Chapter 10

Poisson processes

10.1 Overview

The Binomial distribution and the geometric distribution describe the behavior of two random variables derived from the random mechanism that I have called coin tossing. The name *coin tossing* describes the whole mechanism; the names *Binomial* and *geometric* refer to particular aspects of that mechanism. If we increase the tossing rate to n tosses per second and decrease the probability of heads to a small p, while keeping the expected number of heads per second fixed at $\lambda = np$, the number of heads in a t second interval will have approximately a Bin(nt, p) distribution, which is close to the Poisson(λt). Also, the numbers of heads tossed during disjoint time intervals will still be independent random variables. In the limit, as $n \to \infty$, we get an idealization called a *Poisson process*.

Remark. The double use of the name Poisson is unfortunate. Much confusion would be avoided if we all agreed to refer to the mechanism as "idealized-very-fast-coin-tossing", or some such. Then the Poisson distribution would have the same relationship to idealized-very-fast-coin-tossing as the Binomial distribution has to coin-tossing. Conversely, I could create more confusion by renaming coin tossing as "the binomial process". Neither suggestion is likely to be adopted, so you should just get used to having two closely related objects with the name Poisson.

Definition. A Poisson process with rate λ on $[0, \infty)$ is a random mechanism that generates "points" strung out along $[0, \infty)$ in such a way that

(i) the number of points landing in any subinterval of length t is a random variable with a Poisson (λt) distribution

(ii) the numbers of points landing in disjoint (= non-overlapping) intervals are independent random variables.

It often helps to think of $[0, \infty)$ as time.

Note that, for a very short interval of length δ , the number of points N in the interval has a Poisson($\lambda\delta$) distribution, with

$$\mathbb{P}\{N=0\} = e^{-\lambda\delta} = 1 - \lambda\delta + o(\delta)$$
$$\mathbb{P}\{N=1\} = \lambda\delta e^{-\lambda\delta} = \lambda\delta + o(\delta)$$
$$\mathbb{P}\{N\geq 2\} = 1 - e^{-\lambda\delta} - \lambda\delta e^{-\lambda\delta} = o(\delta).$$

. .

When we pass to the idealized mechanism of points generated in continuous time, several awkward details of discrete-time coin tossing disappear.

Example <10.1> (Gamma distribution from Poisson process) The waiting time W_k to the *k*th point in a Poisson process with rate λ has a continuous distribution, with density $g_k(w) = \lambda^k w^{k-1} e^{-\lambda w} / (k-1)!$ for w > 0, zero otherwise.

It is easier to remember the distribution if we rescale the process, defining $T_k = \lambda W_k$. The new T_k has a continuous distribution with a **gamma**(k) **density**,

$$f_k(t) = \frac{t^{k-1}e^{-t}}{(k-1)!} \mathbb{I}\{t > 0\}$$

Remark. Notice that $g_k = f_k$ when $\lambda = 1$. That is, T_k is the waiting time to the *k*th point for a Poisson process with rate 1. Put another way, we can generate a Poisson process with rate λ by taking the points appearing at times $0 < T_1 < T_2 < T_3 < \ldots$ from a Poisson process with rate 1, then rescaling to produce a new process with points at

$$0 < \frac{T_1}{\lambda} < \frac{T_2}{\lambda} < \frac{T_3}{\lambda} < \dots$$

You could verify this assertion by checking the two defining properties for a Poisson process with rate λ . Doesn't it makes sense that, as λ gets bigger, the points appear more rapidly?

For k = 1, Example <10.1> shows that the waiting time, W_1 , to the first point has a continuous distribution with density $\lambda e^{-\lambda w} \mathbf{1}\{w > 0\}$, which is called the *exponential distribution with expected value* 1/ λ . (You should check that $\mathbb{E}W_1 = 1/\lambda$.) The random variable λW_1 has a *standard exponential distribution*, with density $f_1(t) = e^{-t} \mathbf{1}\{t > 0\}$ and expected value 1.

