial Poisson processes as well.

For example, we might be modelling the locations of nests of some species of bird within some region. We could treat the locations as points in two dimensional space to be modelled by a spatial Poisson process. If some parts of the region are more favourable to the species than others, then we would expect a higher density of nests in these areas, so the model would have a variable rate which is higher in favourable areas and lower elsewhere. If we wanted to model the total number of offspring raised, then a compound spatial Poisson model might be appropriate.