

Chapter 9

Poisson processes

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 m tosses per second and decrease the probability of heads to a small p , while keeping the expected number of heads per second fixed at $\lambda = mp$, the number of heads in a t second interval will have approximately a $\text{Bin}(mt, p)$ distribution, which is close to the $\text{Poisson}(\lambda t)$. Also, the numbers of heads tossed during disjoint time intervals will still be independent random variables. In the limit, as $m \rightarrow \infty$, we get an idealization called a POISSON PROCESS.

•Poisson process

<9.1> **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 $\text{Poisson}(\lambda t)$ distribution
- (ii) the numbers of points landing in disjoint (= non-overlapping) intervals are independent random variables. □

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. Obversely, we 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.

Why bother about Poisson processes? When we pass to the idealized mechanism of points generated in continuous time, several awkward artifacts of discrete-time coin tossing disappear. The Examples and Exercises in this Chapter will illustrate the simplifications.

<9.2> **Exercise.** Find the distribution of the time to the k th point in a Poisson process on $[0, \infty)$ with rate λ .

SOLUTION: Denote the time to the k th point by T_k . It has a continuous distribution, which is specified by a density function. For $t > 0$ and small $\delta > 0$,

$$\begin{aligned} \mathbb{P}\{t \leq T_k < t + \delta\} \\ = \mathbb{P}\{\text{exactly } k - 1 \text{ points in } [0, t), \text{ exactly one point in } [t, t + \delta)\} + \text{smaller order terms} \end{aligned}$$

The ‘smaller order terms’ contribute probability less than

$$\mathbb{P}\{2 \text{ or more points in } [t, t + \delta)\} = \mathbb{P}\{\text{Poisson}(\lambda\delta) \geq 2\} = \frac{e^{-\lambda\delta}(\lambda\delta)^2}{2!} + \dots$$

By the independence property (ii) for Poisson processes, the main term factorizes as

$$\begin{aligned} & \mathbb{P}\{\text{exactly } k-1 \text{ points in } [0, t)\} \mathbb{P}\{\text{exactly one point in } [t, t+\delta)\} \\ &= \frac{e^{-\lambda t} (\lambda t)^{k-1}}{(k-1)!} \frac{e^{-\lambda \delta} (\lambda \delta)^1}{1!} \\ &= \frac{e^{-\lambda t} \lambda^k t^{k-1} \delta}{(k-1)!} + \text{smaller order terms} \end{aligned}$$

That is, the distribution of T_k has density

$$\frac{e^{-\lambda t} \lambda^k t^{k-1}}{(k-1)!} \quad \text{for } t > 0.$$

□

Gamma function and gamma density

You should try this calculation at home.

It is easier to remember the form of the density for T_k if one rescales, using an argument similar to the one for the $N(\mu, \sigma^2)$ distribution in Chapter 7, to show that λT_k has a distribution with density

$$\langle 9.3 \rangle \quad \frac{e^{-t} t^{k-1}}{(k-1)!} \quad \text{for } t > 0.$$

This density is called the **GAMMA(k) DENSITY**.

More generally, for each $\alpha > 0$, the density

$$\frac{e^{-t} t^{\alpha-1}}{\Gamma(\alpha)} \quad \text{for } t > 0$$

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

•gamma(α) density

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

The function $\Gamma(\cdot)$ is called the **GAMMA FUNCTION**. Don't confuse the gamma density with the gamma function.

•gamma function

$\langle 9.4 \rangle$ **Example.** The waiting time T_k from Example $\langle 9.2 \rangle$ has expected value

$$\begin{aligned} \mathbb{E}T_k &= \int_0^\infty \frac{t e^{-\lambda t} \lambda^k t^{k-1}}{(k-1)!} dt \\ &= \frac{1}{\lambda} \int_0^\infty \frac{e^{-x} x^k}{(k-1)!} dx \quad \text{putting } x = \lambda t, \text{ (cf. distribution of } \lambda T_k) \\ &= \frac{k}{\lambda} \quad \text{(Use integration by parts.)} \end{aligned}$$

Does it make sense to you that $\mathbb{E}T_k$ should decrease as λ increases?