Remark. I write the exponential distribution symbolically as "exp, mean $1/\lambda$ ". Do you see why the name $\exp(1/\lambda)$ would be ambiguous? Don't confuse the exponential density (or the exponential distribution that it defines) with the exponential function.

Just as for coin tossing, the independence properties of the Poisson process ensures that the times $W_1, W_2 - W_1, W_3 - W_2, \ldots$ are independent, each with the same distribution. You can see why this happens by noting that the future evolution of the process after the occurence of the first point at time W_1 is just a Poisson process that is independent of everything that happened up to time W_1 . In particular, the standardized time $T_k = \lambda W_k$, which has a gamma(k) distribution, is a sum of independent random variables $Z_1 = \lambda W_1, Z_2 = \lambda (W_2 - W_1), \ldots$ each with a standard exponential distribution.

The gamma density can also be defined for fractional values $\alpha > 0$:

$$f_{\alpha}(t) = \frac{t^{\alpha-1}e^{-t}}{\Gamma(\alpha)} \mathbf{1}\{t > 0\}$$

is called the *gamma*(α) *density*. The scaling constant, $\Gamma(\alpha)$, which ensures that the density integrates to one, is given by

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha - 1} e^{-x} dx \quad \text{for each } \alpha > 0.$$

The function $\Gamma(\cdot)$ is called the *gamma function*. Don't confuse the gamma density (or the gamma distribution that it defines) with the gamma function.

Example <10.2> Facts about the gamma function: $\Gamma(k) = (k-1)!$ for k = 1, 2, ..., and $\Gamma(1/2) = \sqrt{\pi}$.

The change of variable used in Example $\langle 10.2 \rangle$ to prove $\Gamma(1/2) = \sqrt{\pi}$ is essentially the same piece of mathematics as the calculation to find the density for the distribution of $Y = Z^2/2$ when $Z \sim N(0, 1)$. The random variable Y has a gamma(1/2) distribution.

Example <10.3> Moments of the gamma distribution

Poisson processes are often used as the simplest model for stochastic processes that involve arrivals at random times.

Example <10.4> A process with random arrivals

Poisson Processes can also be defined for sets other than the half-line.

Example <10.5> A Poisson Process in two dimensions.

10.2 Things to remember

Analogies between coin tossing, as a discrete time mechanism, and the Poisson process, as a continuous time mechanism:

DISCRETE TIME	\leftrightarrow	CONTINUOUS TIME
coin tossing, prob p of heads	\leftrightarrow	Poisson process with rate λ
$\overline{\operatorname{Bin}(n,p)}$	\leftrightarrow	$Poisson(\lambda t)$
X = #heads in n tosses		X = # points in $[a, a + t]$
$\mathbb{D}(\mathbf{V}_{i}) = (n) i n - i$		$\mathbb{T}(\mathbf{V}_{i}) = \frac{\lambda t}{\lambda} (\lambda_{i}) \frac{i}{\lambda} \frac{i}{\lambda}$
$\mathbb{P}\{X = i\} = \binom{n}{i} p^i q^{n-i}$		$\mathbb{P}\{X = i\} = e^{-\lambda t} (\lambda t)^{i} / i!$
for $i = 0, 1,, n$		for $i = 0, 1, 2$
geometric(p)	\leftrightarrow	(standard) exponential
$N_1 = \#$ tosses to first head;		T_1/λ = time to first point;
$\mathbb{P}\{N_1 = 1 + i\} = q^i p$		T_1 has density $f_1(t) = e^{-t}$
$\mathbb{P}\{N_1 = 1 + i\} = q^i p$ for $i = 0, 1, 2,$		T_1 has density $f_1(t) = e^{-t}$ for $t > 0$
$ \mathbb{P}\{N_1 = 1 + i\} = q^i p $ for $i = 0, 1, 2, \dots$ negative binomial	\leftrightarrow	T_1 has density $f_1(t) = e^{-t}$ for $t > 0$ gamma
$\mathbb{P}\{N_1 = 1 + i\} = q^i p$ for $i = 0, 1, 2, \dots$ negative binomial	\leftrightarrow	T_1 has density $f_1(t) = e^{-t}$ for $t > 0$ gamma
$ \mathbb{P}\{N_1 = 1 + i\} = q^i p $ for $i = 0, 1, 2, \dots$ negative binomial	\leftrightarrow	T_1 has density $f_1(t) = e^{-t}$ for $t > 0$ gamma
$\mathbb{P}\{N_1 = 1 + i\} = q^i p$ for $i = 0, 1, 2, \dots$ negative binomial See HW10	\leftrightarrow	T_1 has density $f_1(t) = e^{-t}$ for $t > 0$ gamma T_k has density
$\mathbb{P}\{N_1 = 1 + i\} = q^i p$ for $i = 0, 1, 2, \dots$ negative binomial See HW10	\leftrightarrow	$T_1 \text{ has density } f_1(t) = e^{-t}$ for $t > 0$ gamma $T_k \text{ has density}$ $f_k(t) = t^{k-1}e^{-t}/k! \text{ for } t > 0$
$\mathbb{P}\{N_1 = 1 + i\} = q^i p$ for $i = 0, 1, 2, \dots$ negative binomial See HW10	\leftrightarrow	$T_1 \text{ has density } f_1(t) = e^{-t}$ for $t > 0$ gamma $T_k \text{ has density}$ $f_k(t) = t^{k-1}e^{-t}/k! \text{ for } t > 0$
$\mathbb{P}\{N_1 = 1 + i\} = q^i p$ for $i = 0, 1, 2, \dots$ negative binomial See HW10 negative binomial as sum of	\leftrightarrow	$T_1 \text{ has density } f_1(t) = e^{-t}$ for $t > 0$ gamma $T_k \text{ has density}$ $f_k(t) = t^{k-1}e^{-t}/k! \text{ for } t > 0$ gamma(k) as sum of
$\mathbb{P}\{N_1 = 1 + i\} = q^i p$ for $i = 0, 1, 2,$ negative binomial See HW10 negative binomial as sum of independent geometrics	\leftrightarrow	$T_1 \text{ has density } f_1(t) = e^{-t}$ for $t > 0$ gamma $T_k \text{ has density}$ $f_k(t) = t^{k-1}e^{-t}/k! \text{ for } t > 0$ gamma(k) as sum of independent exponentials

10.3 Examples for Chapter 10

<10.1> **Example.** Let W_k denote the waiting time to the *k*th point in a Poisson process on $[0, \infty)$ with rate λ . It has a continuous distribution, whose density g_k we can find by an argument similar to the one used in Chapter 7 to find the distribution of an order statistic for a sample from the Uniform(0, 1).

For a given w > 0 and small $\delta > 0$, write M for the number of points landing in the interval [0, w), and N for the number of points landing in the interval $[w, w + \delta]$. From the definition of a Poisson process, M and N are independent random variables with

$$M \sim \text{Poisson}(\lambda w)$$
 AND $N \sim \text{Poisson}(\lambda \delta)$

To have W_k lie in the interval $[w, w + \delta]$ we must have $N \ge 1$. When N = 1, we need exactly k - 1 points to land in [0, w). Thus

$$\mathbb{P}\{w \le W_k \le w + \delta\} = \mathbb{P}\{M = k - 1, N = 1\} + \mathbb{P}\{w \le W_k \le w + \delta, N \ge 2\}.$$

The second term on the right-hand side is of order $o(\delta)$. Independence of M and N lets us factorize the contribution from N = 1 into

$$\mathbb{P}\{M=k-1\}\mathbb{P}\{N=1\} = \frac{e^{-\lambda w}(\lambda w)^{k-1}}{(k-1)!} \frac{e^{-\lambda \delta}(\lambda \delta)^1}{1!}$$
$$= \frac{e^{-\lambda w}\lambda^{k-1}w^{k-1}}{(k-1)!} \left(\lambda \delta + o(\delta)\right)$$