More generally, for $\alpha > 0$,

$$\begin{aligned} \Gamma(\alpha + 1) &= \int_0^\infty e^{-x} x^\alpha dx \\ &= [-e^{-x} x^\alpha]_0^\infty + \alpha \int_0^\infty e^{-x} x^{\alpha-1} dx \\ &= \alpha \Gamma(\alpha) \end{aligned}$$

In particular,

$$\begin{aligned} \Gamma(k) &= (k-1)\Gamma(k-1) \\ &= (k-1)(k-2)\Gamma(k-2) \\ &= \dots \end{aligned}$$

$$\begin{aligned}
 &= (k-1)(k-2)(k-3)\dots(2)(1)\Gamma(1) \\
 &= (k-1)!
 \end{aligned}$$

because $\Gamma(1) = \int_0^\infty e^{-x} dx = 1$. Compare with the fact that the gamma(k) density in <9.3> integrates to one. \square

Exponential distribution

Specializing the gamma(k) to the case $k = 1$ we get the density

$$e^{-t} \quad \text{for } t > 0,$$

which is called the (standard) EXPONENTIAL DISTRIBUTION. The time to the first point in the Poisson process has density

$$\lambda e^{-\lambda t} \quad \text{for } t > 0,$$

an exponential distribution with expected value $1/\lambda$. Don't confuse the exponential density with the exponential function.

Notice the parallels between the negative binomial distribution (in discrete time) and the gamma distribution (in continuous time). Each distribution corresponds to the waiting time to the k th occurrence of something, for various values of k . Just as (see Problem Sheet 4) the negative binomial can be written as a sum of independent random variables, each with a geometric distribution, so can the gamma(k) be written as a sum of k independent random variables, each with an exponential distribution. The times between points in a Poisson process are independent, exponentially distributed, random variables.

The gamma distribution turns up in a few unexpected places.

<9.5> **Exercise.** If Z has a standard normal distribution, with density $\phi(t) = \exp(-t^2/2)/\sqrt{2\pi}$ for $-\infty < t < \infty$, show that $Z^2/2$ has a gamma(1/2) distribution.

SOLUTION: Write Y for $Z^2/2$. It has a continuous distribution concentrated on the positive half line $(0, \infty)$. For $y > 0$, and $\delta > 0$ small,

$$\begin{aligned}
 \mathbb{P}\{y < Y < y + \delta\} &= \mathbb{P}\{2y < Z^2 < 2y + 2\delta\} \\
 &= \mathbb{P}\{\sqrt{2y} < Z < \sqrt{2y + 2\delta} \text{ or } -\sqrt{2y + 2\delta} < Z < -\sqrt{2y}\}
 \end{aligned}$$

Notice the two contributions; the square function is not one-to-one. Students who memorize and blindly apply transformation formulae quite often overlook such multiple contributions.

Calculus gives a good approximation to the length of the short interval from $\sqrt{2y}$ to $\sqrt{2y + 2\delta}$. Temporarily write $g(y)$ for $\sqrt{2y}$. Then

$$\begin{aligned}
 \sqrt{2y + 2\delta} - \sqrt{2y} &= g(y + \delta) - g(y) \\
 &\approx \delta g'(y) \\
 &= \delta/\sqrt{2y}
 \end{aligned}$$

The interval from $-\sqrt{2y + 2\delta}$ to $-\sqrt{2y}$ has the same length. Using the approximation

$$\mathbb{P}\{x < Z < x + \epsilon\} \approx \epsilon\phi(x) \quad \text{for small } \epsilon > 0,$$

deduce that

$$\begin{aligned}
 \mathbb{P}\{y < Y < y + \delta\} &\approx \frac{\delta}{\sqrt{2y}}\phi(\sqrt{2y}) + \frac{\delta}{\sqrt{2y}}\phi(-\sqrt{2y}) \\
 &= \frac{2\delta}{\sqrt{2y}} \frac{1}{\sqrt{2\pi}} \exp\left(-(\sqrt{2y})^2/2\right) \\
 &= \frac{\delta}{\sqrt{\pi}} y^{-1/2} e^{-y}
 \end{aligned}$$

That is, Y has the distribution with density

$$\frac{1}{\sqrt{\pi}} y^{-1/2} e^{-y} \quad \text{for } y > 0.$$

•exponential distribution

cts. time \leftrightarrow discrete time
 gamma \leftrightarrow neg. binomial
 exponential \leftrightarrow geometric
 For counts:
 Poisson \leftrightarrow Binomial

Compare with the gamma(1/2) density,

$$\frac{y^{1-1/2}e^{-y}}{\Gamma(1/2)} \quad \text{for } y > 0.$$

The distribution of $Z^2/2$ is gamma (1/2), as asserted.

Note: From the fact that the density must integrate to 1, we get a bonus:

$$\Gamma(1/2) = \int_0^\infty y^{1/2-1}e^{-y}dy = \sqrt{\pi}$$

Actually, you could arrive at the same conclusion by making the change of variable $y = x^2/2$ in the integral—which is effectively what we have done in finding the density for the random variable $Z^2/2$. \square

The Poisson process is often used to model the arrivals of customers in a waiting line, or the arrival of telephone calls at an exchange. The underlying idea is that of a large population of potential customers, each of whom acts independently of all the others. The next Example will derive probabilities related to waiting times for Poisson processes of arrivals. As part of the calculations we will need to find probabilities by conditioning on the values of a random variable with a continuous distribution. As before, the trick is first to condition on a discretized approximation to the the variable, and then pass to a limit.

Suppose T has density $f(\cdot)$, and let X be another random variable. If $\delta \approx 0$, then

$$\mathbb{E}(X | T = t) \approx \mathbb{E}(X | t \leq T < t + \delta)$$

Break the whole range for T into small intervals. Rule E4 for expectations gives

$$\mathbb{E}X = \sum_{j=-\infty}^{\infty} \mathbb{E}(X | j\delta \leq T < (j+1)\delta) \mathbb{P}\{j\delta \leq T < (j+1)\delta\}$$

Approximate the last probability by $f(j\delta)\delta$. Temporarily writing $g(t)$ for $\mathbb{E}(X | T = t)$, we then get

$$\sum_{j=-\infty}^{\infty} g(j\delta)f(j\delta)\delta$$

as an approximation to $\mathbb{E}(X)$. Think of the sum as an approximation to $\int_{-\infty}^{\infty} g(t)f(t) dt$. As δ tends to zero, the errors of approximation to both the expectation and the integral tend to zero, leaving (in the limit)

$$\mathbb{E}X = \int \mathbb{E}(X | T = t)f(t) dt \quad \text{for each random variable } X.$$

As a special case, when X is replaced by the indicator function of an event, we get

$$<9.6> \quad \mathbb{P}A = \int \mathbb{P}(A | T = t)f(t) dt \quad \text{for each event } A,$$

Rule E4 for expectations strikes again!

<9.7> **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

$$f(t) = \lambda_w e^{-\lambda_w t} \quad \text{for } t > 0.$$

We need to calculate $\mathbb{P}\{N = i\}$ for $i = 0, 1, 2, \dots$. Invoke formula <9.6>, with A equal to $\{N = i\}$.

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

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 N counts. Given $T = t$, the random variable N has a $\text{Poisson}(\lambda_c t)$ conditional distribution. Thus

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

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

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

Compare with the $\text{geometric}(p)$ 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 nice clean result couldn't happen just by accident. Maybe we don't need all the calculus to arrive at the distribution for N . In fact, the properties of the Poisson distribution and Problem 7.3 show what is going on, as I will now explain.

Consider the process of all inquiries, both walk-ins and calls. In an interval of length t , the total number of inquiries is the sum of a $\text{Poisson}(\lambda_w t)$ distributed random variable and an independent $\text{Poisson}(\lambda_c t)$ distributed random variable; the total has a $\text{Poisson}(\lambda_w t + \lambda_c t)$ distribution. Both walk-ins and calls contribute independent counts to disjoint intervals; the total counts for disjoint intervals are independent random variables. It follows that the process of all arrivals is a Poisson process with rate $\lambda_w + \lambda_c$.

Now consider an interval of length t in which there are X walk-ins and Y calls. From Problem 7.3, given that $X + Y = n$, the conditional distribution of X is $\text{Bin}(n, p)$, where

$$p = \frac{\lambda_w t}{\lambda_w t + \lambda_c t} = \frac{\lambda_w}{\lambda_w + \lambda_c}$$

That is, X has the conditional distribution that would be generated by the following mechanism:

- (1) Generate inquiries as a Poisson process with rate $\lambda_w + \lambda_c$.
- (2) For each inquiry, toss a coin that lands heads with probability $p = \lambda_w/(\lambda_w + \lambda_c)$. For a head, declare the arrival to be a walk-in, for a tail declare it to be a call.

A formal proof that this two-step mechanism does generate a pair of independent Poisson processes, with rates λ_w and λ_c , would involve:

- (1') Prove independence between disjoint intervals. (Easy)
- (2') If step 2 generates X walk-ins and Y calls in an interval of length t , show that

$$\begin{aligned}\mathbb{P}\{X = i, Y = j\} &= \mathbb{P}\{X = i\}\mathbb{P}\{Y = j\} \\ X &\sim \text{Poisson}(\lambda_w t) \quad \text{and} \quad Y \sim \text{Poisson}(\lambda_c t)\end{aligned}$$

You should be able to write out the necessary conditioning argument for 2'.

The two-step mechanism explains the appearance of the geometric distribution in the problem posed at the start of the Example. The classification of each inquiry as either a walk-in or a call is effectively carried out by a sequence of independent coin tosses, with probability p of a head (= a walk-in). The problem asks for the distribution of the number of tails before the first head. The embedding of the inquiries into continuous time is irrelevant. \square