Thus

$$\mathbb{P}\{w \le W_k \le w + \delta\} = \frac{e^{-\lambda w} \lambda^k w^{k-1}}{(k-1)!} \delta + o(\delta),$$

which makes

$$g_k(w) = \frac{e^{-\lambda w} \lambda^k w^{k-1}}{(k-1)!} \mathbf{1}\{w > 0\}$$

the density function for W_k .

<10.2> **Example.** The gamma function is defined for $\alpha > 0$ by

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha - 1} e^{-x} dx.$$

By direct integration, $\Gamma(1) = \int_0^\infty e^{-x} dx = 1$. Also, a change of variable $y = \sqrt{2x}$ gives

$$\begin{split} \Gamma(1/2) &= \int_0^\infty x^{-1/2} e^{-x} dx \\ &= \int_0^\infty \sqrt{2} e^{-y^2/2} dy \\ &= \frac{\sqrt{2}}{2} \frac{\sqrt{2\pi}}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{-y^2/2} dy \\ &= \sqrt{\pi} \qquad \text{cf. integral of } N(0,1) \text{ density.} \end{split}$$

For each $\alpha > 0$, an integration by parts gives

$$\Gamma(\alpha+1) = \int_0^\infty x^\alpha e^{-x} dx$$
$$= \left[-x^\alpha e^{-x}\right]_0^\infty + \alpha \int_0^\infty x^{\alpha-1} e^{-x} dx$$
$$= \alpha \Gamma(\alpha).$$

Repeated appeals to the same formula, for $\alpha > 0$ and each positive integer m less than α , give

$$\Gamma(\alpha+m) = (\alpha+m-1)(\alpha+m-2)\dots(\alpha)\Gamma(\alpha).$$

In particular,

$$\Gamma(k) = (k-1)(k-2)(k-3)\dots(2)(1)\Gamma(1) = (k-1)! \quad \text{for } k = 1, 2, \dots$$

<10.3> **Example.** For parameter value $\alpha > 0$, the gamma(α) distribution is defined by its density

$$f_{\alpha}(t) = \begin{cases} t^{\alpha-1}e^{-t}/\Gamma(\alpha) & \text{for } t > 0\\ 0 & \text{otherwise} \end{cases}$$

If a random variable T has a gamma(α) distribution then, for each positive integer m,

$$\mathbb{E}T^{m} = \int_{0}^{\infty} t^{m} f_{\alpha}(t) dt$$

= $\int_{0}^{\infty} \frac{t^{m} t^{\alpha-1} e^{-t}}{\Gamma(\alpha)} dt$
= $\frac{\Gamma(\alpha + m)}{\Gamma(\alpha)}$
= $(\alpha + m - 1)(\alpha + m - 2) \dots (\alpha)$ by Example <10.2>.

In particular, $\mathbb{E}T = \alpha$ and

$$\operatorname{var}(T) = \mathbb{E}(T^2) - (\mathbb{E}T)^2 = (\alpha + 1)\alpha - \alpha^2 = \alpha.$$

Statistics 241/541 fall 2014 ©David Pollard, 2Nov2014

<10.4> **Example.** Suppose an office receives two different types of inquiry: persons who walk in off the street, and persons who call by telephone. Suppose the two types of arrival are described by independent Poisson processes, with rate λ_w for the walk-ins, and rate λ_c for the callers. What is the distribution of the number of telephone calls received before the first walk-in customer?

Write T for the arrival time of the first walk-in, and let N be the number of calls in [0, T). The time T has a continuous distribution, with the exponential density $f(t) = \lambda_w e^{-\lambda_w t} \mathbf{1}\{t > 0\}$. We need to calculate $\mathbb{P}\{N = i\}$ for $i = 0, 1, 2, \ldots$ Condition on T:

$$\mathbb{P}\{N=i\} = \int_0^\infty \mathbb{P}\{N=i \mid T=t\}f(t) \, dt.$$

The conditional distribution of N is affected by the walk-in process only insofar as that process determines the length of the time interval over which Ncounts. Given T = t, the random variable N has a Poisson $(\lambda_c t)$ conditional distribution. Thus

$$\mathbb{P}\{N=i\} = \int_0^\infty \frac{e^{-\lambda_c t} (\lambda_c t)^i}{i!} \lambda_w e^{-\lambda_w t} dt$$
$$= \lambda_w \frac{\lambda_c^i}{i!} \int_0^\infty \left(\frac{x}{\lambda_c + \lambda_w}\right)^i e^{-x} \frac{dx}{\lambda_c + \lambda_w} \quad \text{putting } x = (\lambda_c + \lambda_w)t$$
$$= \frac{\lambda_w}{\lambda_c + \lambda_w} \left(\frac{\lambda_c}{\lambda_c + \lambda_w}\right)^i \frac{1}{i!} \int_0^\infty x^i e^{-x} dx$$

The 1/i! and the last integral cancel. (Compare with $\Gamma(i+1)$.) Writing p for $\lambda_w/(\lambda_c + \lambda_w)$ we have

$$\mathbb{P}\{N=i\} = p(1-p)^i$$
 for $i = 0, 1, 2, ...$

That is, 1 + N has a geometric (p) distribution. The random variable N has the distribution of the number of tails tossed before the first head, for independent tosses of a coin that lands heads with probability p.

Such a clean result couldn't happen just by accident. HW10 will give you a neater way to explain how the geometric got into the Poisson process. \Box

- <10.5> **Example.** A Poisson process with rate λ on \mathbb{R}^2 is a random mechanism that generates "points" in the plane in such a way that
 - (i) the number of points landing in any region of area A is a random variable with a $Poisson(\lambda A)$ distribution

(ii) the numbers of points landing in disjoint regions are independent random variables.

Suppose mold spores are distributed across the plane as a Poisson process with intensity λ . Around each spore, a circular moldy patch of radius r forms. Let S be some bounded region. Find the expected proportion of the area of S that is covered by mold.



Write $\mathbf{x} = (x, y)$ for the typical point of \mathbb{R}^2 . If B is a subset of \mathbb{R}^2 ,

area of
$$S \cap B = \iint_{\mathbf{x} \in S} \mathbb{I}\{\mathbf{x} \in B\} d\mathbf{x}$$

If B is a random set then

$$\mathbb{E} (\text{area of } S \cap B) = \iint_{\mathbf{x} \in S} \mathbb{EI}\{\mathbf{x} \in B\} \, d\mathbf{x} = \iint_{\mathbf{x} \in S} \mathbb{P}\{\mathbf{x} \in B\} \, d\mathbf{x}.$$

If B denotes the moldy region of the plane,

$$1 - \mathbb{P}\{\mathbf{x} \in B\} = \mathbb{P}\{ \text{ no spores land within a distance } r \text{ of } \mathbf{x} \} \\ = \mathbb{P}\{ \text{ no spores in circle of radius } r \text{ around } \mathbf{x} \} \\ = \exp\left(-\lambda\pi r^2\right).$$

Notice that the probability does not depend on \mathbf{x} . Consequently,

$$\mathbb{E}\left(\text{area of } S \cap B\right) = \iint_{\mathbf{x} \in S} 1 - \exp\left(-\lambda \pi r^2\right) \, d\mathbf{x}$$
$$= \left(1 - \exp\left(-\lambda \pi r^2\right)\right) \times \text{ area of } S$$

The expected value of the proportion of the area of S that is covered by mold is $1 - \exp(-\lambda \pi r^2)$.